CCS 2021

CONFERENCE ON COMPLEX SYSTEMS

25-29 OCT. 2021

Centre de Congrès
de Lyon – Hall Terreaux

ccs2021.univ-lyon1.fr
@CCS2021L

BOOK OF ABSTRACTS

Conception Morgane Druet - ENS Média Création graphique

Organizing Committee

Co-Chairs

Pierre Borgnat (CNRS, ENS de Lyon)
Márton Karsai (CEU, Vienna)

Advisor

Alain Barrat (CNRS, Aix-Marseille Université)

Program Committee Co-Chairs

Vittorio Loreto (Sapienza University of Rome)
Kristina Lerman (University of Southern California)

Satellite Co-Chairs

Chiara Poletto (iPLESP, INSERM & UPMC)
Nicola Perra (Greenwich University)

Poster Co-Chairs

Bertrand Jouve (CNRS, University of Toulouse 2, France)
Marta Sales Pardo (SEESLAB, Universitat Rovira i Virgili, Spain)

Sponsor Co-Chairs

Hocine Cherifi (University of Burgundy, France)
Guillaume Deffuant (INRAE, France)

Web, Design and Hybrid Co-Chairs

Rémy Cazabet (Lyon 1 University)
David Chavalarias (ISC-PIF, CNRS, France)

Social Media Co-Chairs

Pablo Jensen (CNRS, ENS de Lyon, France)
Carlos Gershenson (UNAM, Mexico)

Local Committee Lyon

Patrice Abry (CNRS)
Diane Barthelemy (ENS de Lyon)
Fatiha Bouchneb (CNRS)
Rémy Cazabet (Univ. Lyon 1)
Stéphane Chrétien (Univ. Lyon 2)
Sara Franceschelli (ENS de Lyon)
Paulo Goncalves (Inria)
Pablo Jensen (CNRS)
Christine Largeron (UJM)
Gaël Poux-Médard (Univ. Lyon 2)
Stéphane Roux (ENS de Lyon)

Steering Committee of the Society of Complex Systems

President of CCS

Guido Caldarelli (until CCS21)

General chairs of CCS

Panos Argyrakis (until CCS 2022)
Cheong Siew Ann (until CCS 2021)

Member until CCS 2021

Alain Barrat
Sandro Meloni
Yamir Moreno, Chair of the CCS SC
Hiroki Sayama

Members until CCS 2023

Maia Angelova
Alberto Antonioni
Iulia Martina Bulai
Fabrizio de Vico Fallani
Tommaso Gili
Vincenzo Nicosia
Jose Ramasco

Contents

Preface

Serving as Program Committee Co-Chairs has been an honour that humbled us with strong responsibility, especially during these unsettling times. We are grateful to the General Chairs for the trust and the support they gave us throughout the journey. We were positively struck by the enthusiasm of all the contributors, despite the evident uncertainties, disruptions and challenges related to the current pandemic. The selection process was not significantly different from the usual ones before the COVID-19 pandemic. What was different, and indeed inspiring, was seeing the dedication and care of authors, on the one hand, and members of the Program Committee, despite the challenging external conditions. From this perspective, a special thanks go to all members of the Program Committee. Our profound gratitude goes to them for their generosity in donating their time and knowledge to go through up to 10-15 contributions each. Thanks to them and all the authors, this Book of Abstracts is extraordinarily rich in high-level scientific contributions, and we hope you will all appreciate the contents presented and the conference itself, both remotely and in Lyon. As always, the latitude of contributions is vast, spanning all areas of Complexity Science, from basic foundations to innovative applications. One of the distinctive features of our community has always been the inclusiveness, this unique capacity to welcome contributions from different communities, still sharing the same technical tools and the same aim for excellence. This conference does not make an exception, and you will find the same openness in all the contributions. We hope you will enjoy it! We wish you a pleasant and productive conference, and thanks again for helping us shape the Complex Systems community.

Kristina Lerman and Vittorio Loreto
Program Chairs of CCS 2021 Lyon

Plenary and Invited Talks

1 Foundations of Complex Systems

Contents

Higher-order percolation processes on multiplex hypergraphs

Hanlin Sun[1] *and Ginestra Bianconi*[1,2] *,(1) School of Mathematical Sciences, Queen Mary University of London, E1 4NS, United Kingdom, hanlin.sun@qmul.ac.uk (2) The Alan Turing Institute, The British Library, 96 Euston Rd, London NW1 2DB, United Kingdom.*

Higher-order interactions are increasingly recognized as a fundamental aspect of complex systems ranging from the brain to social contact networks. Hypergraph as well as simplicial complexes capture the higher-order interactions of complex systems and allow to investigate the relationship between their higher-order structures and their functions.

In this paper [1], we investigate the interplay between structure and the dynamics of higher-order networks and we provide a comprehensive multilayer framework to study higher-order percolation processes on hypergraphs.

The comprehensive multilayer framework capture the rich correlated structure of random multiplex hypergraphs. Random multiplex hypergraphs are generated with layers formed by hyperedges of a given cardinality m and by nodes with given generalized hyperdegree indicating the number of hyperedges of cardinality m incident to them. As such random multiplex hypergraphs can display relevant interlayer generalized hyperdegree correlations. Here we show that the standard percolation is affected by the non-trivial topology of multiplex hypergraphs and by their interlayer correlations that can be tuned to increase or decrease the percolation threshold of the hypergraph. Most importantly our work reveals how the multiplex nature of the multiplex hypergraph ensembles can be exploited to propose higher-order percolation problems displaying a rich interplay between higher-order topology and dynamics and a rich set of phenomena, including discontinuous hybrid transitions and multiple percolation transitions.

In conclusion, we provide a comprehensive view of the higher-order percolation processes on random multiplex hypergraphs, and we hope that this work can generate further interest in the interplay between the structure of higher-order networks and their dynamics.

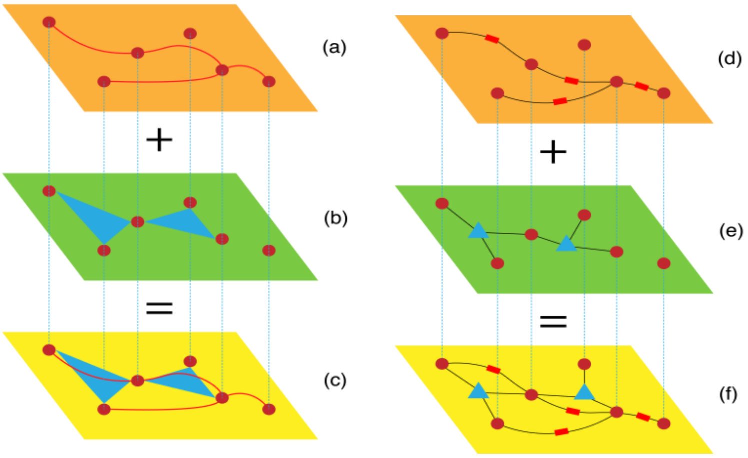


Figure 1: A schematic representation of the multiplex hypergraph with given generalized hyperdegree sequences for hyperedges of cardinality m_1=2 (layer 1) and m_2=3 (layer 2).

[1] Hanlin Sun and Ginestra Bianconi, "Higher-order percolation processes on multiplex hypergraphs," (2021), arXiv: 2104.05457.

Fractal and Seasonal Patterns in Long-Range Time Series

Matthijs Koopmans, *Mercy College, School of Education, 555 Broadway, Dobbs Ferry, NY 10522, USA,*

mkoopmans@mercy.edu

Mandelbrot has argued that the estimation of fractal patterns in time series data reveals adaptive behavior in the system of interest, implying unpredictability. Fractality shows up in time series as self-similar patterns in the variability of the data points, as well as scale invariance in those patterns. Two issues need further attention. First, there are many different methods to estimate fractal patterns in time series, and we need to evaluate the extent to which they yield consistent results. Second, fractal estimates tend to be heavily correlated with non-fractal ones, and therefore, we need to know how capable these methods are distinguishing the two sources of variability. This presentation addresses both of these concerns.

An ARFIMA (p, d, q) simulation was conducted, with one hundred realizations of each of the following types of processes (all of which were stationary): I. White noise ($d = 0$), II. Anti-persistence ($d = -0.30$), III. Persistence ($d = 0.35$), IV. ARIMA (1, 0, 1) at $\varphi = 0.35$ and $\theta = -0.25$, V. Seasonal ARIMA ($\Phi = 0.45$, period = 7) and VI. Seasonal-plus-fractal ARFIMA ($\Phi = 0.45$, period = 7; $d = 0.35$). Simulations I, IV and V are non-fractal and should therefore yield parameter estimates of $d = 0$, while in the other simulations, the differencing parameters should reflect variations in the input values. The response of the following estimation techniques to these six scenarios was evaluated: detrended fluctuation analysis (DFA), re-scaled range analysis (R/S), Higuchi's fractal dimension, smoothed spectral regression, Geweke & Porter Hudak's (GPH) estimator, the spectral periodogram (Sperio) estimate, Whittle's fractional differencing estimate, power spectral density analysis (PSDA), and fractional differencing (FD) with or without ARIMA covariates. Table 1 shows selected results from these analyses.

Table 1

Differencing Parameter Estimates Under Varied Input Conditions (N = 100 Iterations)

Model	Specification (p, d, q) X (P, D, Q)	Input Value of d	Model Estimates of d FD	GPH	Sperio	Whittle
I	(0, 0, 0)	0.00	0.01	-0.00	-0.03	0.00
II	(0, d, 0)	-0.30	-0.40	-0.30	-0.29	-0.41
III	(0, d, 0)	0.35	0.42	0.37	0.34	0.42
IV	(1, 0, 1)	0.00	-0.01	-0.00	-0.02	0.45
V	(0, 0, 0) X $(1, 0, 0)_7$	0.00	0.00	0.13	0.10	0.03
VI	(0, d, 0) X $(1, 0, 0)_7$	0.35	0.25	0.37	0.34	0.33

The table shows that the models investigated here estimate the strength of the fractal dependency, as well as the direction (persistence vs. anti-persistence) quite accurately, although Whittle overestimates fractality in Model IV, while GPH and Sperio overestimate it in Model V. FD results include covariates in Models IV, V and VI. The estimates are near the input values, but underestimate fractality in Model VI. The models that rely on the Hurst exponent ($H = d + 0.50$) for fractal estimation (results not shown here) similarly estimate the direction and strength of the long-range dependency effectively but vary in the extent that they effectively evaluate such dependencies if seasonal and short range autoregressive parameters are added to the simulation conditions. These findings argue for a triangulation of fractal estimates with traditional time series diagnostics (ACF plots, unit root tests, etc.), since both are needed to successfully distinguish fractal patterns from seasonal ones.

Explosive synchronization in interlayer phase-shifted Kuramoto oscillators on multiplex networks

Anil Kumar and S. Jalan* [1] *(1) IIT Indore, Khandwa Road, Indore, India, sarika@iiti.ac.in*

Different methods have been proposed in the past few years to incite explosive synchronization (ES) in Kuramoto phase oscillators. In this work, we show that the introduction of a phase shift α in interlayer coupling terms of a two-layer multiplex network of Kuramoto oscillators can also instigate ES in the layers. As $\alpha \to \frac{\pi}{2}$, ES emerges along with hysteresis. The width of hysteresis depends on the phase shift α, interlayer coupling strength, and natural frequency mismatch between mirror nodes. A mean-field analysis is performed to justify the numerical results. Similar to earlier works, the suppression of synchronization is accountable for the occurrence of ES. The robustness of ES against changes in network topology and natural frequency distribution is tested. Finally, taking a suggestion from the synchronized state of the multiplex networks, we extend the results to classical single networks where some specific links are assigned phase-shifted interactions.

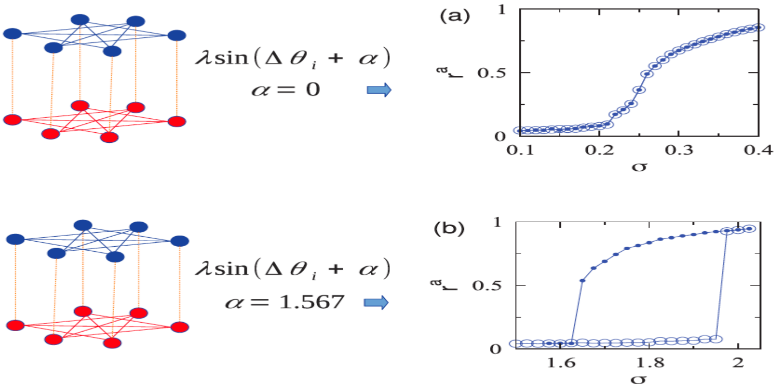


Figure 1: Order parameter as a function of coupling strength (top) second order transition to synchrony without inter-layer phase lag, (bottom) ES with inter-layer phase lag

Acknowledgements (optional)

S.J. acknowledges the Government of India, CSIR grant (No. 25(0293)/18/EMR-II) and the DST grant (No. EMR/2016/001921) for financial support. A.K. acknowledges CSIR, the Government of India for providing SRF.

References

[1] Anil Kumar and Sarika Jalan, *Chaos (Fast track)* 31, 041103 (2021).

Turing Instability In Coupled Nonlinear Relativistic Heat Equations

T. Carletti[1] *and R. Muolo*[1]*. (1) Department of Mathematics & naXys, Namur Institute for Complex Systems, University of Namur, rue Grafé 2, 5000 Namur, Belgium, timoteo.carletti@unamur.be*

Nature is pervaded by regular spatio-temporal patterns, signature of self-organised processes where order emerges from disorder [1]. Very often, the interaction among the coupled microscopic constituting units can be modelled by reaction-diffusion equations governing the time evolution of the concentrations on a complex network [2]. Spatially homogeneous equilibria of the reaction-diffusion system may undergo a symmetry breaking instability eventually driving the system towards a spatially heterogeneous solution, i.e., a pattern, as firstly explained by Alan Turing [3]. Nowadays applications of the *Turing instability* phenomenon go well beyond the original framework of the morphogenesis. It stands for a pillar to explain self-organisation in Nature [4,5], being rooted on an elegant mechanism involving slow diffusing activators and fast diffusing inhibitors.

The reaction-diffusion process at the basis of the Turing mechanism is characterised by an infinite fast propagation and thus it can accurately model the physical phenomenon only in cases of very large diffusivity. To overcome this issue *Cattaneo* proposed to modify the constitutive equation by allowing for a relaxation term (*inertial time*), eventually obtaining a modified diffusion equation responsible for a finite propagation velocity [6].

In this work [7] we extend the conditions for the onset of Turing instability for a reaction-diffusion system defined on top of a complex network and modified according to the Cattaneo recipe, eventually obtaining an *hyperbolic reaction-diffusion system* defined on top of a complex network. We analytically determine the conditions for the onset of the Turing instability and we show that the proposed framework strongly enlarges the parameters region for which Turing instability can emerge. Indeed, Turing instability can set up with a fast activator and slow inhibitor, in contrast with the classical case. Being the patterns solely due to the presence of the inertial times, we propose to call them *inertia-driven Turing instability*.

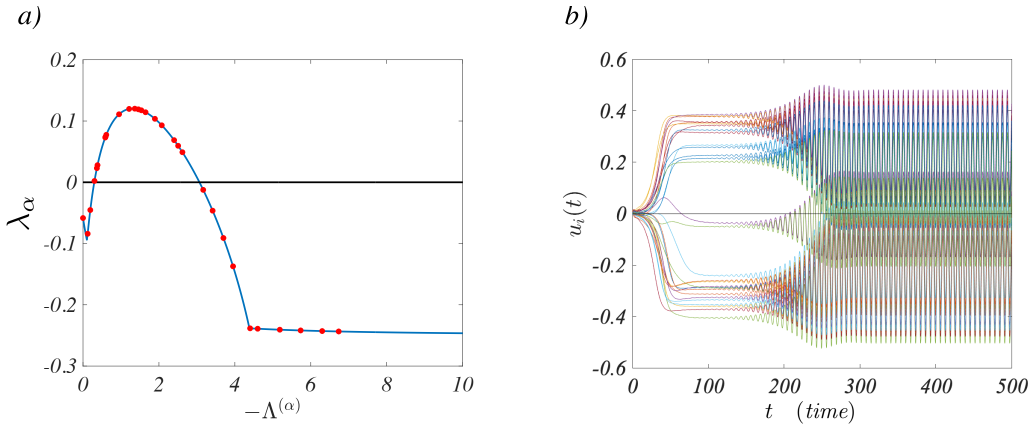


Figure 1: FitzHugh-Nagumo model [8] extended to the framework of hyperbolic reaction-diffusion networked systems. Panel a: the dispersion relation. Panel b: inertia-driven patterns.

References

[1] Nicolis G, Prigogine I. 1977, J. Wiley and Sons.
[2] Pastor-Satorras R, Vespignani A. 2010, Nature Physics 6, 480.
[3] Turing AM. 1952, Phil.Trans.R.Soc.Lond. B 237, 37.
[4] Pecora LM et al. 1997 Chaos 7, 520–543.
[5] Strogatz SH. 2001 Nature 410, 268.
[6] Cattaneo G. 1948 Atti. Sem. Mat. Fis. Univ. Modena p. 83.
[7] Carletti T, Muolo R, preprint https://arxiv.org/abs/2104.04319 (2021).
[8] FitzHugh R. 1961 Biophys. J. 1, 445. Nagumo J, Arimoto S, Yoshizawa S. 1962 Proc. IRE 50, 2061.

Effects of non-normal networks on nonlinear dynamics

R. Muolo[1] *(1) Department of Mathematics and naXys, Université de Namur, Belgium, riccardo.muolo@unamur.be*

In the last 20 years network science has become extremely popular in the framework of dynamical systems and non-linear phenomena. One of the most interesting aspects of such studies is to understand how the network topology, i.e., the fashion in which the nodes are connected to each other, can affect the dynamics: for example it has been shown that it is more likely for a system to synchronize in a small-world network rather than on a regular lattice, or that an asymmetric topology enhances the process of instability leading to Turing patterns [1]. The case examined in our latest studies is that of a non-normal topology. A matrix M is called non-normal when it does not commute with its adjoint, namely $MM^* \neq M^*M$, meaning that there does not exist an orthonormal basis of eigenvector (although a basis may still exist); straightforwardly, a non-normal network is a network whose adjacency matrix is non-normal. It has been shown that real networks, ranging from food-webs to transport networks, are non-normal [2], making this framework extremely attractive for applications. In our investigation, we try to understand qualitatively and quantitatively how non-normality can affect the emergence of instability in reaction-diffusion systems. Such framework is interesting for the study of Turing instability: under certain conditions, a perturbation of the homogeneous stable equilibrium drives the system unstable due to diffusion activation and leads it to an inhomogeneous state, i.e., a pattern. It was recently shown that, when the system is placed on a non-normal network, a stability analysis based on eigenvalues may fail and we can obtain the formation of patterns also with parameters values for which the system is supposed to be stable [3]. Such effect is explainable by the fact that non-normal matrices are more sensible to perturbations, hence, their perturbed spectra (or pseudo-spectra) are significantly larger if compared with their normal analogous [2]. A similar effect has been detected also for the oscillatory case, where now the stability analysis is performed with the Master Stability Function (MSF) [4]. When oscillators are coupled through a non-normal network, synchronization may not be obtained, despite a stable MSF (Figure 1) [5]. The latter results can find many applications, from neuroscience, where synchronization can be an indicator of illnesses such as epilepsy, to power grids, where a lack of synchronization can lead to cascade failures.

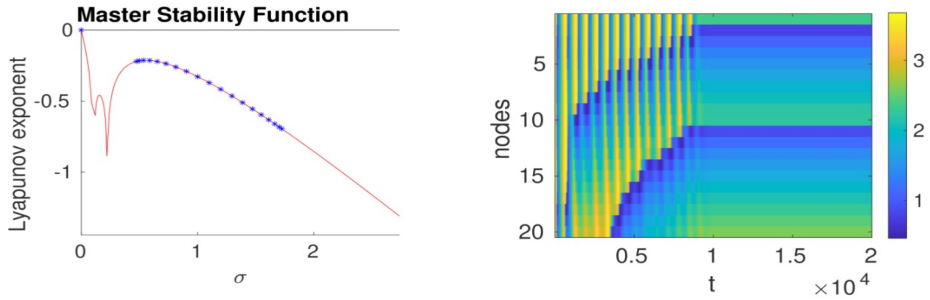


Figure 1: Brusselator model in the limit cycle regime on a non-normal network. Despite a negative (stable) MSF (left), the system loses synchronization and settles in a stationary pattern (right); from [5].

References

[1]Asllani M et al., Nat. Comm. 5 (2014).

[2] Asllani M et al., Sci. Adv. 4, 1–8 (2018) Eaau9403.

[3] Muolo R et al, JTB 480 (2019) 81-91.

[4] Pecora LM, Carroll TL, Phys. Rev. Lett. 80 (1998) 2109.

[5] Muolo R et al., Entropy 26(1) (2021) 36.

Synchronization dynamics in real-world networks

R. Muolo[1], *J. D. O'Brien*[2],*T. Carletti*[1], *J. P. Gleeson*[2], *M. Asllani*[3], *(1) Department of Mathematics and naXys, Université de Namur, Belgium, riccardo.muolo@unamur.be,*
(2) MACSI, Department of Mathematics and Statistics, University of Limerick, Ireland,
(3) School of Mathematics and Statistics, University College Dublin, Ireland

The importance of self-organized collective behaviors in many natural or human-made systems have captured the interest of scientists for a considerable time, leading to the study of stability and, consequently, robustness of the synchronous state. Scholars have investigated, among other parameters, the role that different topologies of the underlying networks have in the stability of the synchronized behavior [1]. One such topological feature is that of non-normality, a structural property known to have strong effects on the dynamics [2], particularly on the occurrence of synchronization. When the coupling network is non-normal, the stable synchronized state is less robust to perturbation and, moreover, a classic spectral approach may fail in predicting the stability [3]. In this work, we propose a thorough analysis of synchronization dynamics in real-world networks. In fact, it was recently shown that non-normality is a feature of such topologies [4] and that an increasing non-normality in such networks is often correlated to the emergence of leader nodes [5], i.e., nodes with only outgoing or incoming links. We show that there is a variety of different phenomena, both caused by the presence of leader nodes and, hence, by the non-normality. For example, in a setting where full synchronization is expected to be robust, a small perturbation permanently destroys the synchronized state. Moreover, as the system moves towards instability (i.e., Lyapunov exponent approaching zero), it exhibits an interesting behaviors: most of the oscillators perfectly synchronize on the main limit cycle, while only a small fraction deviating from the rest of the group (as depicted in Figure 1). We show that such behaviors are ubiquitous in real-world networks and, therefore, universal, which makes our results appealing for future potential applications.

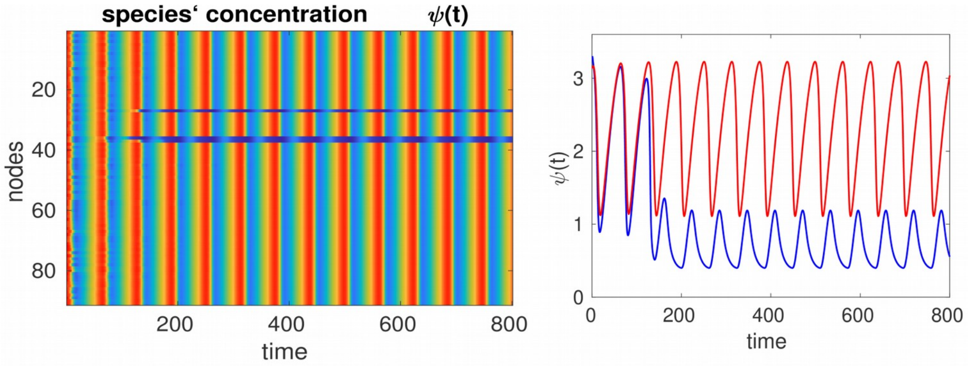


Figure 1: Brusselator model on an empirical neural network. After the perturbation, a small fraction of the nodes synchronizes on a different attractor (blue curve, right panel), while the rest remain synchronized close to the unperturbed solution (red curve).

References

[1] Motter AE, Zhou CS, Kurths J., Europhys. Lett. 69 (2005) 334.
[2] Trefethen LN. Embree M, Princeton University Press, Princeton, NJ (2005).
[3] Muolo R et al., Entropy 26(1) (2021) 36.
[4] Asllani M et al, Sci. Adv. 4, 18. Eaau9403 (2018).
[5] O'Brien JD et al., Phys. Rev. Research (2021) in press.

Classification of Power law Relations in the Intensity Distribution for a Large Class of Self-excited Nonlinear Hawkes Processes

K. Kanazawa[1,2]*, D. Sornette*[3,4]*. (1) The University of Tsukuba, Tsukuba, Ibaraki, Japan, kiyoshi@sk.tsukuba.ac.jp. (2) JST, PRESTO, Kawaguchi, Saitama, Japan. (3) ETH Zurich, Zurich, Switzerland, dsornette@ethz.ch. (4) Southern University of Science and Technology, Shenzhen, China*

Self-excited mechanisms are prevalent in complex systems, such as physical, financial, social, and seismic systems. Consider retweeting dynamics on Twitter as one of the most intuitive examples. If someone retweets a tweet, the tweet's visibility is enhanced, and thus, the tweet is likely to be retweeted shortly. This represents the feedback loop of retweets and induces intermittent behaviour of the retweeting dynamics. Such a dynamical epidemic process can be ubiquitously observed in various complex systems and have attracted the wide interest of physicists and data analysts of social data.

To capture such a self-excited picture, one of the most popular models is the nonlinear Hawkes processes. Hawkes processes are non-Markovian point processes that model the occurrence of events (e.g., retweet on Twitter and earthquake in seismology). The intensity $\lambda(t)$ (the probability per unit time for the occurrence of an event) is defined as

$$\lambda(t) = g\left(\sum_{i=1}^{N(t)} y_i h(t - t_i)\right)$$

with a nonlinear non-negative function $g(v)$, random "marks" $\{y_i\}_{\mathrm{i}}$, total number of events $N(t)$ during $[0, t)$, time of the ith event t_i, and the memory function $h(t)$. Since this Hawkes process is non-Markovian and nonlinear, its theoretical solution has not been established yet, and its theoretical solution is expected to be useful for data analyses of various complex systems.

In this talk, we present our theoretical solution to a wide class of Hawkes processes using field master equations [1, 2]. We first convert the original non-Markovian point dynamics onto a Markovian field dynamics described by a stochastic partial differential equation (SPDE) through the addition of a sufficient number of auxiliary field variables. We then derive the master equation corresponding to the Markovian SPDE as a dynamical equation for the probability density functional. This master equation is solved exactly in the large intensity limit and we classify various power law relations in the asymptotic tail of the steady intensity distributions. For example, we find that (i) a non-universal power law tail holds for the linear Hawkes processes near criticality [1] and that (ii) robust Zipf's law holds for a large class of nonlinear Hawkes processes [2]. This work provides the basic tools to analyse various Hawkes models and further confirm the usefulness of Hawkes processes to model complex systems and their typical power law relations.

Acknowledgements

This work was supported by JST, PRESTO Grant Number JPMJPR20M2, Japan, and the Japan Society for the Promotion of Science KAKENHI (Grand No. 20H05526).

References

[1] K Kanazawa and D Sornette, *Physical Review Letters* **125**, 138301 (2020); *Physical Review Research* **2**, 033442 (2020).

[2] K Kanazawa and D Sornette, arXiv:2102.00242 (2021).

Variance and covariance of distributions on graphs

K. Devriendt[1,2], S. Martin-Gutierrez[3], ***R. Lambiotte***[1,2]*. (1) Mathematical Institute, University of Oxford, UK, renaud.lambiotte@maths.ox.ac.uk. (2) Alan Turing Institute, London, UK. (3) Grupo de Sistemas Complejos, Universidad Politcnica de Madrid, Spain*

We develop a theory to measure the variance and covariance of probability distributions defined on the nodes of a graph, which takes into account the distance between nodes. Our approach generalizes the usual (co)variance to the setting of weighted graphs and retains many of its intuitive and desired properties. Interestingly, we find that a number of famous concepts in graph theory and network science can be reinterpreted in this setting as variances and covariances of particular distributions: we show this correspondence for Kemeny's constant, the Kirchhoff index, network modularity, Markov stability and closeness centrality. Depending on the underlying graph, the measure interpolates between the standard variance defined in Euclidian spaces and the Rényi entropy. As a particular application, we define the maximum-variance problem on graphs with respect to the effective resistance distance, and characterize the solutions to this problem both numerically and theoretically. We show how the maximum-variance distribution can be interpreted as a core-periphery measure, illustrated by the fact that these distributions are supported on the leaf nodes of tree graphs, low-degree nodes in a configuration-like graph and boundary nodes in random geometric graphs. Our theoretical results are supported by a number of experiments on a network of mathematical concepts, where we use the variance and covariance as analytical tools to study the (co-)occurrence of concepts in scientific papers with respect to the (network) relations between these concepts.

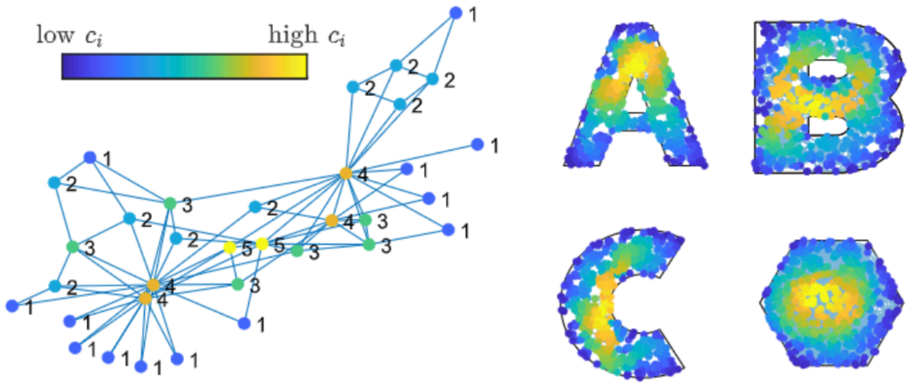


Figure 1: Iterated core-periphery decomposition based on the maximum variance distribution. The method divides the nodes of a graph into layers of increasing centrality or 'coreness', here illustrated for the Karate club network and four random geometric graphs.

Acknowledgements

The acknowledgements should be written here. Thank you for spending your time in reading the instructions. Feel free to contact us for any further information.

References

[1] K. Devriendt, S. Martin-Gutierrez, R. Lambiotte, *arXiv:2008.09155* (2020).

Emerging criticality in heteroclinic dynamics

H. Meyer-Ortmanns[1]*, M. Voit*[1] *(1) Physics and Earth Sciences, Jacobs University Bremen, Germany*
h.ortmanns@jacobs-university.de

Typical features of criticality are a proliferation of the dynamical repertoire, critical slowing down of the dynamics and a high sensitivity to perturbations in the vicinity of a critical parameter value. Such features support the storage, transformation and processing of information; in particular this is assumed and of interest in relation to brain dynamics. A suitable framework to describe the transient dynamics in the brain is heteroclinic dynamics, among other applications it may be applied to cognitive processes of the neural system. We consider various heteroclinic networks and zoom into the dynamics that emerges right at a subset of bifurcation points. The observed features of criticality qualify these points and their immediate vicinity as candidates for working points in systems which store and transfer information. As a novel feature we observe the emergence of a heteroclinic cycle between three manifolds that are densely covered by periodic orbits. Which periodic orbits are selected sensitively depends on the initial conditions due to hidden conservation laws. From a more general perspective, our results add to the observation that emergent features may arise at borderlines. Here these borderlines separate different dynamical regimes in heteroclinic dynamics.

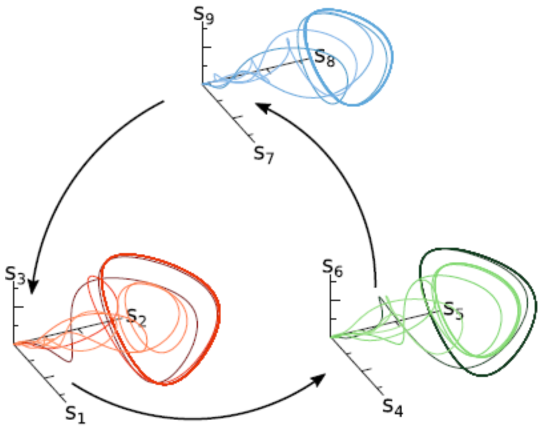


Figure 1: Sketch of the dynamics of a heteroclinic cycle between manifolds (from [1])

Acknowledgements

We thank the German Research Foundation (DFG, Contract No. ME-1332/28-1) for financial support.

References

[1] M. Voit and H. Meyer-Ortmanns, Phys. Rev. Res. **2**, 043097 (2020).

[2] M. Voit and H. Meyer-Ortmanns, Phys. Rev. Res. **1**, 023008 (2019).

Potential-driven random walks on complex networks

B. Benigni,[1, 2] R. Gallotti,[1] and M. De Domenico[1]
[1] *CoMuNe Lab, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo (TN), Italy*
[2] *University of Trento, Department of Information Engineering and Computer Science, Via Sommarive, 9, 38123 Povo,Trento TN*

Routing information on networks is mostly based on shortest paths or random walk protocols. On the one hand, shortest path protocols assume a global knowledge of the network, driving agents to follow routes of minimum cost among the available ones. However, this approach is not always feasible because of lack of global knowledge about the system or for computational issues. On the other hand, random walk protocols rely only on local knowledge and, consequently, could take a long time to reach a destination, providing a misrepresentation of real-world navigation processes. In real applications, efficient navigation strategies must be able to exploit local knowledge and, at the same time, non-local one which, in addition, can vary in time and space because of constraints (e.g. restricted traffic areas in urban networks). Moreover, from the analysis of information flows, non-trivial patterns emerge which are not captured either by broadcasting, such as in random walks, or geodesic routing, such as shortest paths. In fact, alternative models between those extreme protocols are still eluding us.

In this work we define a walk which tends to minimize distances (like a shortest path; a global feature) while keeping some flexibility in random exploration (like a random walk; a local feature). In particular we propose a new process, the *potential-driven random walk*, that effectively interpolates between shortest path and random walk routing protocols thanks to the presence of a potential field defined on the top of the network, acting at each node. In practice, we propose a new type of biased random walk [1, 2] with bias encoded by the potential, which can take different functional forms depending on background assumptions. By considering a generalized Coulomb dependence on the distance on destination(s), we show that it is possible to interpolate between random walk and geodesic routing in a simple and effective way (see Fig.1) . Here, we consider the combination of two fundamental variables: node's degree and the topological distance from the potential. By assuming that the potential has a gravitational-like form and it is centered in a node ℓ, a walker in node i will randomly select a neighbor j while i) being attracted by ℓ inversely proportional to their relative topological distance; and ii) being biased by the presence of network hubs in its neighborhood. We propose a biased transition matrix defined by

$$T_{ij}(\ell) = \frac{c_j(\ell) W_{ij}}{\sum\limits_{k=1}^{N} c_k(\ell) W_{ik}} \tag{1}$$

where $c_j(\ell) \propto d_{j\ell}^{\gamma} k_j^{\beta}$ and $d_{j\ell}$ is the topological distance between j and the potential ℓ and k_j is the degree of node j, γ and β are the bias parameters. We present results concerning the steady-state and the mean first-passage time statistics of this new process, and show its feasibility to model real-world movements in spatial networks, with applications to empirical animal and human movements. From a methodological perspective, our potential-driven random walks open the doors to a broad spectrum of new analytical tools, ranging from new random-walk centralities to geometry induced by potential-driven network processes [3].

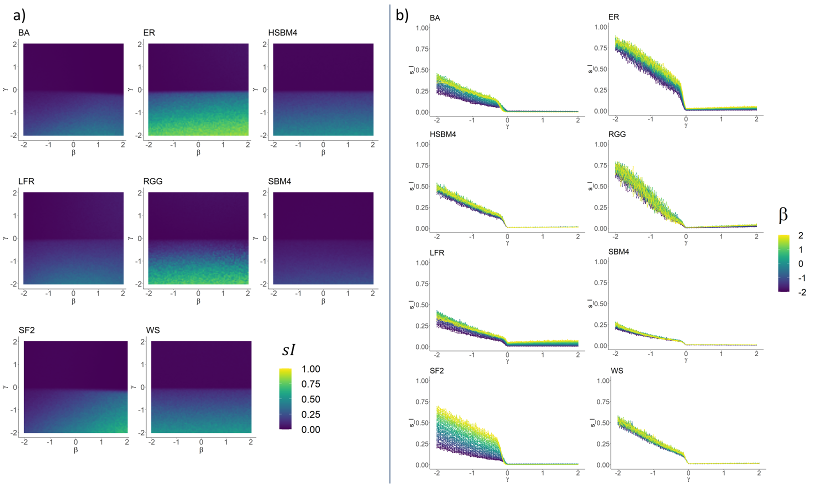


Figure 1. **Straightness index on networks**, sI, quantifying to what extent the potential-driven random walk is close to shortest path ($sI = 1$) or to the random walk ($sI = 0$). In a) we report the values of sI for all the considered network topologies of 256 nodes: Barabási-Albert (BA), Erdös-RÃ©nyi (ER), Hierarchical Stochastic Block Model (HSBM4) Lancichinetti-Fortunato-Radicchi(LFR), Random Geometric Graph (RGG), Stochastic Block Model (SBM4), Scale-Free (SF2) and Watts-Strogatz (WS). The values of bias parameter γ and β are the axes of the heatmap, while tile color encodes the values of sI. In b) we express the values of sI as a function of γ, color line encodes the value of β.

---

[1] M. Bonaventura, V. Nicosia, and V. Latora, Physical Review E **89**, 012803 (2014).
[2] J. Gómez-Gardenes and V. Latora, Physical Review E **78**, 065102 (2008).
[3] B. Benigni, R. Gallotti, and M. De Domenico, Submitted (2021).

An empirical study on classical and community-aware centrality measures in complex networks

S. Rajeh[1]*, M. Savonnet*[1]*, E. Leclercq*[1]*, H. Cherifi*[1]*.* [1]*Laboratoire D'Informatique de Bourgogne, University of Burgundy, 21000 - Dijon, France. E-mail: stephany.rajeh@u-bourgogne.fr.*

Community structure is a ubiquitous feature in natural and artificial systems. Identifying key nodes is a fundamental task to speed up or mitigate any diffusive processes in these systems. Centrality measures aim to do so by selecting a small set of critical nodes. Classical centrality measures are agnostic to community structure, while community-aware centrality measures exploit this property. Several works study the relationship between classical centrality measures, but the relationship between classical and community-aware centrality measures is almost unexplored. In this work [1], we answer two questions: (1) How do classical and community-aware centrality measures relate? (2) What is the influence of the network topology on their relationship? We perform an analysis involving a set of popular classical and community-aware centrality measures on synthetic and real-world networks. Community-aware centrality measures can be classified into two groups. The local ones are relying mainly on the intra-community links. The global ones exploit the connections bridging the communities. First, we calculate the Kendall's Tau correlation between all pair of classical and community-aware centrality measures on a set of synthetic LFR networks with controllable parameters (community structure strength (μ), degree distribution exponent (γ), and community size distribution exponent (θ)). Results show that the main feature driving the correlation variation between classical and community-aware centrality measures is the community structure strength. Furthermore, local community-aware centrality measures correlate with classical ones in networks with a strong community structure, while correlation is weak for global community-aware centrality measures. One observes the opposite behavior in networks with a weak community structure. Second, we calculate the Kendall's Tau correlation using a set of 50 real-world networks from various domains. We perform a linear regression analysis to assess the relationship between the networks' macroscopic and mesoscopic features and the mean correlation value. Results show that the community structure strength and transitivity are the most significant features driving the correlation between classical and community-aware centrality measures. These findings open a new perspective on designing community-aware centrality measures tailored to the network topology.

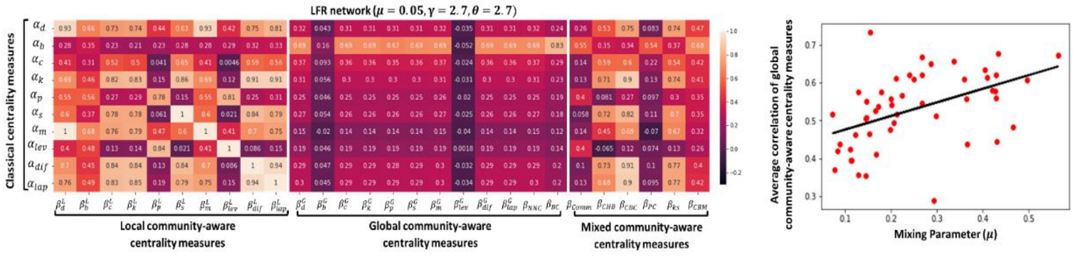


Figure 1: Left figure: Kendall's Tau correlation for all possible combinations between classical (x-axis) and community-aware (y-axis) centrality measures. Right figure: Linear regression between the mean correlation values of the global community-aware centrality measures and the mixing parameter.

References

[1] Rajeh, S., Savonnet, M., Leclercq, E. et al. Characterizing the interactions between classical and community-aware centrality measures in complex networks. Sci Rep 11, 10088 (2021).

Percolation Theory Of Self-Exciting Temporal Processes

D. Notamuzi[1]*, C. Castellano*[2]*, A. Flammini*[1]*, D. Mazzilli*[1]*, F. Radicchi*[1]*. (1) Center for Complex Networks and Systems Research, Luddy School of Informatics Computing and Engineering, 919 E. 10th St., Bloomington, Indiana (47408), dnotarm@iu.edu (2) Istituto dei Sistemi Complessi (ISC-CNR), Via dei Taurini 19 , I-00185 Roma, Italy.*

We investigate how the properties of inhomogeneous patterns of activity, appearing in many natural and social phenomena, depend on the temporal resolution used to define individual bursts of activity. To this end, we consider time series of microscopic events produced by a self-exciting Hawkes process, and leverage a percolation framework to study the formation of macroscopic bursts of activity as a function of the resolution parameter. We find that the very same process may result, when analyzed at different resolutions, in different distributions of avalanche size and duration, which are understood in terms of the competition between the one-dimensional percolation and the branching process universality classes. Pure regimes are observed at specific values of the resolution parameter, corresponding to the critical points of the percolation phase diagram. A regime of crossover characterized by a mixture of the two universal behaviors is observed in a wide region of the phase diagram. Such hybrid scaling appears to be a likely outcome for an analysis of the time series based on a reasonably chosen, but not precisely adjusted, value of the resolution parameter.

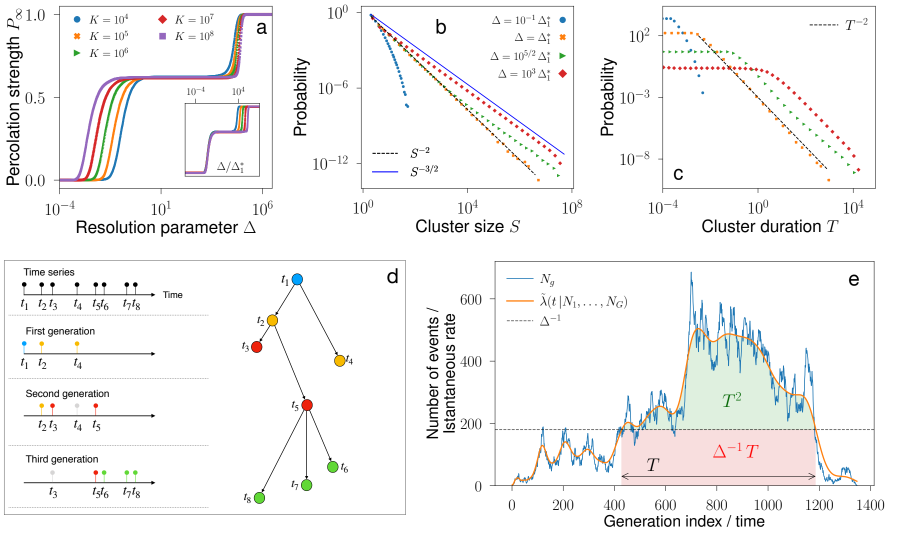


Figure 1: Summary figure of our main results. Given a time series with K events we define a cluster of size S as a sequence of S events separated in time no more than Δ, the resolution parameter. The duration T of the cluster is the time lag between its last and first event. (a) The double percolation transition associated to pure scaling regimes. (b-c) Distribution of S and T, displaying pure scaling and a neat crossover (b). (d) The latent geometrical structure of the Hawkes process allows to understand the crossover as a threshold phenomenon (e).

References

[1] D. Notarmuzi, C. Castellano, A. Flammini, D. Mazzilli, F. Radicchi, Phys. Rev. E 103, L020302.

Behaviour Recognition Algorithms In Temporal Graphs

H. Hourcade[1,2], *F.Hamissi*[1], *C. Miachon*[2], *B.M. Bui-Xuan*[1]. *(1) LIP6, Paris, France, hugo.hourcade@lesfurets.com. (2) Courtanet, Paris, France.*

A temporal graph G is a sequence of static graphs indexed by a set of integers T representing time instants. Such a graph can be used to represent the interactions between given agents and the evolution of those interactions through time. Algorithms able to discern behaviours of those agents are therefore useful to characterize populations. We focus on two behaviour recognition problems, and the algorithmic solutions we devised for each one. The first one is called Δ-Module[1]. Given Δ an integer, a group of vertices A form a Δ-module if they have the same neighbourhood outside of A for at least Δ consecutive time instants. A solution to this problem permits the recognition of similarities of behaviour between multiple agents. The second one is called TemporalSubpathIsomorphism[2]. Given P a temporal graph called the pattern graph, P is subgraph isomorphic to G if there exists a bijection $f\colon V(P) \rightarrow V(G)$ such that, given t_0 a time instant of G, there is an edge between u and v, vertices of P, at time instant t if and only if there exists an edge between $f(u)$ and $f(v)$ at time instant $t+t_0$ in G. A solution to this problem permits the recognition of behavioural pattern between various agents. We present exact algorithms solving specific cases of both these problems.

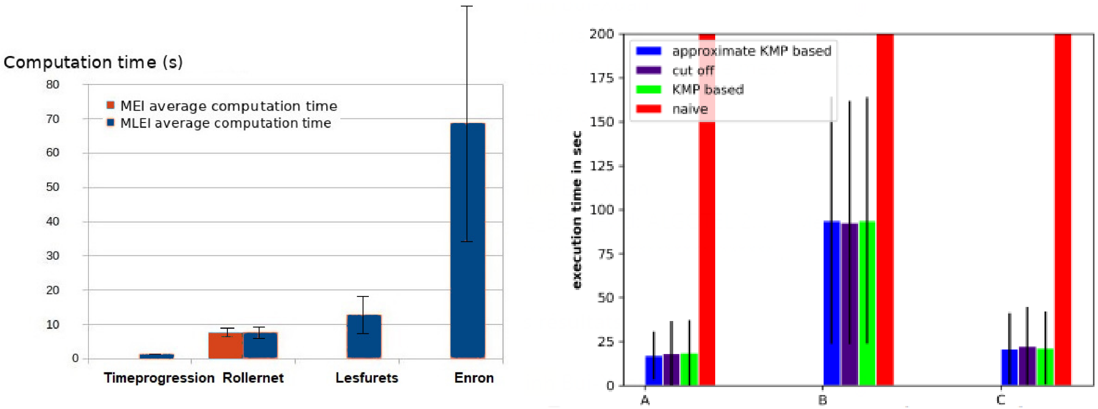


Figure 1: (Left) Computation time of Δ-modules of size 2. (Right) Computation time solving temporal subgraph isomorphism when pattern graph P is a temporal path.

Our implementations are available at *https://github.com/DaemonFire/deltaModules* and at *https://github.com/reihan35/temporal_subpath_isomorphism*.

Acknowledgements

We are grateful to Emmanuel Chailloux for helpful discussions and pointers.

References

[1] B-M. Bui-Xuan, H. Hourcade, C. Miachon, *Computing Temporal Twins in Time Logarithmic in History Length*, International Conference on Complex Networks and Their Applications IX (2020) 651-663.

[2] F. Hamissi, *Subpath stream isomorphism in tree streams,* Master's thesis, Sorbonne Université, 2020.

Random Matrix Analysis Of Multiplex Networks

Tanu Raghav[1]*, Sarika Jalan*[1]*. (1) Complex Systems Lab, Discipline of Physics, Indian Institute of Technology Indore, Khandwa Road, Simrol, Indore, 453552, India, tanu@iiti.ac.in.*

We investigate the spectra of adjacency matrices of multiplex networks under random matrix theory (RMT) framework. Through extensive numerical experiments, we demonstrate that upon multiplexing two random networks[1], the spectra of the combined multiplex network exhibit superposition of two Gaussian orthogonal ensemble[2] (GOE)s for very small multiplexing strength followed by a smooth transition to the GOE statistics with an increase in the multiplexing strength. Interestingly, randomness in the connection architecture, introduced by random rewiring to 1D lattice[1], of at least one layer may govern nearest neighbor spacing distribution (NNSD)[2] of the entire multiplex network, and in fact, can drive to a transition from the Poisson to the GOE statistics or vice versa. Notably, this transition transpires for a very small number of the random rewiring corresponding to the small-world transition[1]. Ergo, only one layer being represented by the small-world network is enough to yield GOE statistics for the entire multiplex network. Spectra of adjacency matrices of underlying interaction networks have been contemplated to be related with dynamical behavior of the corresponding complex systems, the investigations presented here have implications in achieving better structural and dynamical control to the systems represented by multiplex networks against structural perturbation in only one of the layers.

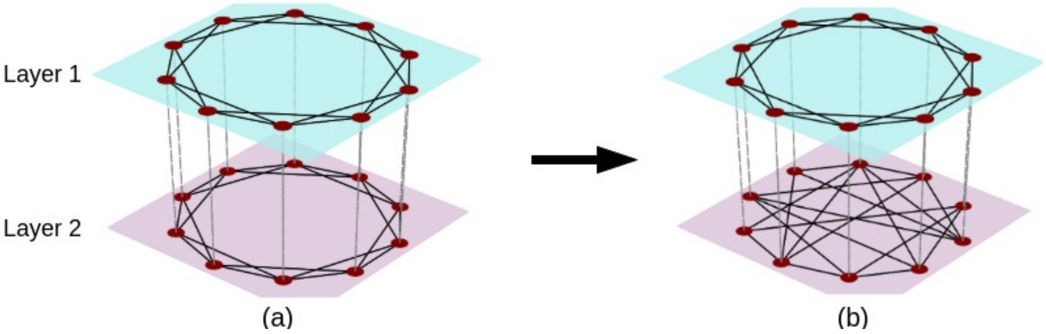


Figure 1: Schematic diagram representing a multiplex network consisting of two layers. (a) Both layers are represented by 1D lattice. (b) Edges are rewired in the second layer.

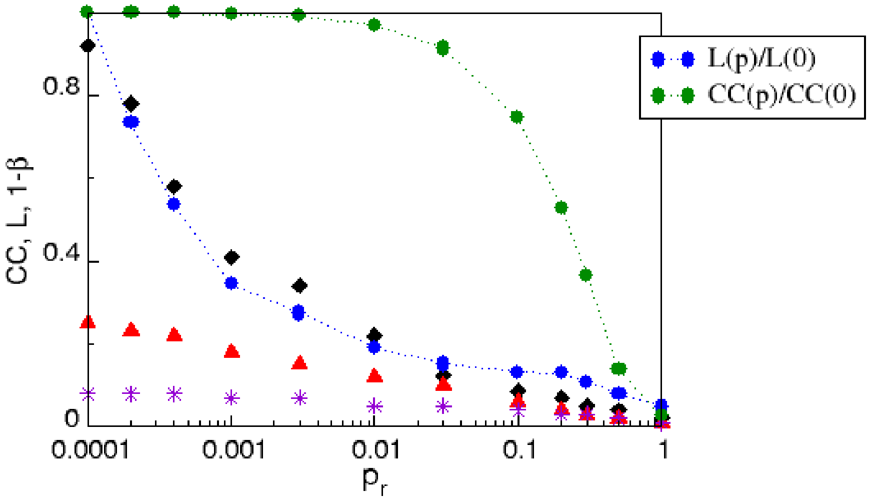


Figure 2: The shifted Brody parameter[2] (1 − β) of the three multiplex networks (with one layer fixed as 1D, random and small-world network while the second layer undergoes transition from 1D lattice to random network using Watts-Strogatz algorithm[1]) is compared with the two network parameters, normalized characteristic path length (L) and normalized clustering coefficient (CC) of a rewired network.

Acknowledgements (optional)

SJ and Tanu acknowledge Govt. of India, BRNS Grant No.37(3)/14/11/2018-BRNS/37131 for financial support and JRF fellowship, respectively. We thank Shashi Chandralal Srivastava (VECC) for interesting discussions throughout the project, and Ranveer Singh (IIT Indore) for useful suggestions.

References (optional)

[1] R. Albert, A. L. Barabási, Reviews of Modern Physics, 74 (2002) 47.

[2] M. L. Mehta. Random Matrices. Academic Press. 1991

Multi-level and Multi-scale Reconstruction of Knowledge Dynamics with Phylomemies

D. Chavalarias[1,2], ***Q. Lobbé***[1], *A. Delanoë*[1]. *(1) Complexe Systems Institute of Paris IdF, 113 Rue Nationale, Paris, France, quentin.lobbe@iscpif.fr. (2) CAMS - Centre d'Analyse et de Mathématique sociales*

In this presentation, we address the question of “drawing science” by taking advantage of the massive digitization of scientific production, and focusing on its body of knowledge. We demonstrate how we can reconstruct, from the massive digital traces of science, a reasonably precise and concise approximation of its dynamical structures that can be grasped by the human mind [1] and explored interactively [2]. For this purpose, we formalize the notion of level and scale of knowledge dynamics as complex systems and we introduce a new formal definition for phylomemetic networks as dynamical reconstruction of knowledge dynamics. We propose a new reconstruction algorithm for phylomemetic networks that outperforms previous ones and demonstrate how this approach also makes it possible to define a new temporal clustering on dynamical graphs. Finally, we show in case studies that this approach produces representations of knowledge dynamics close to the ones that can be obtained by synthesizing the points of view of experts on a given domain.

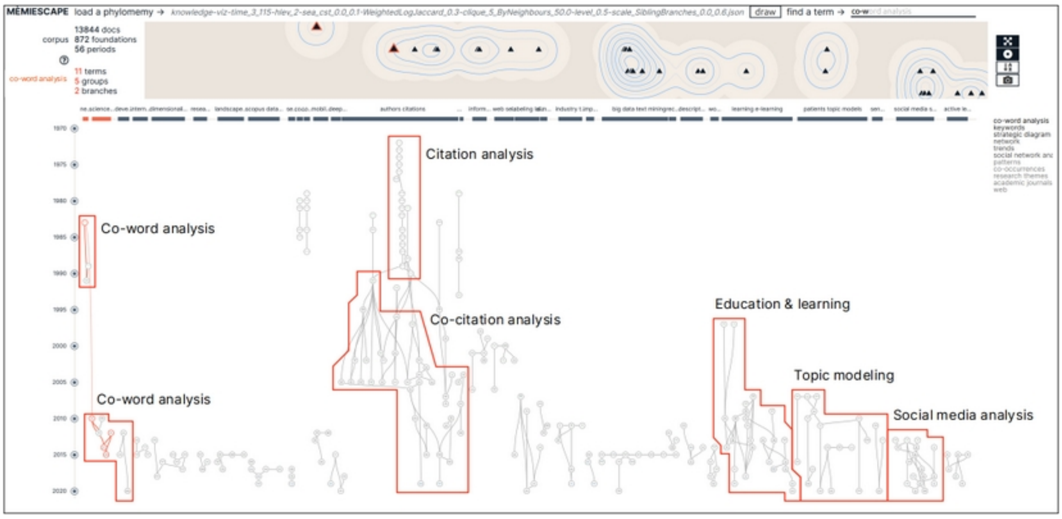


Figure 1: Phylomemy of the literature related to science and knowledge mapping, extracted from the Web of Science.

References

[1] David Chavalarias, Quentin Lobbe, Alexandre Delanoë. Draw me Science - multi-level and multi-scale reconstruction of knowledge dynamics with phylomemies (forthcoming), https://halshs.archives-ouvertes.fr/hal-03180347/

[2] Quentin Lobbé, Alexandre Delanoë, David Chavalarias. Exploring, browsing and interacting with multi-scale structures of knowledge (forthcoming), https://hal.archives-ouvertes.fr/hal-03181233v1

Emergence of functional communities in non-linear random walks on complex networks

Jean-François de Kemmeter[1], *Timoteo Carletti*[1], *Malbor Asllani*[2]. *(1) naXys, Namur Institute for Complex Systems, University of Namur, rue Grafé 2, B 5000 Namur, Belgium, jean-francois.dekemmeter@unamur.be. (2) School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland.*

Network science has emerged in recent years as a powerful discipline that studies the structure of the interactions of complex systems and their influence on the collective dynamics [1]. A current trend has been the investigation of the role that compelling nonlinear interactions have in the emergence of novel exotic states in the dynamical behavior of such systems. An irrefutable example of a network-based dynamical process that spans applications from neuroscience to finance, from human mobility to animal foraging, is that of random walks [2]. A modern perspective of such dynamics considers the biased random walk processes where single individuals are confined to move in the network following nodes' pre-assigned attributes. Such an idea has been recently extended and redefined to take into account the finite-carrying capacities of the hosting node [3,4]. Based on this concept, in this work, we have introduced the novel class of *self-recruitment random walks* where the movement of the walkers is proportionally oriented by the density of individuals in the target nodes and simultaneously constrained by their finite sizes. Interestingly enough, we observe that functional communities, namely connected subgraphs of filled nodes surrounded by a sea of empty ones, spontaneously emerge (see fig. 1). We predict the emergence of these states and show that the process can exhibit multistability, at odds with other random walk processes. A striking characterizing feature of the emerging communities is that the fraction of vacant nodes can occur in a quantized manner (see fig. 1). Such results can potentially shed light on a better understanding of the emergence of vacant niches and their segregation in ecology and the formation of urban prairies in the cities' evolution.

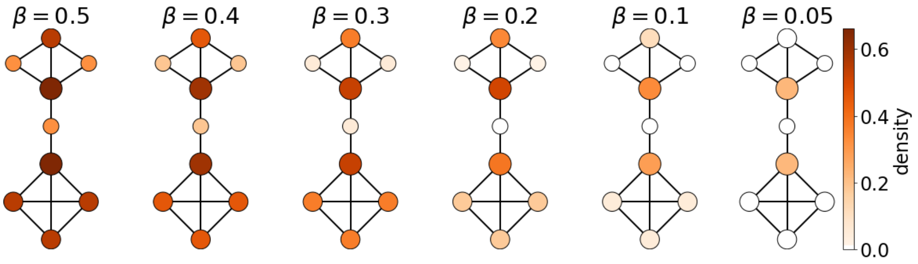


Figure 1: Emergence of functional communities separated by empty nodes (white) as the average node crowding β is reduced. The node density is displayed (see colorbar).

References

[1] Newman, M. (2018). *Networks*. Oxford university press.

[2] Barrat, A., Barthelemy, M., & Vespignani, A. (2008). *Dynamical processes on complex networks*. Cambridge university press.

[3] Asllani, M., Carletti, T., Di Patti, F., Fanelli, D., & Piazza, F. (2018). Hopping in the crowd to unveil network topology. *Physical review letters*, *120*(15), 158301.

[4] Carletti, T., Asllani, M., Fanelli, D., & Latora, V. (2020). Nonlinear walkers and efficient exploration of congested networks. *Physical Review Research*, *2*(3), 033012.

Complexity and Quantum Information Science

***M. Swan**. University College London, Malet Place, London WC1E 6BT UK, swanm@blockchainstudies.org.*

Quantum computing is a fast-moving physics and mathematics-driven research frontier whose implementation is expanding the range of approaches to complexity. This work provides a comprehensive articulation of the emerging canon of quantum information science techniques and how they are related to the study of complex systems. Quantum computing is currently available (including as one free minute per month) and easily accessed via cloud services (the current idea is not that quantum desktops and quantum smartphones with supercooled quantum information processors (QPUs) would become the norm).

Why quantum? Scalability, energy-efficiency, and natural three-dimensional modeling

The main argument for quantum computing is that it provides a more capacious next-generation computational paradigm that is more scalable and energy-efficient than current methods of classical computing and supercomputing. Further, as Nobel physicist Feynman noted, the most efficient computational models are those built in the structure of reality [1]. Quantum computing is the use of engineered quantum systems to perform computation. Such quantum machines are literally comprised of quantum objects (atoms, ions, photons) that are manipulated through logic gates to perform computation. This approach delivers a mode of computing that more naturally corresponds to the three-dimensional structure of atomic reality.

Quantum information science canon

The primary quantum information science methods harness the quantum mechanical properties of superposition, entanglement (entropy), and phase transition [2], all of which feature in the understanding of complex systems. One surprise is that quantum computing can be used to model classical systems to reveal otherwise inaccessible correlations in the data or underlying system (the correlations are expressed as complex-valued wavefunction amplitudes). As a field, quantum information science can be organized into four topics in the areas of describing, evolving, transmitting, and deploying the underlying system, with various constituent methods (Table 1). Some of the central themes relating quantum information science to complexity are investigating systems on the basis of ordered and disordered regimes, entropic uncertainty relations, and many-body dynamics that impact criticality and multi-phase behavior.

Describe the System	Evolve the System	Transmit the System	Deploy the System
Information content Entropy Fourier duals Subsystem correlations Information compression Uncertainty relations	Quantum probability Quantum dynamics Quantum master equations Operatorial relations Many-body localization Unitary operations	Quantum teleportation Quantum walks Error correction Quantum channels Entanglement generation End-to-end qubits	Quantum machine learning: neural networks, tensor networks, kernel Wavefunctions Simulation Benchmarking

Table 1: Quantum Information Science Techniques

References

[1] R.P. Feynman, *Int J Theor Physics*, 21 (1982) 6:467-88.

[2] M. Swan, R.P. dos Santos, F. Witte, *Quantum Computing: Physics, Blockchains, and Deep Learning Smart Networks*, London: World Scientific, 2020.

Long Wavelength Coherence in Well-Connected Power Grids

J. Fritzsch[1] *(1) Department of Quantum Matter Physics, University of Geneva, 1211 Geneva, Switzerland, julian.fritzsch@etu.unige.ch*

Power grids are standardly modeled as coupled oscillators where each oscillator represents an electric generator or a load. It has been observed that groups of generators in one area sometimes start to oscillate coherently against a geographically separated group of generators in a different area. Such long wavelength coherent effects are well-understood in networks that consist of weakly-coupled well-defined areas. However, these oscillations have also been observed in large scale power grids. Fig. 1 shows the coherent response of the generators and loads in Spain and Portugal to a noisy perturbation on power production in Greece.

These Inter-area oscillations are problematic because they can lead to grid instabilities. It is therefore of utmost importance to understand better their properties, how they emerge, and how they can be controlled.

In this talk, using perturbation theory we show that such inter-area modes generically occur in systems of coupled oscillators, even in large-scale well-connected graphs, and that they are triggered by any perturbation affecting either area. We finally discuss how to geographically distribute primary control to optimally damp them.

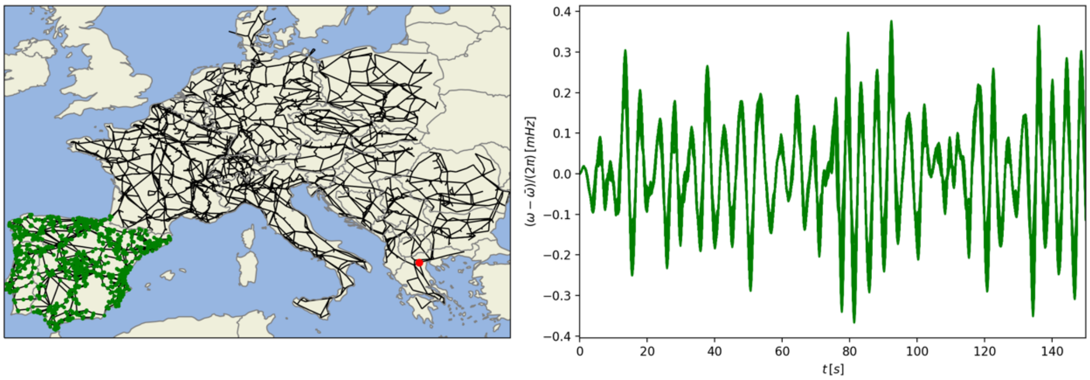


Figure 1: Reaction of the generators and loads in Spain and Portugal to a noisy perturbation on the power production in Greece. The node marked by the red dot denotes the perturbed generator.

Balanced Triad Formation explained by Homophily-induced Dyadic Interactions

Tuan Pham[1,2], Jan Kobel[1,2], Rudolf Hanel[1,2], and Stefan Thurner[1,2,3]

[1] Section for the Science of Complex Systems, Medical University of Vienna, Vienna, Austria,
[2] Complexity Science Hub Vienna, Josefstadterstrasse 39, A-1090 Vienna, Austria
[3] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA
`tuan.pham@meduniwien.ac.at`

The evolution of social (signed) triads towards the so-called balanced state with either three or one positive link [1] often results in the formation of clusters of positively-linked agents. We argue that such cluster formation can also emerge from *dyadic* interactions if these interactions are determined by homophily between agents. Without knowledge of triads in their neighbourhoods, the agents modify their opinions so as to minimize a social tension defined via the weighted sum of opinion overlaps with friends and opinion discordance with enemies. The model exhibits a transition from unbalanced- to balanced society at a critical temperature which depends on the number of *independent* binary opinions, G, the mean degree, K, and the relative strength of positive interactions to that of negative ones, α. As α exceeds 1/2, a transition between the absorbing states with different fractions of balanced triads occurs.

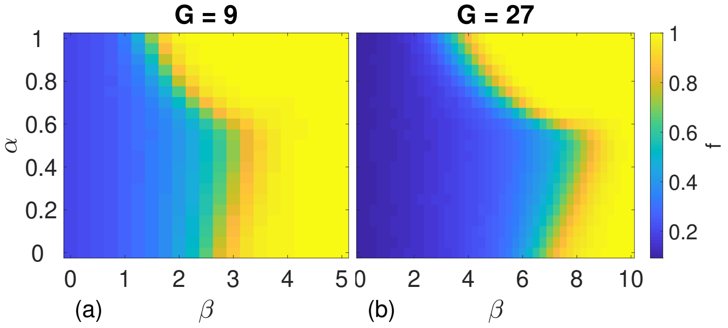


Fig. 1. Order parameter, $f = (n_+ - n_-)/(n_+ + n_-)$, where n_+ and n_- are the numbers of balanced (i.e. with an even number of negative links) and unbalanced ((i.e. with an odd number of negative links) triads, respectively, as a function of the relative strength of positive interactions to that of negative ones, α, and the inverse temperature, β, for two different numbers of independent opinions: (a) $G = 9$ and (b) $G = 27$. Results are averaged over 100 runs on ring networks with mean degree $K = 8$ and the number of nodes $N = 1000$.

References

1. Heider, F.: Attitudes and Cognitive Organization. Journal of Psychology 21, 107 (1946)

Voluntary safety commitments provide an escape from over-regulation in AI development

T.A. Han[1]*, T. Lenaerts*[2]*, F.C. Santos*[3]*, L.M. Pereira*[4] *(1) Teesside University, Middlesbrough, UK, Email: T.Han@tees.ac.uk; (2) Université Libre de Bruxelles, Brussels, Belgium (3) Universidade de Lisboa, Lisbon, Portugal; (4) Universidade Nova de Lisboa, Caparica, Portugal.*

With the introduction of Artificial Intelligence (AI) and related technologies in our daily lives, fear and anxiety about their misuse, as well as the hidden biases in their creation, have led to a demand for regulation to address such issues. Yet, blindly regulating an innovation process that is not well understood may stifle this process and reduce benefits that society might gain from the generated technology, even under the best of intentions. Starting from a baseline game-theoretical model that captures the complex ecology of choices associated with a race for domain supremacy using AI technology [1], we show how socially unwanted outcomes may be produced when sanctioning is applied unconditionally to risk-taking, i.e., potentially unsafe behaviours [2]. As an alternative to resolve the detrimental effect of over-regulation, we propose a voluntary commitment approach, wherein technologists have the freedom of choice between independently pursuing their course of actions or else establishing binding agreements to act safely, with sanctioning of those that do not abide to what they have pledged [3]. Overall, this work reveals for the first time how voluntary commitments, with sanctions either by peers or by an institution, leads to socially beneficial outcomes in all scenarios that can be envisaged in the short-term race towards domain supremacy through AI technology. These results are directly relevant for the design of governance and regulatory policies that aim to ensure an ethical and responsible AI technology development process.

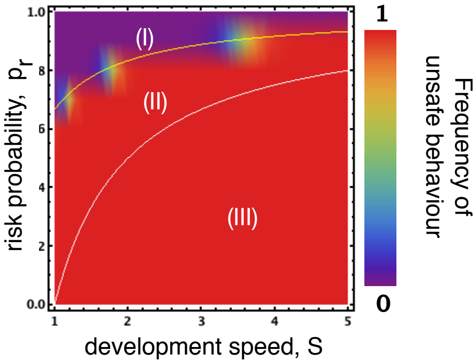


Figure 1: Frequency of unsafe behaviour as a function of development speed and the disaster risk, in absence of incentives [1]. In regions (**I**) and (**III**), safe and unsafe/innovation, respectively, are the preferred collective outcome also selected by natural selection, thus no regulation being required. Region (**II**) requires regulation as safe behaviour is preferred but not the one selected. This talk explores how to promote safe behaviour in this dilemma region by using voluntary commitments, without engendering over-regulation in other regions [3].

References

[1] Han, T. A., Pereira, L. M., Santos, F. C., & Lenaerts, T. (2020). To Regulate or Not: A Social Dynamics Analysis of an Idealised AI Race. *J. Artificial Intell. Research*, *69*, 881-921.

[2] Han, T. A., Pereira, L. M., Lenaerts, T., & Santos, F. C. (2021). Mediating artificial intelligence developments through negative and positive incentives. *PloS one*, *16*(1), e0244592.

[3] Han, T. A., Lenaerts, T., Santos, F. C., & Pereira, L. M. (2021). Voluntary safety commitments provide an escape from over-regulation in AI development. arXiv preprint arXiv:2104.03741.

Estimation of static community memberships from multiplex and temporal network data

K. Avrachenkov[1], ***M. Dreveton***[1], *L. Leskelä*[2]. *(1) Inria Sophia-Antipolis, 2004 Route des Lucioles, 06902 Valbonne France, maximilien.dreveton@inria.fr, (2) Department of Mathematics and Systems Analysis, Aalto University, Otakaari 1, 02015 Espoo Finland*

Data sets in many application domains consist of pairwise interactions observed over time. Pair interactions are often characterized by the type of interacting objects, and a set of objects with a common type is called a community. *Community recovery* or *clustering* is the unsupervised task of inferring the community memberships from the observed pair interactions.

As an information-theoretic benchmark, we study data sets generated by a homogeneous block model where the pairwise interactions take values in a measurable space *S*. The interactions within a block are distributed according to f_{in} fin, and interactions between blocks according to f_{out} . We derive a lower on the expected number of misclassified nodes made by *any* clustering algorithm. This naturally extends the recent results of [3] to a non-asymptotic setting which makes no regularity assumptions on fin, fout nor the underlying space *S* of interaction types. In particular, it can be used when the number of snapshots *T* grows with the number of nodes *N*.

Then, we also show that the lower bound can be achieved by an *ad-hoc* algorithm. If we denote by D the Renyi-divergence between f_{in} and f_{out} , then for same-size clusters *almost exact recovery* (the expected proportion of misclassified nodes going to zero) is possible if $ND \ll 1$, and is impossible otherwise. This provides a natural extension to known results in block models [4]. We later apply those results to dynamic networks where the interaction kernel has a Markov structure, generalizing the results of [2] on a multiplex SBM with independent layers.

In the second part, we propose several clustering algorithms, both offline and online, which fully utilize the temporal nature of the observed data. The online procedure for community detection is based on likelihood estimation. It runs in $O(K N^2 T)$, and can be further reduced in sparse networks to *O(K m T)* where *m* is the average number of edges in a layer.

For further details and proofs, we refer to our preprint [1]. The simulation code is available at https://github.com/maximiliendreveton/MarkovSBM.

References (optional)

[1] Avrachenkov, K., Dreveton, M., & Leskelä, L. (2020). Estimation of Static Community Memberships from Temporal Network Data. *arXiv preprint arXiv:2008.04790*.

[2] Paul, S., & Chen, Y. (2016). Consistent community detection in multi-relational data through restricted multi-layer stochastic blockmodel. *Electronic Journal of Statistics*, *10*(2), 3807-3870.

[3] Xu, M., Jog, V., & Loh, P. L. (2020). Optimal rates for community estimation in the weighted stochastic block model. *The Annals of Statistics*, *48*(1), 183-204.

[4] Zhang, A. Y., & Zhou, H. H. (2016). Minimax rates of community detection in stochastic block models. *The Annals of Statistics*, *44*(5), 2252-2280.

Flow Network Problems on the *n*-torus with Asymmetric Couplings

R. Delabays*, S. Jafarpour, and F. Bullo.*
UC Santa Barbara, Santa Barbara, CA 93106, United States.
robindelabays@ucsb.edu.

The last decades have seen the increase of the scientific interest in systems of diffusively coupled dynamical agents, and in particular networks of nonlinearly coupled oscillators. Indeed, such models are well-suited to describe the behavior of a wide class of physical and virtual systems, ranging from power grids to synapses in the brain, to name just two of them. The literature shows, as can be expected, that the network structure underlying the pairwise couplings between oscillators can have strong impact on the overall behavior of the system as a whole.

In some of these systems, it is natural to consider that the coupling between two individual agents is symmetric, either for realism, when the coupling is truly symmetric, or for mathematical convenience, when such an approximation is not too wrong. Nevertheless, some systems of coupled oscillators have intrinsically asymmetric coupling. In power grids for instance, dissipation implies that the amount of power that is withdrawn at one end of a line is larger than the power received at the other end.

In this talk, we will explore how far the lessons learned for symmetric couplings can be extended to systems with asymmetric couplings. First of all, we will need to set up a framework that allows us to conveniently rephrase the problem. We will then see that, under rather fair assumptions, we are able to fully characterize the relation between network structures and steady states of networks of coupled oscillators.

Physical study of collective movements in complex environments: A Case Applied to Fish Schools

R. Larrieu[1], *C. Sebastiao*[2], *C. Graff* [3], *P. Peyla*[1], *A. Dupont*[1]. *(1) LIPhy, 140 Rue de la Physique, Grenoble, France, renaud.larrieu@univ-grenoble-alpes.fr , (2) École Normale Supérieure de Lyon, France, (3) Laboratoire de Psychologie et Neurocognition, Grenoble, France*

The collective behavior of some animals is a fascinating example of coordinated movement that spontaneously forms on a large scale despite the limited communication capacities between individuals. Different phases of collective movement (polarized, rotational, disordered) have been characterized by different dynamic models of out-of-equilibrium systems. Previously, we have shown that social interactions improve rheotaxis sensitivity [1] using a numerical modeling. Here, we propose to experimentally test the interactions between individuals and the strategies of non-hierarchical groups of small fish (*Paracheirodon innesi*) in a set of complex and controlled conditions. First, we study the passage of fish through a constriction that the group cannot cross at once (Figure 1). By applying a stress, we can create an important clustering of fish close to the circular constriction and study the statistics of passage. In a different experiment, we focus on the global polarization of the group placed in a dense array of obstacles. This work is only at its beginning and is part of the long term with in parallel links that will be made with numerical modelling.

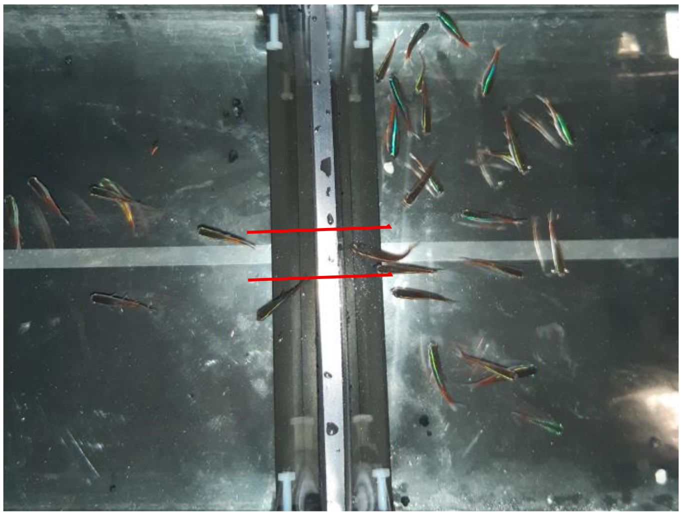


Figure 1: Fish passage experiment in a circular slot under light stress. The red lines represent the opening of the slot.

References

[1] R. Larrieu, C. Quilliet, A. Dupont, P. Peyla. Collective orientation of an immobile fish school, effect on rheotaxis, *Physical Review E*, 2 (2021) 103.

Directed Percolation in Temporal Networks

***A. Badie-Modiri*[1]**, *A. K. Rizi*[1], *M. Karsai*[2,3], *M. Kivelä*[1]. *(1) Aalto University School of Science, Espoo, Finland, arash.badiemodiri@aalto.fi. (2) Department of Network and Data Science, Central European University, 1100 Vienna, Austria. (3) Alfréd Rényi Institute of Mathematics, 1053 Budapest, Hungary*

Percolation theory forms the backbone of our understanding of connectivity in networks. It connects phenomena such as spreading processes to the broader concept of critical phenomena and allows theoretically grounded ways to characterise phase transitions. This theory, however, has not been extended yet to temporal networks, leaving us without an understanding of connectivity and related dynamical phenomena in the same way as it has been established on static networks. Here we show that connectivity, as induced by a limited waiting-time process, can be interpreted as directed percolation. We employ recent theoretical advances in temporal event graphs [1] to define analogues quantities and derive analytical solutions for different types of synthetic temporal networks. Further, we use a new estimation algorithm to compute reachability [2] to confirm our analytical result empirically. We find that critical exponents of the system are the same as the mean-field exponents of directed percolation (see Figure 1 for examples of the finite-size scaling analysis).

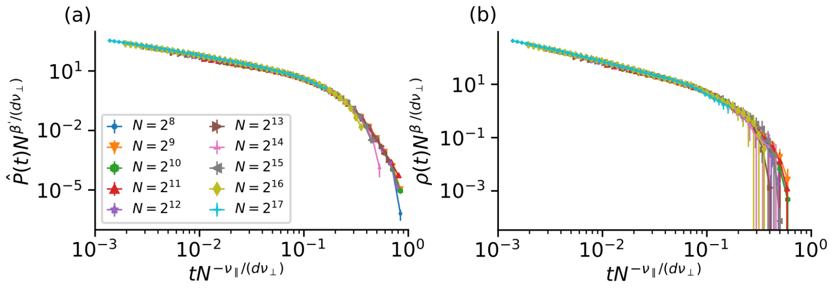


Figure 1: Finite-size scaling collapse of probability of survival in a single source spreading process (a) and occupation density with a fully occupied initial condition (b) on a temporal network constructed from random 9-regular networks of size N with a Poisson process $\lambda=1$ governing activations. The collapse of the trajectories confirm that the system can be explained using the mean-field exponents of directed percolation $\beta=\beta'=\nu_{//}=1$ and $\nu_\perp=1/2$.

Many dynamical processes can be interpreted as a limited waiting-time process: in an epidemic, infected individuals recover after some time, on social networks people may forget a piece of news after some times, or in transportation passengers may not wait forever for a connection. If such processes are present in temporal networks, their behaviour as characterised by an order parameter (e.g. probability of survival of a spreading process) goes through a directed percolation phase-transition controlled by the maximum allowed waiting time. Beyond the theoretical foundation, we demonstrate this mapping on real-world temporal networks related to communication and transport by measuring the thermodynamic quantities adapted to temporal networks which reveals reachability transitions and their time-scales in these systems. Our results allow the use of machinery developed for directed percolation to analyse temporal networks and provides a new tool for the understanding and precise modelling of reachability and spreading phenomena.

References

[1] M Kivelä, J Cambe, J Saramäki, M Karsai, *Scientific reports* 8.1 (2018): 1-9.

[2] A Badie-Modiri, M Karsai, M Kivelä, *Physical Review E*, 101.5 (2020): 052303.

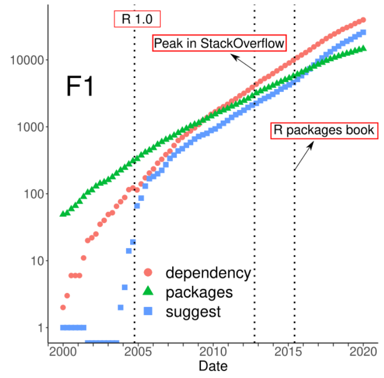
The R Network evolution: characterization of a collaborative network of software

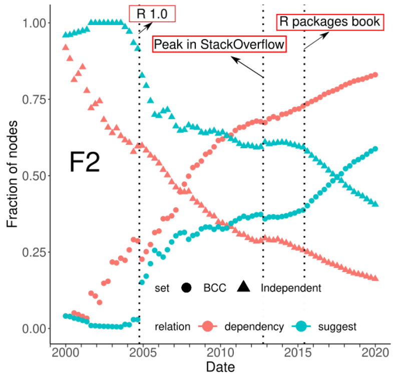
A. Salgado[1], *I. Caridi*[1]. *(1) Instituto de Cálculo UBA-CONICET, Buenos Aires, Argentina, asalgado@df.uba.ar*

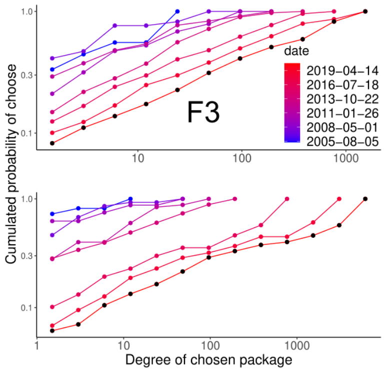
We present the evolution of the *R* project's collaborative software network, based on the official site data where *R*-users upload their package contributions (Comprehensive R Archive Network, CRAN). *R* is a free open source programming language created in the late '90s from *S*, a statistical analysis-oriented software. Nowadays, *R* packages cover a wide range of topics, including statistics, economics, machine-learning, biology, ecology, physics, geography, and many others. A package consists of a set of functions designed to provide a specific tool and can take advantage of other available packages. CRAN has grown from a few packages in 2000 to more than 14,000 interrelated packages today (F1). This growth results from an ever-increasing collaboration between developers from different fields and the emergence of a community of *R*-users worldwide, which develop packages and build other resources on *R* like books, tutorials, FAQ, etc.

We study the evolution of this package network made up of three elements: the nodes (packages) and two types of links among them (dependencies and suggestions) (F1). Dependencies are software-directed relationships, similar to paper citations, where a package depends on another if it uses its functionalities to work. Suggestions indicate the existence of examples or tutorials using both packages. While the network was sparse in its origin, the number of connections surpasses the number of packages nowadays, making the network mostly connected. We characterize the network's changes in time using macroscopic measures accounting for the size of the biggest connected component (BCC), and the fraction of independent packages (*BCC* and *independent* packages in F2). We analyze how the relationships of both types are distributed between packages, finding long-tailed distributions. The process of package addition is characterized using preferential attachment (PA) notions, finding sub and superlinear PA in dependencies and suggestions, respectively (F3, upper and lower panel, respectively). The number of relationships (dependencies and suggestions) of a new package is described with a unique distribution, using the BCC in each network as its unique parameter.

We associate the network increase in connectivity with external "R events" like book publications, journal creations, and interest in StackOverflow. Changes in the network connectivity can be related to both changes in the CRAN guidelines for development and the availability of resources for learning *R*. F1 and F2 show as an example the publication of R 1.0, most *R*-related questions in StackOverflow, and the publication of the *R packages* book.







A Multi-Type Branching Process Method for Modelling Complex Contagion on Clustered Networks

***L. A. Keating**, J. P. Gleeson, D. J.P. O'Sullivan. MACSI, Department of Mathematics and Statistics, University of Limerick, Limerick V94 T9PX, Ireland, leah.keating@ul.ie.*

Online social networks such as Twitter, Facebook, Instagram and TikTok serve as media for the spread of information between their users. We are interested in developing models for this information diffusion to gain a greater understanding of its drivers. Some models for the spread of online behaviour and information assume that the information behaves similarly to a virus, where infection is equally likely after each exposure, these dynamics are known as a simple contagion. In a simple contagion, the exposures are independent of each other. However, online adoption of some behaviour and content has been empirically observed to be more likely after multiple exposures from their network neighbours [1-2], the exposures are not independent of each other, we refer to this as a complex contagion. Analytically tractable descriptions of complex contagions have been developed for continuous-time dynamics. These extend mean-field and pair approximation methods to account for clustering in the network topologies [3]; however, no such analogous treatments for discrete-time cascade processes exist using branching processes. We describe a novel definition of complex contagion adoption dynamics and show how to construct multi-type branching processes which account for clustering on networks. We achieve this by tracking the evolution of a cascade via different classes of clique motifs which account for the different numbers of active, inactive and removed nodes, as shown in Fig. 1. This description allows for extensive Monte Carlo simulations (which are faster than network-based simulations), accurate analytical calculation of cascade sizes, determination of critical behaviour and other quantities of interest.

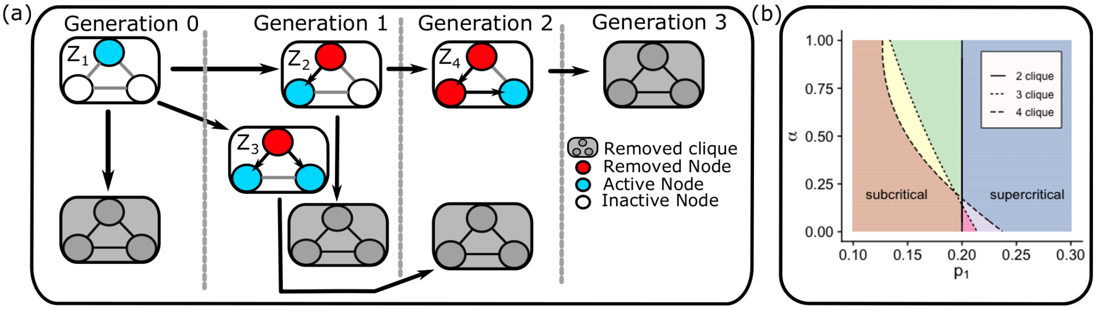


Figure 1: (a) The possible motifs in a network composed with 3-cliques. Active nodes are shown in red, inactive nodes are white and removed nodes are grey. The arrows represent the direction of the diffusion. (b) Critical behaviour for k-regular clique network we consider.

References

[1] D. Centola, The spread of behavior in an online social network experiment, Science 329, 1194 (2010).
[2] D. M. Romero, B. Meeder, and J. Kleinberg, Differences in the mechanics of information diffusion across topics: idioms, political hashtags, and complex contagion on twitter, in Proceedings of the 20th international conference on World wide web (2011) pp. 695–704.
[3] D. J. P. O'Sullivan, G. J. O'Keeffe, P. G. Fennell, and J. P. Gleeson, Mathematical modeling of complex contagion on clustered networks, Frontiers in Physics 3,10.3389/fphy.2015.00071 (2015).

Clarifying the Link Between Degree Entropies, Degree-Degree Correlations and Robustness of Complex Networks

***C. Jones*[1]*, K. Wiesner*[2]*. (1) University of Bristol, School of Mathematics, Bristol, United Kingdom, christopher.jones@bristol.ac.uk. (2) University of Potsdam, Potsdam Institute for Climate Impact Research, Potsdam, Germany*

It has previously been hypothesized that a complex network's robustness against node removal may be accounted for by the heterogeneity of its degree distribution. Wang et al [1] claimed that maximizing a scale-free network's heterogeneity as measured by the degree distribution entropy is equivalent to optimizing the network's robustness against random failures. This is on the condition that the number of nodes and edges in the network remain constant. First, we expand upon their analysis by considering the relationship between degree distribution entropy and robustness for a wide variety of networks, demonstrating that degree distribution entropy sets a lower bound for a network's robustness when no other constraints are applied (Figure 1). Second, we compare network robustness to heterogeneity measured by the remaining degree entropy for networks with the same numbers of nodes and edges. We show that a network's robustness against random attacks is optimized by maximizing the remaining degree entropy as opposed to the degree distribution entropy. Finally, we examine how another entropic measure, mutual information, relates to robustness for networks with different configurations. We demonstrate that mutual information, which measures degree-degree correlations, does not necessarily vary as robustness changes when a network's configuration is altered, showing that degree-degree correlations are not indicative of network robustness.

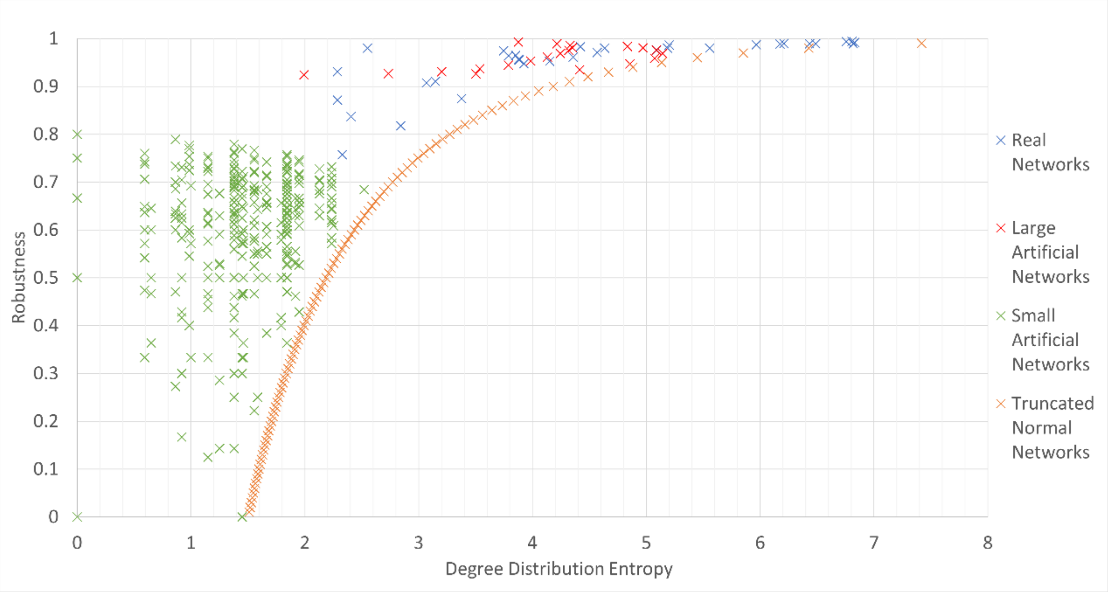


Figure 1: Robustness against Degree Distribution Entropy for various networks

References

[1] B Wang, H Tang, Z Xiu. *Physica A: Statistical Mechanics and its Applications,*

363(2):591-596, 2006.

Self-Organization In Stellar Evolution: Size-Complexity and Rate-Complexity Rules

G. Y. Georgiev[1,2,3]*, T. Butler*[1]*. (1) Department of Physics, Assumption University, 500 Salisbury St., Worcester, MA 01609, USA, ggeorgie@assumption.edu. (2) Department of Physics, Worcester Polytechnic Institute, 100 Institute Road, Worcester, MA 01609, USA. (3) Department of Physics, Tufts University, 4 Colby St., Medford, MA 02155, USA.*

Complexity Theory is highly interdisciplinary, therefore any regularities must hold on all levels of organization, independent on the nature of the system. An open question in science is how complex systems self-organize to produce emergent structures and properties. We explore the quantity-quality transition in natural systems. This is to say that the properties of a system depend on its size. More recently, this has been termed the size-complexity rule, which means that to increase their size, systems must increase their complexity, and that to increase their complexity they must grow in size. This rule goes under different names in different disciplines and systems of different nature, such as the area-speciation rule, economies of scale, scaling relations (allometric) in biology and for cities, and many others. We apply the size-complexity rule to stars to compare them with other complex systems in order to find universal patterns of self-organization independent of the substrate. As a measure of complexity of a star, we are using the degree of grouping of nucleons into atoms, which reduces nucleon entropy, increases the variety of elements, and changes the structure of the star [1]. As a measure of the rate of self-organization and increase in complexity, we are using the average rate of nucleosynthesis over the lifetime of stars as a function of their mass. As seen in our previous work [2], complexity, using action efficiency, is in power law proportionality of all other characteristics of a complex system, including its size. Here we find that, as for the other systems studied, the complexity of stars and the rates of increase of complexity are in a power law proportionality with their size. The bigger a system is, the higher its level of complexity is and the faster the rate of its increase - despite differing explosion energies and initial metallicities from simulations and data, which confirms the size-complexity and size-rate of increase of complexity rules and our model. The complexity and the rate of increase of complexity of a star are also in a power-law relation with each other.

Acknowledgements

We would like to thank the Assumption University Honor's Program and the Department of Biological and Physical Sciences for financial support, as well as Assumption students Thanh Vu, Simon Trcka and Lamberto Qako for help with data collection. We would also like to thank Assumption students Nicole Bramlitt and Rebecca Gilchrist for long discussions and for proofreading the paper, and to Prof. Steven Theroux, Prof. Michael Levin, Clement Vidal and Claudio Flores Martinez for valuable feedback.

References

[1] T. Butler, G.Y. Georgiev, *Self-Organization In Stellar Evolution: Size-Complexity Rule,* in print “Efficiency in Complex Systems” Springer Proceedings in Complexity, CCS17, Cancun, Mexico. Editors: G.Y. Georgiev, M. Shokrollahi-Far, 2020.

[2] G.Y. Georgiev, K. Henry, T. Bates, E. Gombos, A. Casey, H. Lee, M. Daly, and A. Vinod, *Mechanism of organization increase in complex systems*, Complexity, 21(2), 18-28, Wiley, DOI: 10.1002/cplx.21574 7/25 (2015).

Disentangling Internal Interactions From Noisy Environments Through Mutual Information

G. Nicoletti[1], *D. M. Busiello*[2], *S. Suweis*[1], *S. Azaele*[1,] *A. Maritan*[1].
(1) Laboratory of Interdisciplinary Physics, Department of Physicsand Astronomy "G. Galilei", University of Padova, Padova, Italy giorgio.nicoletti.1@phd.unipd.it. (2) Institute of Physics, École Polytechnique Fédérale de Lausanne - EPFL, 1015 Lausanne, Switzerland

Real-world systems are characterized by complex interactions of their many internal degrees of freedom and, at the same time, they live in ever-changing and noisy environments. It is often the case that we have only access to the internal states of the system and we know nothing about the environment, and the net effect of this ignorance about the environmental states is the appearance of new, effective couplings. These two sources of interaction are usually tightly entangled, and singling outthe dominant one starting from data is a long-standing challenge. Here, we show that disentangling internal interactions from the effective ones stemming from noisy environments can be a feasible task in a general linearized stochastic model. Indeed, by computing the mutual information under suitable limits, only the contribution from the internal degrees of freedom survives and the one from the environmental, effective couplings vanishes. Our results elucidate how and how much it is possible to infer from the ever-growing wealth of data on real-world interacting systems.

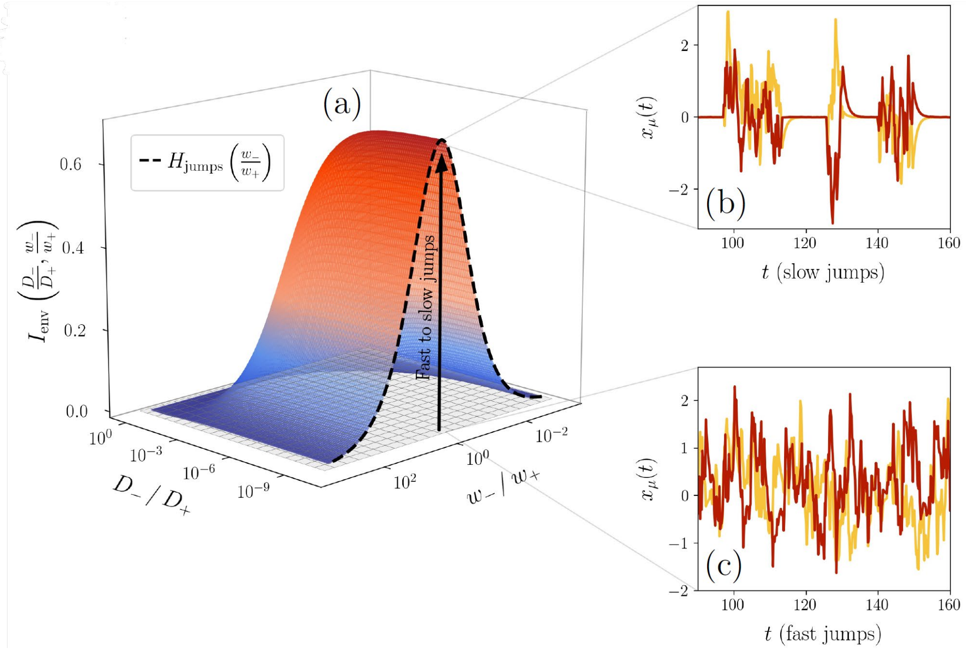


Figure 1: A quantitative example of our modeling approach. (a) Plot of the mutual information between two non-interacting Ornstein-Uhlenbeck processes $x_{1,2}(t)$ that share the same environment. The environment action is modeled through a switching diffusion coefficient that follows a jump process between two states D_- and D_+, with jump rates w_- and w_+. (b-c) Depending on the timescale separation between the two processes and the environment, the system can exhibit a bursty, coordinated behavior. In this limit, where the jumps are slow, a non-vanishing mutual information emerges and it is bounded by the entropy of the jumps, which reflects our ignorance about the stochastic environment. On the other hand, in the fast jump limit the contribution of the environment to the mutual information between $x_1(t)$ and $x_2(t)$ vanishes.

Discovering Causal Structure with Reproducing-Kernel Hilbert Space ε-machines

N. Brodu[1], *J. P. Crutchfield*[2]*. (1) Inria, Geostat team, 200 avenue de la vieille tour, 33405 Talence, France, nicolas.brodu@inria.fr. (2) Complexity Sciences Center and Department of Physics and Astronomy, University of California at Davis, One Shields Avenue, Davis, CA 95616, USA.*

In [1], we merge computational mechanics' definition of causal states (predictively-equivalent histories) with reproducing-kernel Hilbert space (RKHS) representation inference. The result is a widely-applicable method that infers causal structure directly from observations of a system's behaviors whether they are over discrete or continuous events or time. A structural representation—a finite- or infinite-state kernel ε-machines—is extracted by a reduced-dimension transform that gives an efficient representation of causal states and their topology. In this way, the system dynamics are represented by a stochastic (ordinary or partial) differential equation that acts on causal states. We introduce an algorithm to estimate the associated evolution operator. Paralleling the Fokker-Plank equation, it efficiently evolves causal-state distributions and makes predictions in the original data space via an RKHS functional mapping. We first demonstrate these techniques, together with their predictive abilities, on artificial data sets generated by known processes, both discrete and continuous cases, including a chaotic attractor. The method robustly estimates causal structure and recovers the process attractor in the presence of varying external and measurement noise levels. We then apply the technique to measured time series of monthly numbers of sunspots.

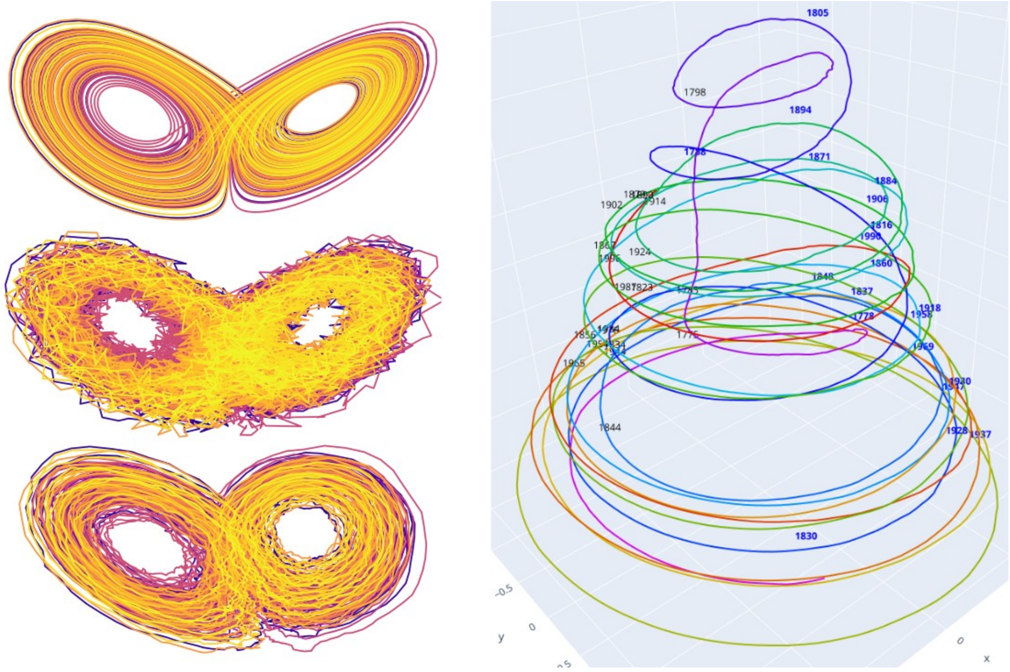


Figure 1: Left : Lorenz attractor, a strongly corrupted version of it by additive noise, the reconstructed attractor with our technique, from corrupted data only, in a surrogate parameter space. The attractor is well recovered despite the noise. Right: Empirical attractor reconstructed from monthly sunspots measurements. Years in blue/black are solar maxima/minima. The algorithm finds three main descriptive parameters: the first two encode the 11-year cycle and its phase, the third encodes the long-term fluctuations of the maxima.

Reference

[1] N. Brodu, J.P. Crutchfield, *ArXiv.org:2011.14821*

Prediction of New Scientific Collaborations Through Multiplex Networks

*M. Tuninetti[1], **A. Aleta**[1], D. Paolotti[1], Y. Moreno[1] & M. Starnini[1] (1) ISI Foundation, Turin, Italy, albertoaleta@gmail.com*

The establishment of new collaborations among scientists fertilizes the scientific environment, fostering novel discoveries. Understanding the dynamics driving the development of scientific collaborations is thus crucial to characterize the structure and evolution of science. In this work[1], we leverage the information included in publication records and reconstruct a categorical multiplex networks to improve the prediction of new scientific collaborations. Specifically, we merge different bibliographic sources to quantify the prediction potential of scientific credit, represented by citations, and common interests, measured by the usage of common keywords. We compare several link prediction algorithms based on different dyadic and triadic interactions among scientists, including a recently proposed metric [2] that fully exploits the multiplex representation of scientific networks. Our work paves the way for a deeper understanding of the dynamics driving scientific collaborations, and validates a new algorithm that can be readily applied to link prediction in systems represented as multiplex networks.

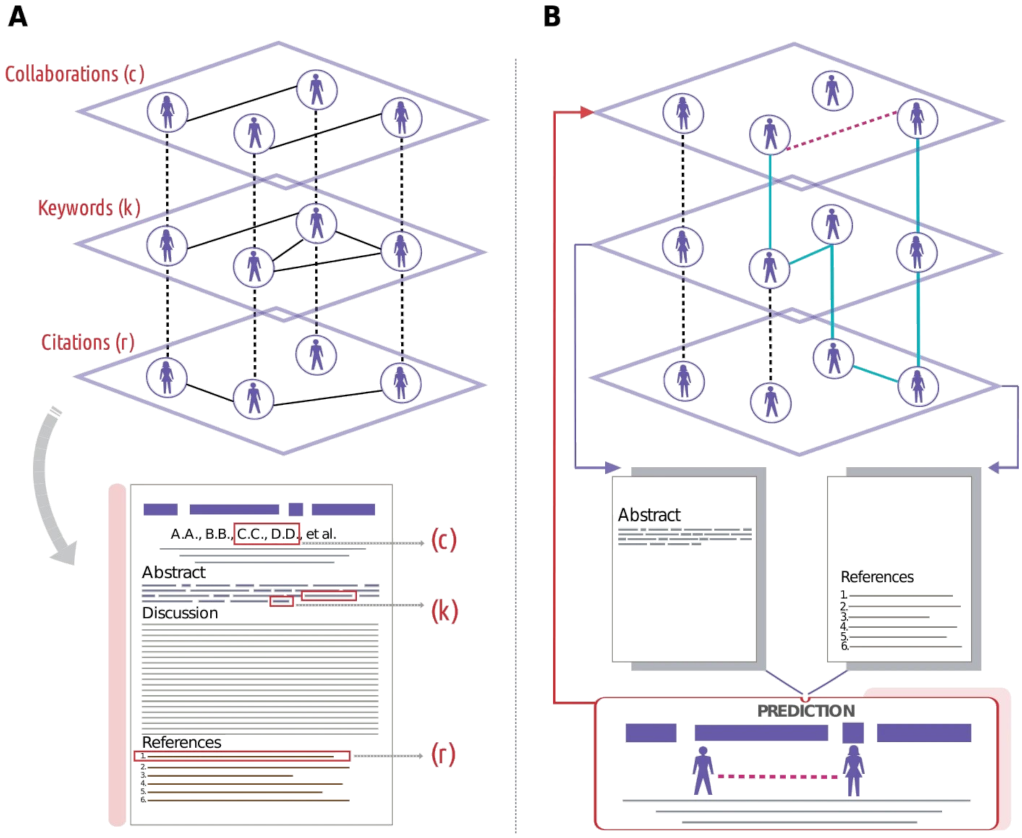


Figure 1: Prediction of new collaborations using multiplex networks. We first build a multiplex network using three different kinds of relational data between scientists (panel A) and then use two of them (panel B) to predict new collaborations among them.

References (optional)

[1] M. Tuninetti, *EPJ Data Science*, doi: 10.1140/epjds/s13688-021-00282-x

[2] A. Aleta, *Physical Review Research*, doi: 10.1103/PhysRevResearch.2.042029

Universal dynamics of ranking

Gerardo Iñiguez^{1,2,3,*}, Carlos Pineda^{4}, Carlos Gershenson^{3,5,6,7}, and Albert-László Barabási^{7,8,1,*}. (1) Department of Network and Data Science, Central European University, 1100 Vienna, Austria. iniguezg@ceu.edu (2) Department of Computer Science, Aalto University School of Science, 00076 Aalto, Finland. (3) Centro de Ciencias de la Complejidad, Universidad Nacional Autónoma de México, 04510 CDMX, Mexico. (4) Instituto de Física, Universidad Nacional Autónoma de México, 04510 CDMX, Mexico. (5) Instituto de Investigaciones en Matemáticas Aplicadas y en Sistemas, Universidad Nacional Autónoma de México, 04510 CDMX, Mexico. (6) Lakeside Labs GmbH, Lakeside Park B04, 9020 Klagenfurt am Wörthersee, Austria. (7) Network Science Institute, Center for Complex Network Research & Department of Physics, Northeastern University, 02115 Boston, MA, USA. alb@neu.edu (8) Channing Division of Network Medicine & Department of Medicine, Brigham and Women's Hospital, Harvard Medical School, 02215 Boston, MA, USA

Virtually anything can be and is ranked; people and animals, universities and countries, words and genes. Rankings reduce the components of highly complex systems into ordered lists, aiming to capture the fitness or ability of each element to perform relevant functions, and are being used from socioeconomic policy to knowledge extraction. A century of research has found regularities in ranking lists across nature and society when data is aggregated over time. Far less is known, however, about ranking dynamics, when the elements change their rank in time. To bridge this gap, here we explore the dynamics of 30 ranking lists in natural, social, economic, and infrastructural systems, comprising millions of elements, whose temporal scales span from minutes to centuries. We find that the flux governing the arrival of new elements into a ranking list reveals systems with identifiable patterns of stability: in high-flux systems only the top of the list is stable, while in low-flux systems the top and bottom are equally stable. We show that two basic mechanisms — displacement and replacement of elements — are sufficient to understand and quantify ranking dynamics. The model uncovers two regimes in the dynamics of ranking lists: a fast regime dominated by long-range rank changes, and a slow regime driven by diffusion. Our results indicate that the balance between robustness and adaptability characterizing the dynamics of complex systems might be governed by random processes irrespective of the details of each system.

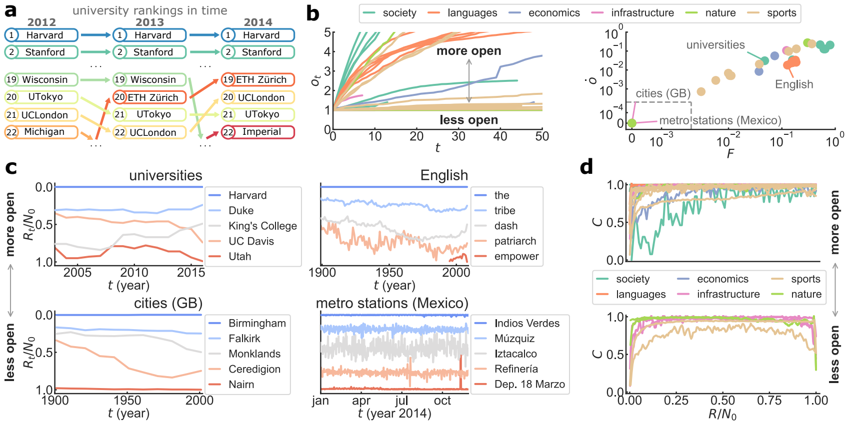


Figure 1: Ranking lists in nature and society show generic patterns in their dynamics. (a) Yearly top ranking of universities. (b) (Left) Rank openness for studied systems (Right) Correlation between mean openness rate and mean flux. (c) Time series of rank occupied by elements across the ranking list in selected systems. (d) Rank change across ranking lists.

Reference

Iñiguez, G., Pineda, C., Gershenson, C., and Barabási, A.-L. (2021). Universal dynamics of ranking. *arXiv*:2104.13439.

An Analytic Approach to Estimate the Connectivity of a Boolean Network from its Noise Sensitivity

O. Zapata[1]. *(1) CUAIEED, UNAM, Cto. Exterior s/n, C.U., Coyoacán, 04510 Ciudad de México, México, octavio_zapata@cuaieed.unam.mx.*

Given a Boolean function *f*, the *noise sensitivity NS(f;δ)* of *f* at δ is the probability that *f* takes different values on two random inputs selected in the following way. The first input is a random binary vector chosen uniformly; the second input is a random perturbation of the first input, obtained by changing the value of each coordinate independently with probability δ. This concept was introduced in 1988 by Kahn, Kalai, and Linial [1]. It has applications in many areas such as combinatorics, complexity theory, computational learning, social choice, and statistical physics [2–7].

A Boolean function is said to be a *k-junta* if it only depends on *k* input coordinates. We say that a Boolean function *f* is an *(ε,k)-junta* if there exists a *k*-junta *g* such that the probability that *f* and *g* take different values on a random input is at most ε. In other words, an (ε,*k*)-junta is a function that is ε-close to depending on *k* coordinates. Juntas are "noise-resistant" in the sense that if *f* is an (ε,*k*)-junta, then *NS(f;δ)* is bounded from above by $2\varepsilon + k\delta$. A sufficient condition for being close to a junta in terms of noise sensitivity was given in 2002 by Bourgain [8]. Bourgain's theorem says that any function with sufficiently low noise sensitivity must be close to a junta.

A *Boolean network* is a family of *k*-juntas mapping binary vectors to binary vectors. This notion was proposed in 1969 by Kauffman [9] as a model for gene regulatory networks. In this talk, we extend the definition of noise sensitivity to the context of Boolean networks and present a vectorial version of Bourgain's result. As an application, we show how to estimate the connectivity of a Boolean network from its noise sensitivity.

References

[1] J. Kahn, G. Kalai, N. Linial, *Proc. 29th Ann. IEEE Symp. on Foundations of Comp. Sci.*, 1988, 68–80.

[2] J. Håstad, *Proc. 29th Ann. ACM Symp. on the Theory of Computing*, 1997, 1–10.

[3] I. Benjamini, G. Kalai, O. Schramm, *Inst. Hautes Études Sci. Publ. Math.*, 90 (1999) 5–3.

[4] N. Bshouty, J. Jackson, C. Tamon, *Proc. 12th Ann. Workshop on Comp. Learning Theory*, 1999, 75–80.

[5] O. Schramm, B. Tsirelson, *Elect. Comm. in Probab.*, 4 (1999) 39–49.

[6] R. O'Donnell, *Ph.D. thesis*, Massachusetts Institute of Technology, 2003.

[7] C. Garban, J. E. Steif, *Cambridge University Press,* 5 (2014).

[8] J. Bourgain, *Israel Journal of Mathematics*, 131 (2002) 269–276.

[9] S.A. Kauffman, J. *Theoret. Biol.*, 22 (1969) 437–467.

The distance backbone of complex networks

Tiago Simas[1], Rion Brattig Correia[2,3,4], ***Luis M. Rocha***[2,3,5]**. (1) Universidade Lusófona Lisboa, Portugal. (2) Center for Social and Biomedical Complexity, Luddy School of Informatics, Computing & Engineering, Indiana University, Bloomington IN, USA. (3) Instituto Gulbenkian de Ciência, Oeiras, Portugal. (4) CAPES Foundation, Ministry of Education of Brazil, Brasília, DF, Brazil. (5) Department of Systems Science and Industrial Engineering, Binghamton University, Binghamton NY, USA. *rocha@binghampton.edu. (2)*

Redundancy needs more precise characterization as it is a major factor in the evolution and robustness of networks of multivariate interactions. We investigate the complexity of such interactions by inferring a connection transitivity that includes all possible measures of path length for weighted graphs. The result, without breaking the graph into smaller components, is a distance backbone subgraph sufficient to compute all shortest paths. This is important for understanding the dynamics of spread and communication phenomena in real-world networks. The general methodology we formally derive yields a principled graph reduction technique and provides a finer characterization of the triangular geometry of all edges---those that contribute to shortest paths and those that do not but are involved in other network phenomena. We demonstrate that the distance backbone is very small in large networks across domains ranging from air traffic to the human brain connectome, revealing that network robustness to attacks and failures seems to stem from surprisingly vast amounts of redundancy [1].

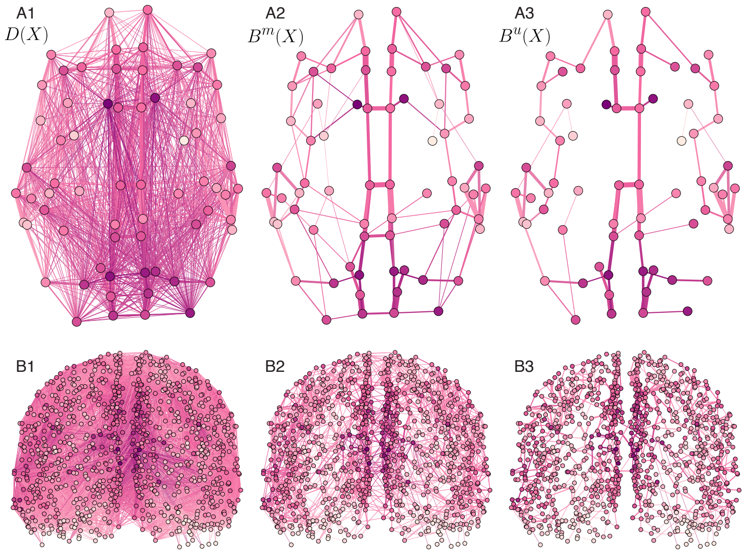


Figure 1: Human Connectome Network (HCN) and Backbones. (**A1-3**): HCN-Coarse. (**B1-3**): HCN-Fine. (**A1,B1)**: Original distance Networks [2], whose distance weights are inversely proportional to the volume of cortico-cortical axonal pathways between brain regions (nodes), obtained via diffusion spectrum imaging. (**A2,B2**). Metric backbone with only 9.23% and 17.57% of original edges HCN-Coarse and HCN-Fine, respectively. (**A3,B3**). Ultrametric backbone with only 5.66% and 5.53% of original edges for HCN-Coarse and HCN-Fine, respectively.

References

[1] T. Simas, R.B. Correia, L.M. Rocha [2021]. "The distance backbone of complex networks". *Journal of Complex Networks*. **In Press**. arXiv:2103.04668.

[2] Hagmann, P., Cammoun, L., Gigandet, X., Meuli, R., Honey, C. J., Wedeen, V. J. & Sporns, O. (2008) Mappingthe structural core of human cerebral cortex.PLoS Biol,6(7), e159.

Unraveling the Role of Node Metadata in Network Robustness: the Feature-Based Percolation Model

Oriol Artime[1,], *Manlio De Domenico*[1] *(1) Bruno Kessler Foundation, Via Sommarive 18, 38123 Povo (TN) Italy*

Percolation is an emblematic model used to understand the robustness of interconnected systems. Despite being a model broadly studied in statistical physics and mathematics, from a theoretical perspective it is usually investigated in relatively simple scenarios, such as the removal of the system's units in random order –simulating unpredictable site failures– or sequentially ordered –simulating targeted attacks– by specific topological descriptors, the simplest one being the number of node connections. However, in the vast majority of empirical applications, it is required to dismantle the network following more sophisticated protocols than the aforementioned ones, such as based on more convoluted topological properties or even non-topological node metadata obtained from the application domain.

In this work we propose a novel mathematical framework to fill this gap: a network is enriched with features and its nodes are then removed according to their importance in the feature space. Percolation analysis is performed, theoretically and numerically, as a function of the feature distribution, finding an excellent match between the analytical results and the simulations. Several degree-feature relations of diverse nature are explored to show the applicability of the theory. We start from ad hoc degree-feature distributions that capture the main characteristics of correlations observed in empirical systems, moving to features that arise naturally in the process of network creation and ending with the case in which features are coupled to dynamical processes running on top of the network, such as epidemics or biochemical dynamics, among others. Both synthetic and real-world networks of different nature are considered in the analysis. Moreover, we show the potential of our model by employing state-of-the-art Bayesian probability techniques that are able to give the most plausible closed-form expression for the degree-feature distribution when it cannot be computed analytically. By feeding these most plausible expressions into the equations of our model, we can study feature-based percolation in systems for which it is only known the feature and the degree of the individual nodes, instead of the entire degree-feature joint probability distribution (see Fig.). This considerably broadens the applicability of the theory and bridges our theory, grounded on statistical physics, with Bayesian machine learning techniques suitable for knowledge discovery. The article has been recently published in *Nature Communications* **12**(1): 1-12 (2021).

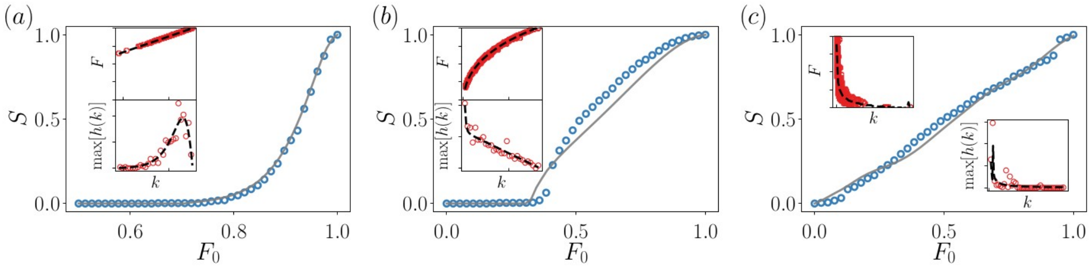


Figure 1: Percolation for (a) mutualistic dynamics in symbiotic ecosystems, (b) population dynamics and (c) mass-action kinetics in biochemistry (solid lines are the theory, points are simulations). *F0* is a parameter of the feature distribution, *S* is the size of the giant component. The insets are different variables fitted using machine-learning techniques that are necessary to compute *S*. The dynamical equations are not given due to space constrains. All dynamics are simulated in real networks.

Comparing Methods to Quantify Multivariate Synchrony with multiSyncPy, using Simulated Noise, Kuramoto Oscillators and Human Movement Data

D. Hudson[1,2], *T. J. Wiltshire*[1], *M. Atzmueller*[2]. *(1) Tilburg University, Warandelaan 2, 5037 AB, Tilburg, The Netherlands. (2) Osnabrueck University, P.O. Box 4469, 49069 Osnabrueck, Germany, daniel.dominic.hudson@uni-osnabrueck.de.*

Across physical, biological, and social systems, interacting components of complex systems coordinate, and at times, synchronize their behavior. Despite being a seemingly universal property exhibited by complex systems [1] there is still much uncertainty about how synchronization originates in systems, how it changes over time and, as we will discuss, what forms it takes in groups larger than two (or between more than two variables) [2]. To advance the scientific understanding of these phenomena, accessible methods are needed to quantify the degree of synchronization observed in multi-component and multi-modal complex systems. To this end, we developed multiSyncPy, an open-source software package for computing multivariate synchrony. Our package includes state-of-the-art multivariate methods including symbolic entropy, multidimensional recurrence quantification, coherence, the cluster-phase 'Rho' metric, and a statistical test based on the Kuramoto order parameter. In this presentation, we will demonstrate the use of our multivariate synchrony metrics on a series of synthetic datasets and two real-world empirical datasets, showcasing the results obtained from a variety of situations. This is a novel combination of methods, each of which has previously only been examined in isolation from one another. Thus, we see this as an important initial investigation of how different multivariate synchrony metrics perform. In particular, we examine how our different metrics perform compared to a surrogate baseline, when applied to: data generated from autoregressive processes with correlated noise, Kuramoto oscillators, and two empirical data sets of body movement during group interactions. An example comparison of these metrics on Kuramoto oscillators with increasing coupling strength is shown in Figure1. multiSyncPy is freely available at: https://github.com/cslab-hub/multiSyncPy, and at the Python package index. Additional details of the package and analyses can be found in our pre-print: https://psyarxiv.com/abquk/.

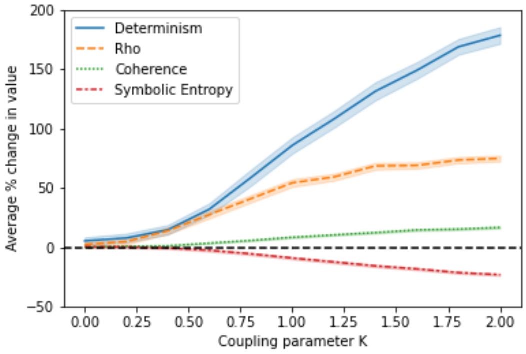


Figure 1: Change over baseline of synchrony metrics as coupling strength is increased

References

[1] P. G. Amazeen, "From physics to social interactions: Scientific unification via dynamics," *Cognitive Systems Research*, 52 (2018) 640–657.

[2] S. Hoehl, M. Fairhurst, A. Schirmer. "Interactional synchrony: Signals, mechanisms and benefits," *Social Cognitive and Affective Neuroscience*, 16 (2021) 5-18.

Effects of local interactions in epidemics outbreaks in networks of structured populations

__C. Simal__[1,2], B. Siebert[2], T. Carletti[1], J. P. Gleeson[2], M. Asllani[3] (1) naXys, Namur Institute for Complex Systems, & Department of Mathematics, University of Namur, Rue Grafé 2, B5000, Belgium (2) MACSI, Department of Mathematics and Statistics, University of Limerick, Limerick V94 T9PX, Ireland (3) School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland (3) School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland cedric.simal@student.unamur.be

Spreading phenomena constitute a class of processes that model applications spanning from the diffusion of information to infectious diseases [1,2]. Depending on the spatial extension of the social interactions, the mathematical modeling has been focused either on the role that the pairwise interactions between individuals have on the overall dynamics or geometry of the spatial support where the agents move before they enter into contact. However, little is known about the effects that the local and global network structures simultaneously have on the dynamics outcomes, although the unequal interactions between these two parts might strongly affect the overall behavior. In this work, we introduce a mean-field model that explicitly considers the role that local interactions play in spreading models of structured populations [3]. Such a concatenated network model has been named a metaplex network [4]. Based on linear stability analysis, we show that the degree of the local networks can capture their contribution in the global reaction-diffusion system and detect which of them drives the spreading of infection. Furthermore, we show that counterintuitively the outbreak of epidemics is not necessarily related to the spatial patched where the disease is seeded, neither their local properties. Instead, it is grounded on the global characteristics of the diffusion network. We also test our model on empirical transport networks and obtain a good validation of our results.

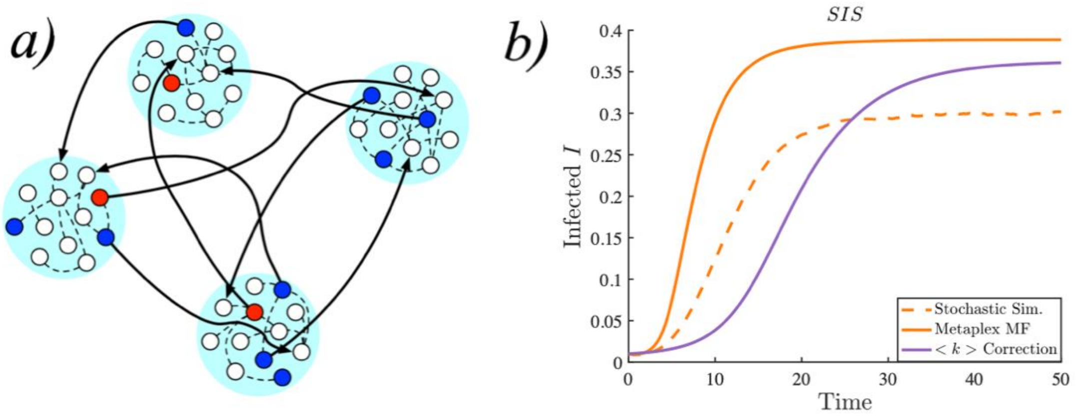


Figure 1: (Left) A metaplex network model (right) models' comparison of evolution of infection in a SIS spreading process

References

[1] I. Z. Kiss, J. C. Miller, and P. L. Simon, *Mathematics of network epidemics: from exact to approximate models*, (Springer, 2017).

[2] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, and A. Vespignani, *Rev. Mod. Phys.* **87**(3), 925 (2015).

[3] J. T. Davis, N. Perra, Q. Zhang, Y. Moreno, and A. Vespignani, *Nat. Phys.* **16**(15), 590–596 (2020).

[4] M. Asllani, B. R. da Cunha, E. Estrada, and J. P. Gleeson, *New J. Phys.* **22**, 063037 (2020).

Causality in Reversed Time Series: Reversed or Conserved?

J. Kořenek[1,2], *J. Hlinka*[1,3]. *(1) Institute of Computer Science, Czech Academy of Sciences, Pod vodarenskou vezi 271/2, 182 07, Prague, Czech Republic, korenek@cs.cas.cz (2) Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, Brehova 7, 115 19, Prague, Czech Republic (3) National Institute of Mental Health, Topolova 748, 250 67, Klecany, Czech Republic*

The inference of causal relations between observable phenomena is paramount across scientific disciplines, however the means for such enterprise without experimental manipulation are limited. A commonly applied principle is that of the cause preceding and predicting the effect, taking into account other circumstances [1]. Intuitively, when the temporal order of events is reverted, one would expect the cause and effect to apparently switch roles [2]. This has been demonstrated in bivariate linear systems and utilized in design of improved causal inference scores [3], while such behavior in linear systems has been put in contrast with nonlinear chaotic systems where the inferred causal direction appears unchanged under time reversal.

We follow the observations in linear processes in more depth. In particular, we extend the analysis to multivariate systems, i.e. networks. For arbitrary network size, we derive the necessary and sufficient condition of causal structure reversal under time reversal for vector autoregressive processes of order 1; it turns out that the key property is the normality (orthogonality) of the causal interaction matrix. In the previously studied bivariate case, we provide an analytic proof that the causal structure reversal in fact never happens perfectly, but only approximately (apart from two rather trivial cases with particular coupling symmetry). We consequently show that the minimal unidirectionally coupled network for which the causal structure reversal appears has 3 subsystems, and document this by an example.

We finally turn to real-world data and simulation scenarios, studying the extent to which the causal structure reversal appears in brain and climate network under time reversal. The results show that already the linear approximations of both brain and climate systems show imperfect causal structure reversal under time reversal, while the extent to which this property is broken is closely predicted by deviation of the coupling matrix from normality in both realizations of randomly connected networks as well as these real-world scenarios. Finally, we discuss the relevance of these findings in a wider context, including how it problematizes the use of comparison with time-reversed version of a given process for inference of causal structures.

Acknowledgements

This work was supported by the Czech Science Foundation projects No. 19-11753S, No. 21-32608S, No. 21-17211S and by Grant Agency of the Czech Technical University in Prague, grant No. SGS20/183/OHK4/3T/14.

References

[1] C. W. Granger, Econometrica, 37 (1969), 424–438

[2] M. Paluš, A. Krakovská, J. Jakubík, M. Chvosteková, Chaos: An Interdisciplinary Journal of Nonlinear Science, 28 (2018), 075307

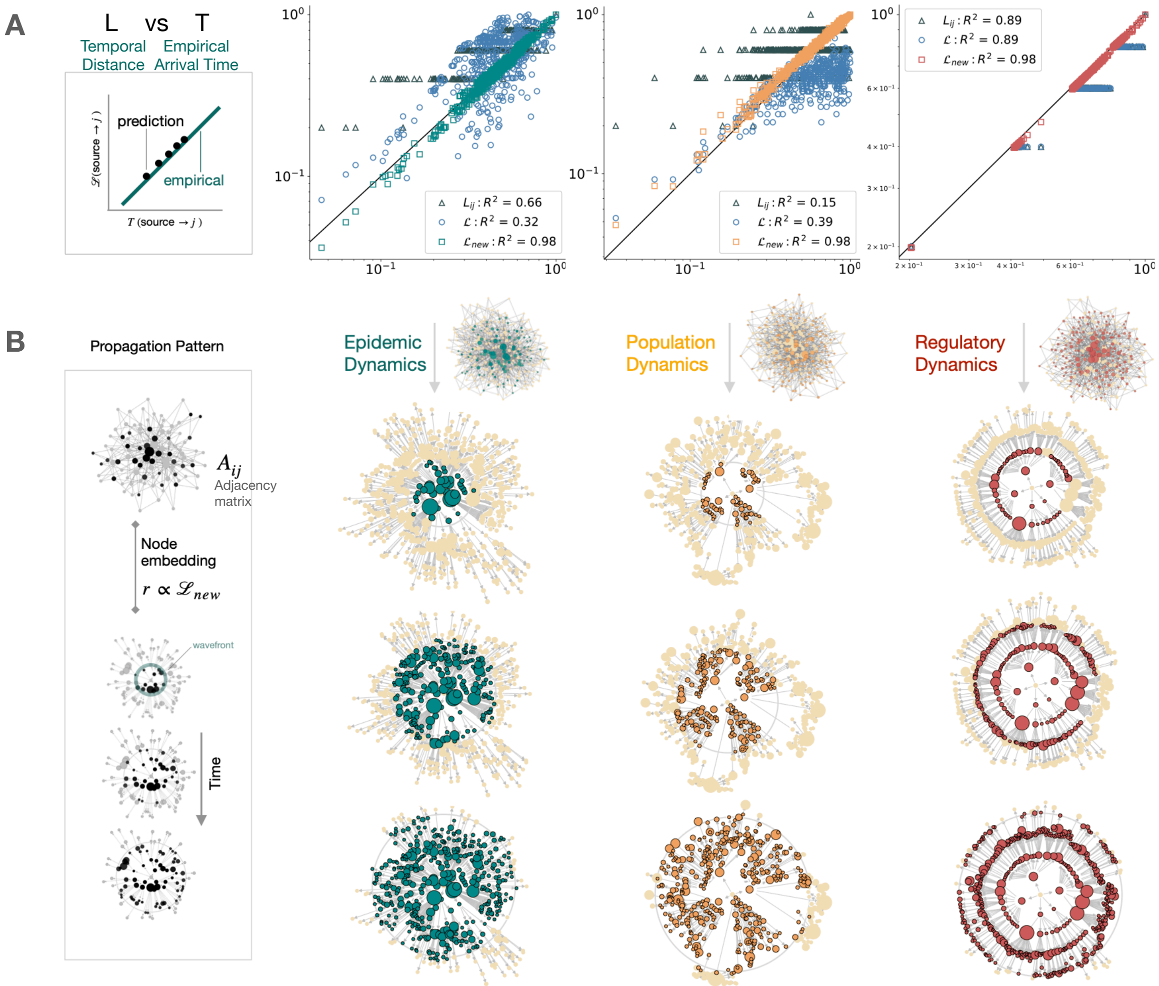
[3] I. Winkler, D. Panknin, D. Bartz, K. R. Muller, S. Haufe, IEEE Transactions on Signal Processing, 64 (2016), 2746–2760

Multi-pathways temporal distance unravels the hidden geometry of network-driven processes

S. Bontorin (1,2) and M. De Domenico (2)
1 - Department of Physics, University of Trento, Via Sommarive 14, 38123 Povo (TN), Italy
2 - Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo, Italy

Network-based interactions are widely adopted to model the behavior of technological and natural systems, where understanding information flow between nodes is of paramount importance in predicting its functioning. As different dynamic processes give rise to profoundly different, and apparently random, perturbation patterns on top of the same topology of connections, the interplay between network connectivity and the dynamics on top of it has been the focus of recent studies [1].

The identification of dynamical regimes and predictive observables via scaling laws sheds light over signal propagation in complex systems, and overcomes the paradigm of information flow being solely dependent on network topology [2]. In this paper we perform a path-driven analysis of perturbation propagation on steady-state systems, establishing the relevant paths for information transfer given the exponential decay of correlations. We exploit known scaling laws [3] to define a new multi-pathways temporal distance between nodes that predicts propagation times and improves over the existing definition. We validate these measures on synthetic models, simulating dynamical processes pertaining to different regimes, measuring the goodness in predicting empirical arrival times. This measure [4] fully predicts the time of arrival of a perturbation and naturally encodes the concerted behavior of the ensemble of paths connecting two nodes in conveying propagation. The dynamics induces a latent geometry in which the signal propagation resembles the traveling wave solution [5] of reaction-diffusion systems.



***Predictions of temporal distance and wavefront propagation**. **A)** Predictions vs Empirical $T(j \to i)$ for Population, Epidemics (SIS model) and Regulatory models on a Scale Free network (N=500,$\gamma = 3$) - Standard topological shortest path length, the existing measure $\mathscr{L}$ and the new metric $\mathscr{L}_{new}$ are compared with an R^2 determination coefficient. **B)** The geometry predicted by $\mathscr{L}_{new}$ allows to transform the complex propagation in an ordered concentric pattern. Node size is proportional to its degree.*

References:

- [1] B. Barzel et Al., *Nature Physics* **9,** 673–681 (2013)
- [2] U. Harush et Al., *Nature Communications* **8,** 2181 (2017)
- [3] C. Hens et Al., *Nature Physics* **15,** 403–412 (2019)
- [4] S. Bontorin and M. De Domenico, submitted (2021)
- [5] D. Brockmann et Al., *Science* ***342** issue 6164*, 1337-1342 (2013)

The Temporal Rich Club Phenomenon

Nicola Pedreschi[1,2], Demian Battaglia[2], Alain Barrat[1,3]

[1]) Aix Marseille Univ, Université de Toulon, CNRS, CPT, Turing Center for Living Systems, Marseille, France
[2]) Aix Marseille Univ, Inserm, INS, Institut de Neurosciences des Systèmes, Marseille, France
[3]) Tokyo Tech World Research Hub Initiative (WRHI), Institute of Innovative Research, Tokyo Institute of Technology. Japan

Identifying the hidden organizational principles and relevant structures of networks representing physical systems, ranging from the cellular scale to that of networks of social interactions, is fundamental for understanding their properties. In temporal networks one must further investigate the dynamic nature of such structures to uncover the key mechanisms of the time evolution of these systems and the processes unfolding on them.

In static networks, the "rich-club" phenomenon, i.e., the tendency of well-connected nodes to also connect to each other, has been extensively investigated. However, a well-connected structure in a static network might correspond in temporal networks to links that are never simultaneously present. Therefore, we propose here a novel measure to investigate the simultaneity and stability over time of such structures in temporal networks: the "*temporal rich club*". Specifically, we consider a temporal network in discrete time on a time interval $[0, T]$, and we denote by $G = (V, E)$ the corresponding aggregated graph. Our aim is to quantify whether the $N_{>k}$ nodes with degree larger than k in G (the sub-graph $S_{>k}$) are more connected *simultaneously* than by chance. Performing such analysis on a sub-network $S_{>k}$ with a large value of k allows to investigate the stability over time of the ties connecting the nodes with high centrality in the aggregated graph, i.e., whether a "*temporal rich club phenomenon*" is present. To this aim, we define $\epsilon_{>k}(t, \Delta)$ as the number of ties between nodes of $S_{>k}$ that remain stable over the time interval $[t, t+\Delta]$, normalized by the maximal density $|S_{>k}|(|S_{>k}|-1)/2$, and we measure the maximal observed cohesion $M(k, \Delta)$ within $S_{>k}$ that remains stable during an interval of length Δ:

$$M(k, \Delta) \equiv \max_t \epsilon_{>k}(t, \Delta)$$

We test our novel measure on a range of temporal networks, ranging from temporal networks of face-to-face interactions in various contexts, to air transportation temporal networks and to brain functional connectivity networks.

For a given temporal network, we plot $M(k, \Delta)$ vs. k and Δ and compare it with a reshuffled version of the temporal graph (Figure 1.a-b). Figure 1.A shows that higher values of cohesion are observed, in face-to-face contact data, for $\Delta = 1$, even at low values of k, and that the cohesion increases with k. When the time-interval Δ increases, the temporal cohesion of sub-networks $S_{>k}$ drops rapidly for low values of k. For sub-networks composed by nodes of higher k, the cohesion decreases more slowly. We test the significance of such behaviour by performing the same analysis on a randomized version of the dataset (Figure 1.b): the randomization preserves the statistical properties of the aggregated graph G and the activity timeline of the temporal network, while randomizing the instantaneous edges. The comparison of original and reshuffled data highlights that the maximal cohesion $M(k, \Delta)$ reflects purely temporal structures. The dynamic evolution of such structures can be captured by studying the instantaneous cohesion $\epsilon_{>k}(t, \Delta)$ for different values of k and Δ (Figure 1.c): in this example, by plotting the cohesion of the sub-network $S_{>k}$ with $k = 88$, we find how this group of students is composed of children belonging to different classes, as the moments of highest instantaneous cohesion correspond to recess or lunchtime.

The analysis of temporal rich clubs can therefore help shed light on the existence and stability of *simultaneous* connections between the most connected nodes of a temporal network, and on their relation with specific moments of interest in the evolution of a temporal network.

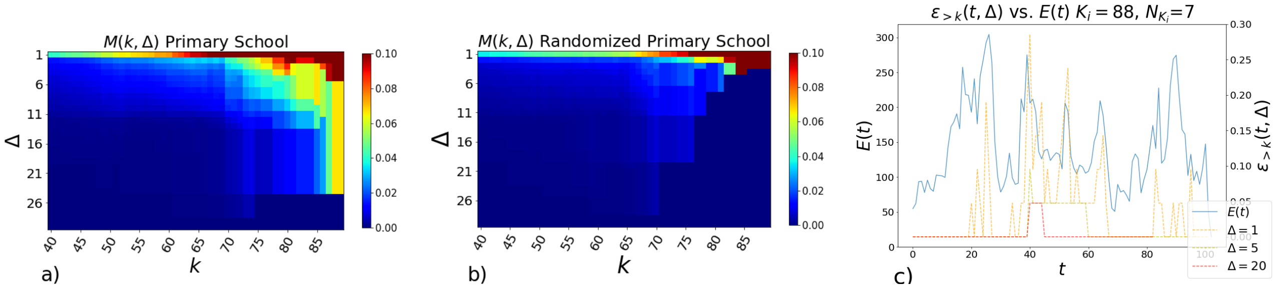


Figure 1: a) $k - \Delta$ diagram of the cohesion $M(k, \Delta)$ of a temporal network describing contacts between children in a primary school (www.sociopatterns.org); b) $M(k, \Delta)$ diagram for the same values of k and Δ as in the previous plot; c) Instantaneous values of the cohesion $\epsilon_{>k}(t, \Delta)$ of $S_{>k}$ for various values of the temporal resolution Δ, together with the instantaneous number of edges of the network $E(t)$ (in blue).

Information-based Granger Causality in High-Dimensional Discrete Time Series and Application to World Financial Markets

***M. Papapetrou**, E. Siggiridou, D. Kugiumtzis Department of Electrical and Computer Engineering, Thessaloniki 54124, Greece, mariapap@ece.auth.gr.*

The study of complex systems has gained great interest recently due to the abundance of big data and the advances of the methodology for their analysis. Different Granger causality measures have been used to estimate the causal relationships among the observed variables of a complex system. If the number of observed variables is large, resulting in high-dimensional time series, dimension reduction approaches are employed to estimate direct causal effects. In the case of high-dimensional continuous-valued time series the measure of partial conditional mutual information from mixed embedding (PMIME) [1] is found to be appropriate for the estimation of direct causality. In this work, this measure is adapted appropriately for discrete-valued time series (called discrete PMIME, DPMIME). Further, the DPMIME is evaluated in terms of its ability to determine the connections of a high-dimensional complex network of the underlining dynamical system or stochastic process. For this, a simulation study is conducted, and the discrete-valued time series are generated as follows: (a) discretizing time series from continuous-valued dynamical systems, such as coupled Hénon maps, (b) directly from transition probability matrices estimated on discretized time series, (c) from multivariate integer-autoregressive systems (MINAR) [2], (d) from mixture transition distribution models (MTD) [3,4]. The subsystems of the generating system are coupled according to a given network topology, including random, small-world and scale-free. In the estimation of the respective causality network using DPMIME, the binary directed connection between two nodes of the studied network is obtained by the statistical significance of the respective DPMIME value. For the statistical significance of DPMIME, the parametric significance test, based on the respective test for conditional mutual information developed by our team in previous related studies, is compared to the resampling (randomization/surrogate) test and it is shown that the parametric test is equally accurate to the resampling test but much faster to compute. Further, the DPMIME is used to investigate whether Brexit and the global financial crisis had effect on the causality (inter-dependence) network of the financial world market.

References

[1] D. Kugiumtzis, *Physical Review E*, 87 (2013) 062918.

[2] X. Pedeli, D. Karlis, *Computational Statistics and Data Analysis*, 67 (2013) 213.

[3] A.E. Raftery, *Journal of the Royal Statistical Society*, 47 (1985) 528.

[4] W. Ching, M. Ng, E. Fung, *Linear Algebra and its Applications*, 428 (2008) 492.

Information-based Estimation of Causality Networks from High Dimensional Multivariate Time Series

A. Fotiadis[1]*, D. Kugiumtzis*[1]*. (1) Aristotle University of Thessaloniki, Thessaloniki 54124, Greece, fotiakyl@auth.gr*

One of the most challenging aspects in the study of the complex dynamical systems is the identification of their underlying interdependence structure from time series data. Being in the era of Big Data, this problem gets even more complicated since the number of available variables may be very large. Under this setting, the estimation of direct causality with a Granger causality-type measure inevitably has to involve some kind of dimension reduction. The measure should, ideally, also be capable to detect nonlinear effects, persistently present in real world complex systems. The model-free information-based measure of partial mutual information from mixed embedding (PMIME) [1] has been developed to address these issues and it was found to perform well on multivariate time series of moderately high dimension. Specifically, it was found that the estimated causality network by PMIME matches well the complex network of the interdependence structure of the original complex system (Figure 1, taken from [2]). Here, the problem of forming complex networks from direct, possibly nonlinear, causality estimation from high dimensional time series at the order of hundreds is investigated. Moreover, the effect of the dimension of the system on the estimation accuracy of the original complex network is studied in relation to the sparsity of the original network. The performance of the measure PMIME is tested on two coupled dynamical systems, one in discrete time (coupled Henon maps) and one in continuous time (coupled Mackey-Glass delay differential equations), for an increasing number of coupled subsystems (dimension) up to 200 and fixed low density as well as increasing density of the original network. It is concluded that the correct detection of the underlying causality network depends mainly on the network density rather than on its dimension.

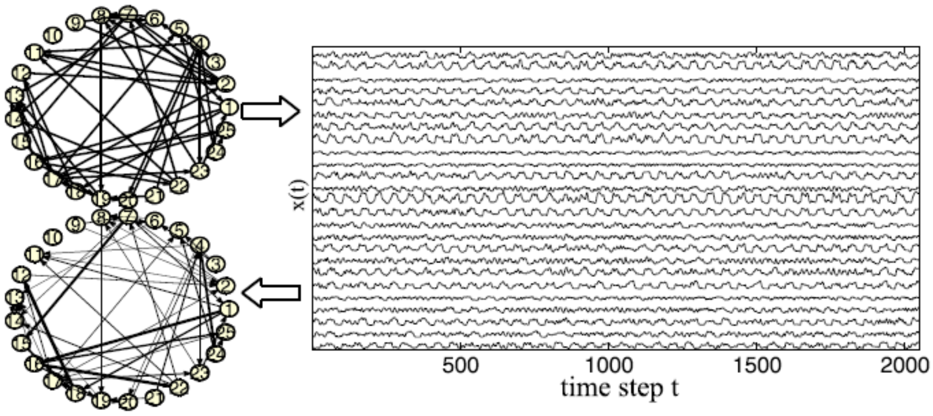


Figure 1: Formation of causality network given network topology and coupling dynamics [2].

Acknowledgements (optional)

This work was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the HFRI PhD Fellowship grant (Fellowship Number: 99510).

References (optional)

[1] D. Kugiumtzis, *Physical Review E*, 87 (2013) 6.

[2] C. Koutlis, D. Kugiumtzis, *Chaos*, 26 (2016) 093120.

Information-based Measure of Partial Mutual Information from Mixed Embedding (PMIME) Revisited

A. Fotiadis[1], *I. Vlachos*[1], *D. Kugiumtzis*[1]*.(1) Aristotle University of Thessaloniki, Thessaloniki 54124, Greece, fotiakyl@auth.gr*

The measure of partial mutual information from mixed embedding (PMIME) is an information theory-based measure to accurately identify the direct and directional coupling between variables or subsystems of a high dimensional dynamical, thus complex, system, without any a priori assumptions about the nature of the coupling relationship [1,2]. In its core, it is a forward selection procedure that aims to iteratively identify the lag-dependence structure of a given variable (response) to all the other observed variables (candidate drivers). This model-free approach is capable of detecting nonlinear interactions, abundantly present in real-world complex systems, and it was shown to perform well on multivariate time series of moderately high dimension. Although, the PMIME performs better than many other Granger causality-type measures in the analysis of time series with predominant nonlinear dynamics, it has been observed that the measure underperforms when applied on strongly stochastic (linear or nonlinear) systems, with a propensity to falsely detect non-existent relationships. Additionally, and by construction, the measure cannot detect purely synergistic relationships present in a system. In the current work, the issue of false detections is addressed by introducing an improved resampling significance test and a procedure of rechecking the identified drivers (backward revising). Regarding the inability to detect synergistic relationships, the PMIME is further enhanced by checking pairs of candidate drivers for the response variable after having considered all drivers individually. The effects of these modifications are investigated in a systematic simulation study on properly designed systems involving strong stochasticity, regressor terms regarding synergetic effects and system dimension ranging from K=3 to K=30. The overall results of the simulations indicate that these modifications indeed improve the performance of PMIME and alleviate to a significant degree the issues of the original algorithm. Guidelines for balancing between accuracy and computational efficiency are also given, particularly relevant for real-world applications.

Acknowledgements (optional)

This work was supported by the Hellenic Foundation for Research and Innovation (HFRI) under the HFRI PhD Fellowship grant (Fellowship Number: 99510).

References (optional)

[1] I. Vlachos, D. Kugiumtzis, *Physical Review E*, 82 (2010) 016207.

[2] D. Kugiumtzis, *Physical Review E*, 87 (2013) 062918.

Symmetry-breaking Mechanism for the Formation of Cluster Chimera Patterns

M. Asllani[1]*, B. A. Siebert*[2]*, A. Arenas*[3]*, J. P. Gleeson*[2] *(1) University College Dublin, Ireland*
malborasllani@gmail.com. (2) University of Limerick, Ireland, (3) Universitat Rovira i Virgili, Spain

The full understanding of the dynamics of complex systems is a significant challenge in physics this century. The emergence of collective behavior and the coexistence of order and disorder are fundamental key players in this process. A well–known phenomenon of collective behavior is that of synchronization, where a set of coupled oscillators spontaneously synchronize their phases. Synchronization is also related to an exotic behavior that has triggered scientists' curiosity for a long time: the emergence of chimera states, long-lasting dynamical patterns of coexistence of synchronous and asynchronous clusters of nodes in networked systems. Although many features have been understood so far, many questions remain unanswered.

In this work, we use a symmetry-breaking mechanism to derive two robust mechanisms for finding both stable cluster synchronization, and chimera states. The model makes use of a three-species reaction-diffusion model capable of generating spatially extended oscillations. In the alternative, the global oscillations can also emerge by employing a (highly) directed network. In both cases, a decisive requirement is that the network should be strictly modular.

Formulating the synchronization dynamics from a pattern formation approach means that the final shape of the equilibrium pattern depends on the dominant eigenvectors of the Laplacian of the network. Using a perturbative approach, we show that the Laplacian eigenvectors of the N-1 largest non-positive eigenvalues will have entries segregated respective to the N modules.

The linear stability analysis will select what eigenvalues will be positive and consequently what will be the eigenvector that will shape the final pattern. For instance, if an eigenvalue corresponding to a segregated eigenvector dominates, then the global oscillatory pattern will be cluster synchronized; otherwise, the network will have an asynchronous pattern. On the other side, if we balanced two dominant eigenvalues, respectively, with segregated and scattered eigenvectors, then the global pattern will be a chimera-like one where synchronous clusters of nodes will coexist with fully asynchronous ones.

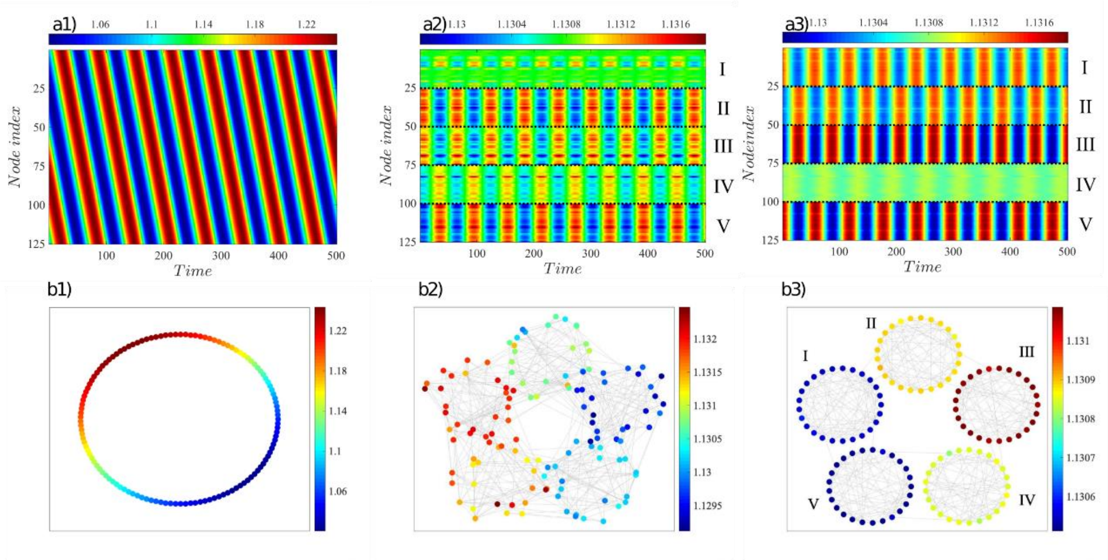


Figure 1: Evolution of oscillatory behavior as network transitions to a modular topology

https://arxiv.org/pdf/2102.03635.pdf

The Role Of Speed Adaptation In Collective Motion

L. Tillert[1] *(*lindatillert@online.de*), L. Gómez-Nava*[1,2]*(*luis.gomez@hu-berlin.de*), P. Romanczuk*[1,2]*(*pawel.romanczuk@hu-berlin.de*).*

(1) Institute for Theoretical Biology, Humboldt-Universität zu Berlin, 10115 Berlin, Germany.
(2) Excellence Cluster Science of Intelligence, Technische Universität Berlin, 10587 Berlin, Germany.

Collective motion in animal groups is among the most interdisciplinary areas of complex systems, where fundamental topics like self-organization and complex networks up to social complexity are studied. To achieve this goal, many computational models have been proposed to explain the collective patterns displayed by moving animal groups, like the Vicsek model [1] or the zone model [2]. However, the vast majority of models rely on the simplifying assumption of individuals moving with a fixed speed, and do not explicitly account for fundamental kinematic properties such as speed-dependent turning rate due to inertia. Here, using a generic agent-based model, we study the effect of the ability of individual agents to modulate their speed on coordination in collective group movement, where each individual can have a different preferred speed. We quantify the effects of the social interactions between individuals of the group on the collective motion by measuring global observables like the spatial structure of the group, the center of mass motion and the polarization. We demonstrate how the ability of individuals to adjust the speed may have a profound impact on the group structure and dynamics. We interpret the results in the context of predator evasion and explain why heterogeneity in the preferred speeds is beneficial for the group.

References

[1] T. Vicsek, A. Czirók, E. Ben-Jacob, I. Cohen, and O. Shochet, *Physical Review Letters* **75** 126, 1995.

[2] I. D. Couzin, J. Krause, R. James, G. D. Ruxton and N. R. Franks, *Journal of Theoretical Biology* **218** (1), 2002.

Modeling species population spread

N. Mukherjee[1], *S. Sinha*[2].

Computational Biology group,

The Institute of Mathematical Sciences,

4th Cross Street, CIT Campus, Tharamani, Chennai, Tamil Nadu, India

nayanam@imsc.res.in

Modeling population spread over the course of time and space is a growing topic of interest among researchers. The basic modeling approach using Fisher-KPP reaction-diffusion system for the single population situation along with introduction of Allee effect, density-dependent dispersal, time delay, advection etc. are discussed in [1]. In this talk I will discuss the existing modeling approaches and introduce a new approach of modeling human population spread along with relevant analytical and numerical simulation results.

Acknowledgements (optional)

The first author is funded from SERB-National Postdoctoral Fellowhip with File Number PDF/2020/001876 .

References (optional)

[1] J. Steele, *Human Biology, Human dispersals: mathematical models and the archaeological record.* 3, page no. 121--141, 2009.

AI MORPHING INTO A SCIENCE OF COMPLEX SYSTEMS

Kemal A. Delic, Jeffrey H. Johnson, *The Open University, Milton Keynes, Walton Hall ,* MK7 6AA, UK,
kemal.delic@open.ac.uk

The field of AI initially created controversy by its very word coinage, while surviving two winters of disillusionment and major cuts of funding. However, one may argue, it has inspired many new developments in the computer sciences. These days, we are witnessing, yet another wave of high interest, important investments, and unrealistic expectations of AI research. The original goal was to construct machines indistinguishable from or behaving similarly to humans - hard and soft AI. The current wave of AI has deviated considerably from these goals being based on technology that is weakly analogous to brains but in reality an ill-defined and dangerous ad-hoc blend of non-intelligent computational processes and human intelligence. Its successes and failures are clear from recommender systems that can hit commercial targets perfectly but be so absurd that their decisions would not be made by the least intelligent person. The dangers of the current wave illustrate people being unjustly and inexplicably disadvantaged by computer-says-no technology while others are persecuted by government systems to detect human emotions and thought crime. AI has lost its way. After nearly a century of research and experimentation the science of AI still has to jump the enormous gap between what computers cannot do and what humans can do with ease. We postulate here that AI research must evolve into a multidisciplinary science of complex systems. Humans are highly complex systems being explored by genetics, neuroscience, and psychology as sapient beings. When arranged into society, reliant on technologies, they become hyper complex systems of systems of systems that are very difficult to model, impossible to understand fully and highly unpredictable. Merged better with the life sciences, AI might be placed into appropriate scientific contexts, might bring some relevant insights and methods from computing, and ultimately become recognized as a science by itself - a science of complex systems. This important change of direction may require a long time to happen.

Diffusion geometry of multiplex and interdependent systems.

***G. Bertagnolli**[1,2], M. De Domenico[1]. (1) CoMuNe Lab, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo (TN), Italy, gbertagnolli@fbk.eu. (2) University of Trento, Department of Mathematics, Via Sommarive, 14, 38123 Povo (TN), Italy.*

Random walks (RWs) represent a powerful tool for the exploration of complex systems. Different dynamics, from random search to information exchange and synchronisation, can be modelled or approximated by RWs and, once they take place *on* networks, they provide both, new insights about their structure and function, and a geometric tool for their analysis [1,2]. In [3] we generalise the framework of diffusion geometry [1] to multilayer networks, which enable the representation of different types of interactions between pairs of units (through layers), and of the interdependencies of these interactions (inter-layer connections). This richer structure contributes to the functional shape of the system, as shown in Fig.1. Given the synthetic multiplex represented in (a), in (b) we first aggregate the layers and then compute diffusion distances, while in (c) distances are evaluated on the multiplex and are then aggregated (by parallel sum). Although both matrices display two main blocks, in (c) we can immediately match the functional clusters with the actual network structure: nodes 20–30 are weakly connected to the main module in layer 1 and are not connected to the first ten (but 8) nodes of layer 2. Hence, preserving the multiplexity of (a) and then aggregating the geometric information in a grounded manner—i.e. treating the distances between state nodes *(i,α)*, *(j,α)* in each layer *α* as parallel resistances between the physical nodes *i*, *j* and defining the (equivalent) distance between *i*, *j* in the same spirit the equivalent resistance is computed on electric circuits—provides a more accurate functional description of the mesoscale of the multiplex. Finally, in [3] we also extended the family of diffusion distances to different RW dynamics. In Fig.1(d), we show the supra-distance matrices w.r.t. the diffusive RW (DRW) and the physical random walk with relaxation, *r*, (PrRW). The transition probabilities in the DRW are normalised on the maximum multilayer out-strength over all nodes and layers, so that the random walker has a larger probability of remaining in state nodes with a small out-strength, while in the PrRW the switching between replicas, i.e. from *(i,α)* to *(i,β)*, does not cost time and the walker can jump to one of its neighbours in the same layer or in other layers, with probability *1-r* and *r*, respectively. It can be seen in Fig.1(d) that the two dynamics induce different geometries, hence the functional shape of the network depends not only on the structure, but also on the exploration dynamics and the two (structure and dynamics) cannot be analysed in separately.

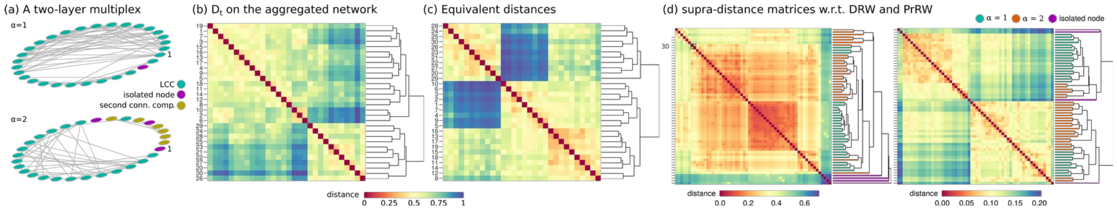


Figure 1: Diffusion distances on a synthetic multiplex. (a) A synthetic two-layer multiplex. (b)-(c) Distance matrices and (d) supra-distance matrices w.r.t different RW dynamics on (a).

References

[1] De Domenico, M. (2017). Phys. Rev. Lett. 118, 168301, arXiv.
[2] Boguna, M. et al. (2021). Nat. Rev. Phys. 3, 114–135, arXiv.
[3] Bertagnolli, G. and De Domenico, M. (2021). Phys. Rev. E, 103(4), 042301, arXiv.

Random walks on hypergraphs: from unbiased to maximum entropy random walks

Guilherme Ferraz de Arruda ^{1,*} and Yamir Moreno^{1,2,†}

^{1}ISI Foundation, Via Chisola 5, 10126 Torino, Italy
^{2}Institute for Biocomputation and Physics of Complex Systems (BIFI) & Department of Theoretical Physics, University of Zaragoza, 50018 Zaragoza, Spain
*`gui.f.arruda@gmail.com`

Random walks are paradigmatic, being interesting both from theoretical and practical points of view. Despite that, most of the study of these processes focuses on networks and, therefore, are limited to pairwise interactions. This assumption is often an oversimplification as many real processes cannot be faithfully represented only by pairwise interactions. To overcome this limitation, we use hypergraphs, which impose virtually no structural constraints and focus on two classes of random walks, the unbiased random walk (URW), where only local information is used by the walker, and the maximal entropy random walk (MERW), where global information is used to construct the transition probabilities. Formally, the walker in an unbiased random walk makes a succession of uniformly random decisions using only local information (the node degree). On the other hand, in a maximal entropy random walks (MERW), the walker does not make random decisions, and the randomness represents the lack of knowledge (1). In the MERW, the walker uniformly chooses a path that maximizes the entropy among all the possible paths of fixed length (2). The construction of such a process requires complete knowledge about the structure, here expressed by the leading eigenvector of the adjacency matrix. Comparatively, the first is a local process, while the second is a non-local dynamics. On top of that, due to the higher-order nature of hypergraphs, we can define more than one type of walk. Thus we focus on two extreme cases, the projected and what we call the higher-order step. In the first, the walker travels in a specific projection that does not differentiate the hyperedges cardinality, while the second accounts for it. Our construction of the random walk provides a different interpretation of the process if compared to (3), as we explicitly define the step as a sequence of two processes, similar to (4). By taking advantage of the network theory results, we extended the maximal entropy random walks and focused our analysis on comparing this random walk and the unbiased version. We focused our efforts on the stationary distributions and the hitting times. Another contribution of our work is our numerical experiments, in which we provide a series of examples, ranging from a small toy example, emphasizing key peculiarities of random walks in hypergraphs, to artificial and real cases. For the toy example, please see Fig. 1. In summary, our main contribution is the generalization of maximal entropy random walks to hypergraphs, revealing different behaviors and extending network results to random walks on hypergraphs. We establish our derivations from the observation that random walks on complex networks are equivalent to the same processes on hypergraphs, up to some small details and constrains (4). Moreover, we focused our efforts on the stationary distributions and the hitting times. We hope that our findings motivate further research extending the analysis to different classes of random walks and practical applications as the characterization of real systems and even the application in machine learning methods.

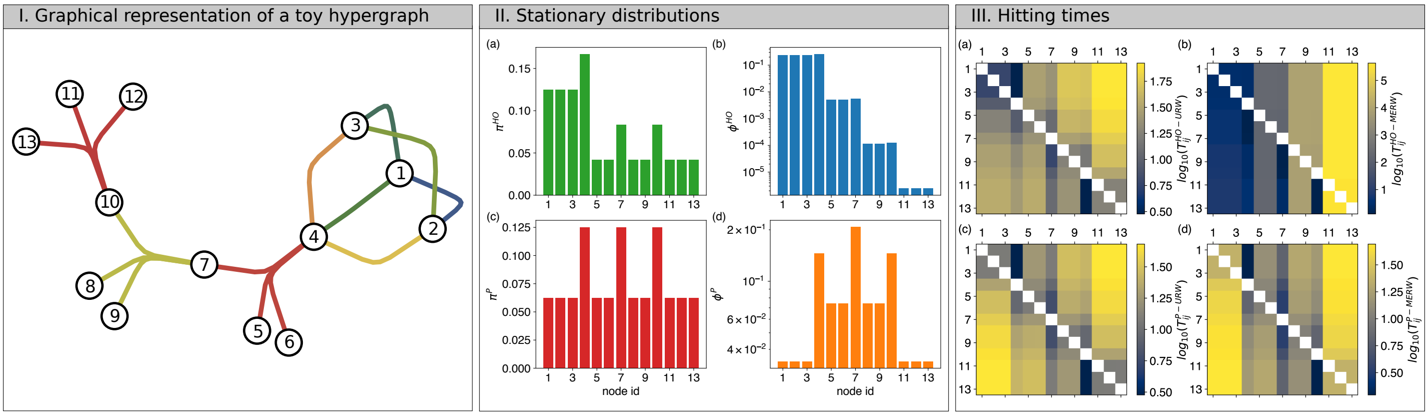


Figure 1: Panel I: A toy hypergraph, $\mathcal{H} = \{\mathcal{V}, \mathcal{E}\}$, with $N = 13$ nodes, $M = 9$ hyperedges, where $\mathcal{E} = \{e_1, e_2, ..., e_9\}$ and $e_1 = \{1, 2\}$, $e_2 = \{1, 3\}$, $e_3 = \{1, 4\}$, $e_4 = \{2, 3\}$, $e_5 = \{2, 4\}$, $e_6 = \{3, 4\}$, $e_7 = \{4, 5, 6, 7\}$, $e_8 = \{7, 8, 9, 10\}$, $e_9 = \{10, 11, 12, 13\}$. The hyperedges are color coded. Panle II: Stationary distribution for the unbiased random walk with the higher-order step in (a), the maximal entropy (MERW) with the same step in (b), the unbiased random walk (URW) with projected step in (c), and the unbiased random walk with the same step in (d). Panel III: Hitting times, T_{ij}, for the different types of random walk studied. In (a) and (c) for the unbiased random walk and in (b) and (d) for the maximum entropy random walk. Complementary, in (a) and (b) the higher-order step, while in (c) and (d) the projected step.

References

[1] J. Duda, Journal of Physics: Conference Series **361**, 012039 (2012).
[2] Z. Burda, J. Duda, J. M. Luck, and B. Waclaw, Phys. Rev. Lett. **102**, 160602 (2009).
[3] T. Carletti, F. Battiston, G. Cencetti, and D. Fanelli, Phys. Rev. E **101**, 022308 (2020).
[4] U. Chitra and B. J. Raphael (PMLR, Long Beach, California, USA, 2019), vol. 97, pp. 1172–1181.

Suppression of Thermoacoustic Oscillations through Self-Coupling

Sneha Srikanth[1], *Samadhan A. Pawar*[2], *Krishna Manoj*[2 a], *R. I. Sujith*[2]. *(1) Dept. of Mechanical Engineering, Indian Institute of Technology Madras, Chennai 600036, India, snehasrikanth3101@gmail.com. (2) Dept. of Aerospace Engineering, Indian Institute of Technology Madras, Chennai 600036, India.*

Suppression of limit cycle oscillations (LCO) through coupling has generated considerable interest due to its applicability in various domains ranging from electrical engineering to biology. One such application is the quenching of the detrimental high amplitude self-sustained oscillations in thermoacoustic systems. Recent theoretical and experimental studies have utilized the concepts from nonlinear dynamics and synchronization theory to study the oscillatory instabilities in single and mutually coupled horizontal Rijke tubes, which are prototypical thermoacoustic systems [1,2,3]. In the present study, we investigate the suppression of LCO in a mathematical model of a horizontal Rijke tube through self-coupling. Self-coupling can easily be implemented in a practical setup by connecting a Rijke tube oscillator to itself using a coupling tube. Correspondingly in the model, a delayed self-feedback term is included for this purpose [4,5]. Coupling parameters, such as coupling strength and coupling delay, and system parameters, such as the amplitude of the LCO prior to coupling, are varied in the model. With the systematic variation of these parameters, we provide the first evidence of suppression of thermoacoustic oscillations through self-coupling mechanism. Through numerically obtained bifurcation plots, we observe that the oscillations are most easily suppressed when their amplitude is low in the uncoupled state, for intermediate values of coupling strength, and for values of coupling delay that are roughly odd multiples of half-period of the oscillations. These trends are supported by the results obtained through analytical methods such as linear stability analysis and method of averaging. Linear stability analysis of the reduced order model of the system shows the interplay between the coupling and system parameters in determining the coupled dynamics of the oscillator. Using the method of averaging, we show that self-coupling suppresses LCO in the system by causing a shift in the Hopf point and thereby, increasing the stability of the steady state in the system.

Acknowledgements

S. S. is thankful to the support offered by the Young Research Fellow Program of Indian Institute of Technology Madras (Project ID: 202025), India. R. I. S. gratefully acknowledges the J. C. Bose fellowship (No. JCB/2018/000034/SSC) from the Department of Science and Technology (DST), Government of India for the financial support.

References

[1] P. Subramanian, R. I. Sujith, P. Wahi, *Journal of Fluid Mechanics,* 715 (2013) 210.

[2] S. Dange, K. Manoj, S. Banerjee, S. A. Pawar, S. Mondal, R. I. Sujith. *Chaos,* 29(9) (2019) 093135.

[3] N. Thomas, S. Mondal, S. A. Pawar, R. I. Sujith. *Chaos*, 28(3) (2018) 033119.

[4] F. M. Atay, *Journal of Sound and Vibration*, 218(2) (1998) 333-339.

[5] T. Biwa, Y. Sawada, H. Hyodo, S. Kato, *Physical Review Applied,* 6(4) (2016) 044020.

---

[a] *The author is currently at the Dept. of Mechanical Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, USA.*

Dynamics of Random Fuzzy Networks

O. Zapata[1], *M. Franco*[1], *C. Gershenson*[1], *D. A. Rosenblueth*[1]. *(1) UNAM, 04510 Ciudad de México, México, octavio_zapata@cuaieed.unam.mx.*

Random fuzzy networks are a generalization of random Boolean networks where the state of each node is a rational number between 0 and 1. Formally, a *fuzzy network* consists of a directed graph and a family of Boolean functions associated with the nodes of the graph. We use fuzzy logic to evaluate the functions replacing the Boolean operators AND, OR, and NOT with MIN, MAX, and $1-x$, respectively. The *average connectivity* of a network is the mean of the number of variables on which the function of each node depends. For investigating the passage from Boolean to continuous, we restrict the number of possible values that each node can take. This number is called the *base*: a fuzzy network base 2 is a Boolean network. A network evolves in discrete time by updating the value of all the nodes synchronously. We say that a network is *random* if the graph and the functions are chosen uniformly at random.

For studying the dynamics of a random fuzzy network, we simulate trajectories in its state-space and measure their entropy as follows. We start from a random state and let the network evolve $2T$ time steps. Then, we discard the first part of this trajectory and only keep the last T states. Along this sequence of length T, we consider the time series of the value of each node independently and calculate its entropy. Finally, we take the average entropy over all the nodes and repeat this process from a different initial state. We also consider the entropy (base 3) of the *transitions* between the states. These characterize the type of activity observed on the values taken by the nodes. If the value of a node decreases, does not change, or increases, the corresponding transition would be –1, 0, or +1, respectively.

The entropy of states in Figure 1 suggests that even though fuzzy networks allow additional values, the probability that a single node takes more than two values is very low. This is because if X is a random variable that takes values in a finite set A and Y is a random variable that takes values in a set B, and if there is an injective function f from A into B such that $P[X = a] = P[Y = f(a)]$ for every a in A (and so $P[Y = b] = 0$ for every b in B not in the image of f), then the entropy of Y is equal to the entropy of X. The entropy of the transitions confirms that the dynamical behavior of these networks is similar across different bases.

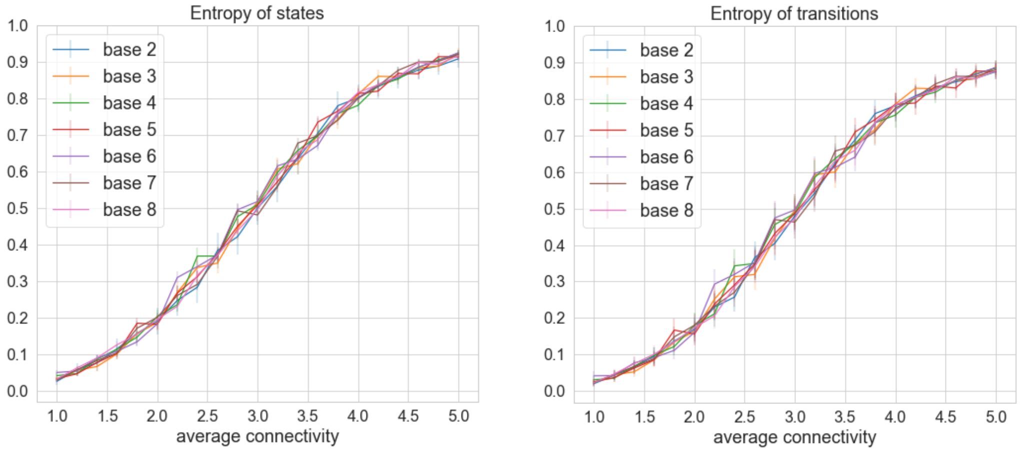


Figure 1: Mean average entropy obtained from an ensemble of 64 random fuzzy networks with 32 nodes, 1024 initial states, and $T = 256$.

Hopf bifurcation in SIHR dynamics

Arash Roostaie[1,2], Hadi Barzegar[1,2], Fakhteh Ghanbarnejad[1,3]. *fakhteh.ghanbarnejad@gmail.com*.

(1)Department of Physics, Sharif University of Technology, P.O. Box 11165-9161, Tehran, Iran
(2) Department of Mathematical Sciences, Sharif University of Technology, P.O. Box 11165-9161, Tehran, Iran
(3)Quantitative Life Sciences (QLS), The Abdus Salam International Centre for Theoretical Physics (ICTP), Strada Costiera, 11, I-34151 Trieste, Italy

In this work, the SIR (susceptible-infected-recovered) dynamics is extended by considering another compartmental which represents hospitalization of the critical cases. So a system of differential equations with four blocks is considered when there is intensive care unit (ICU) to cure critical cases. Outgoing rate of survived infected individuals is divided into *nI* and *bI/(b+I)*. The second term represents the rate of critical cases who enter ICUs. The system of differential equations, which describes this dynamic and definition of the parameters are shown below. S(t), I(t), H(t) and R(t) indicate the number of susceptible individuals, infected individuals, individuals of ICU and recovered individuals respectively.

$$
\begin{cases}
\frac{\mathrm{d}S}{\mathrm{d}t} = A - dS - \beta IS \\[2ex]
\frac{\mathrm{d}I}{\mathrm{d}t} = \beta IS - dI - \alpha I - nI - \frac{bI}{I+b} \\[2ex]
\frac{\mathrm{d}H}{\mathrm{d}t} = \frac{bI}{I+b} - dH - n'H \\[2ex]
\frac{\mathrm{d}R}{\mathrm{d}t} = nI + n'H - dR
\end{cases}
$$

parameter	definition
β	infectious rate
b	number of beds
A	birth rate
d	natural death rate
α	death rate caused by infectious
n	natural recovery rate
n'	recovery rate of individuals of ICU

We have proved that there are forward, backward and Hopf bifurcation in different regimes of parameters. Figure 1 shows where there is a limit cycle for two different initial conditions.

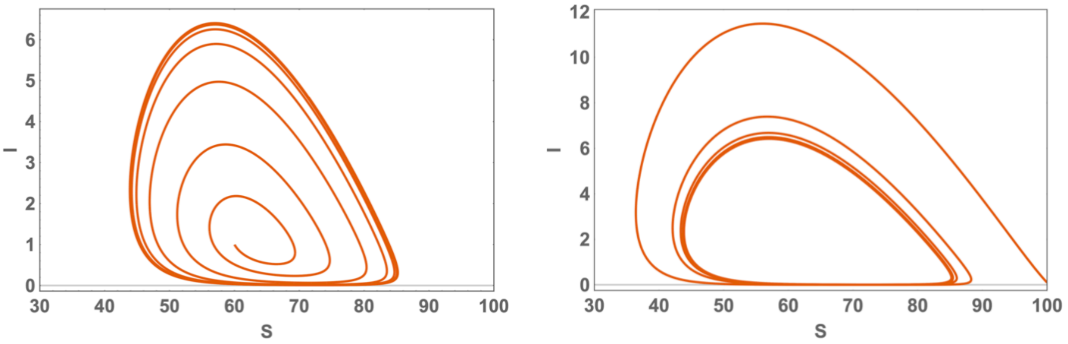


Figure 1: The stream plot *S-I* , when there is a limit cycle, while the parameters are $b = 1$, $A = 10$, $d = .1$, $\alpha = 1.9$, $n = 1$, $n' = 1$ and $\beta = .055$ and initial conditions ($X_0 = (S_0, I_0)$) are $X_0 = (60,1)$ for the left panel and $X_0 = (100,.1)$ for the right panel.

Simplicial contagion in temporal higher-order networks

Sandeep Chowdhary[1], Aanjaneya Kumar[2], Giulia Cencetti[3], Iacopo Iacopini[4], and Federico Battiston[1]

[1]Department of Network and Data Science, Central European University, Vienna
[2]Department of Physics, Indian Institute of Science Education and Research, Dr. Homi Bhabha Road, Pune 411008, India
[3]Fondazione Bruno Kessler, Trento, Italy
[4]Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France

Over the past decades, a variety of complex systems has been successfully described as networks whose interacting pairs of nodes are connected by links. Yet, from human communications to chemical reactions and ecological systems, interactions can often occur in groups of three or more nodes and cannot be described simply in terms of dyads. Until recently little attention has been devoted to the higher-order architecture of real complex systems. However, a mounting body of evidence is showing that taking the higher-order structure of these systems into account can enhance our modeling capacities and help us understand and predict their dynamical behavior.

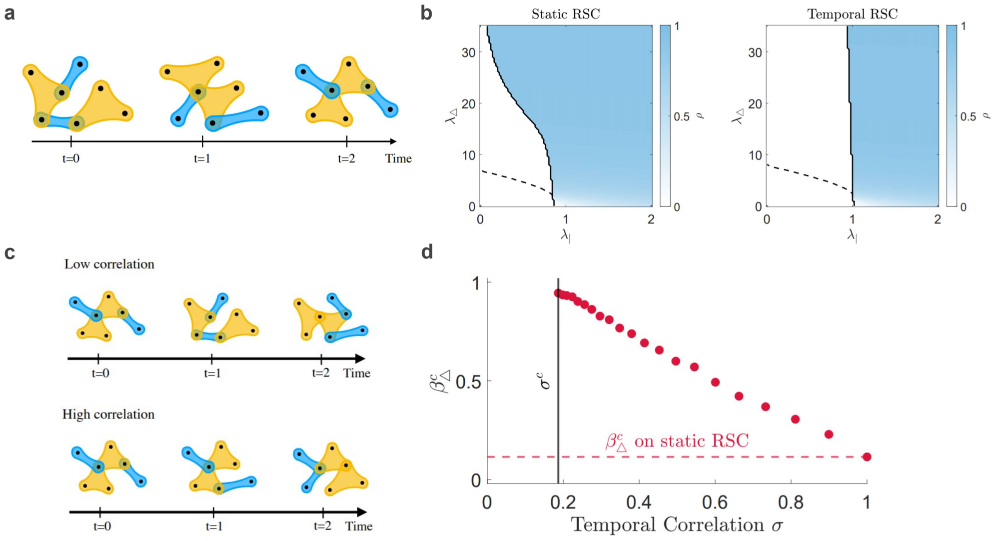


Figure 1: **a** Temporal higher-order networks. Schematic of a time-varying higher-order network where both pairwise and higher-order interactions evolve over time. **b** Contagion on static and temporal simplicial complexes. The fraction of infected nodes, ρ, at the equilibrium starting from a single infected node as a function of rescaled pairwise $\lambda_|$ and simplicial $\lambda_\triangle$ infection rates for static and temporal simplicial complexes. In the static case, the epidemic onset (solid black line) as a function of $\lambda_|$ is anticipated as we increase $\lambda_\triangle$. For time-evolving higher-order networks such effect is not observed, and we find a suppression of the endemic phase which can not be reached for low values of $\lambda_|$, independently on the value of $\lambda_\triangle$. As a consequence, the relevance of simplicial effects is strongly mitigated when we consider temporality, a realistic feature of many real-world social systems. **c** Temporal correlations in higher-order networks. Schematic of temporal simplicial complexes with low and high temporal correlations. **d** Effect of temporal correlations in higher-order networks. Critical simplicial infection rate computed as a function of the temporal correlation σ. The critical value $\beta^c_\triangle$ is higher for decreasing values of σ, and the epidemic threshold disappears below a critical value of temporal correlation σ^c (grey line). This indicates that simplicial effects are stronger in highly correlated higher-order networks.

A paradigmatic case is that of contagion processes. Recently a new model of simplicial contagion was proposed [1], where group interactions of any order are described by the simplices of a simplicial complex. The presence of higher-order interactions leads to explosive epidemic transitions and bistability which can not be obtained when only dyadic ties are considered, meaning that the long-term behavior of the system is determined by the size of the initial seed of infectious nodes. What are the feature of simplicial contagions when higher-order interactions change over time?

In this work [2], we provide a systematic answer to such a question. We extend simplicial contagion to temporal networks, where pairwise and higher-order simplices can be created or destroyed over time. Surprisingly, we find that by following a microscopic Markov chain approach the same number of initially infected agents might or might not lead to an endemic stationary state, depending on the temporal properties of the underlying network structure. In particular, we show that persistent temporal interactions anticipate the onset of the endemic state in finite-size systems. In the static case, as we increase rescaled simplicial infectivity $\lambda_\triangle$, the epidemic onset occurs for progressively smaller values of $\lambda_|$ in finite-size systems. This means that the seed of infectious nodes of fixed size $\frac{1}{N}$ belongs to the basin of attraction of the infection-free state for small values of $\lambda_\triangle$, while it moves to the basin of the endemic state upon increasing $\lambda_\triangle$. Coherently with the results obtained with the mean-field formalism [1], above a critical value of $\lambda_|$ ($\lambda_| = 1$ in the mean-field approximation) the system always reaches a non-zero fraction of infected agents which grows together with $\lambda_\triangle$. More interestingly, before this critical value, it is still possible to end up in the endemic state due to the higher-order contributions, but only if the seed of infectious nodes is big enough (critical mass). Surprisingly, by contrast, $\lambda_\triangle$ does not affect the onset of the epidemics in temporal simplicial complexes of finite size. Notice indeed that critical mass effects are completely suppressed, and below $\lambda^c_|$ the same seed of infectious nodes can never sustain the epidemics –as opposed to what happens in the static case for sufficiently high values of $\lambda_\triangle$.

Thus, introducing time-evolving structures can significantly impact contagion on higher-order networks, by altering the basin of the infection-free state in finite-size simplicial complexes. However, the way in which network structures evolve can be different. For instance, a social system may change more or less quickly, giving rise to different temporal correlations among networks at consecutive times. In order to investigate how the evolution of the network affects the spread of contagion, we introduce a model to systematically tune temporal correlations in simplicial complexes and find that observe that critical simplicial infection rate, $\beta^c_\triangle$ decreases monotonically with the temporal correlation σ and it takes its minimum value for maximally correlated RSCs, corresponding to a static simplicial complex. Finally, we characterize this behavior on heterogeneous simplicial complexes, showing that temporality again limits the effect of higher-order spreading, but in a less pronounced way than for homogeneous structures. Taken together, our work suggests the importance to incorporate temporality, a realistic feature of many real-world systems, when investigating dynamical processes beyond pairwise interactions.

---

[1] Iacopini, I., Petri, G., Barrat, A., Latora, V. (2019). Simplicial models of social contagion. Nature communications, 10(1), 1-9.

[2] Chowdhary, S., Kumar, A., Cencetti, G., Iacopini, I., Battiston, F. (2021). Simplicial contagion in temporal higher-order networks. arXiv preprint arXiv:2105.04455.

Ordering Dynamics and the Path to Consensus in Multi-State Voter Models

Lucía Ramirez, Maxi San Miguel, Tobias Galla*. Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Campus Universitat Illes Balears, E-07122 Palma de Mallorca, Spain, luciaramirez@ifisc.uib-csic.es*

We study the ordering dynamics in the multi-state voter model (MSVM)[1]. The system consists of N interacting agents placed on the nodes of a network. Each agent can be in one of the S possible opinion states (S ≤ N). The agents interact through an imitation process in which a randomly chosen agent adopts the opinion of a randomly chosen neighbor. We are interested mainly in the time-evolution of the density of active links ρ.

When averaged over realizations of the voter process, the density of active links in networks shows an initial decay after which the density stays in a plateau value, <ρ>=ξ, which we calculate within the pair approximation. In the thermodynamic limit the system remains in this metastable state. For finite N however, the system reaches one of the absorbing states eventually. We show that <ρ> decays exponentially to zero in this process and we obtain the characteristic time associated with this process analytically. We show that this time depends only on N, but not on S. For Erdős–Rényi (ER) graphs the time associated with this process scales as τ~N (Fig. 1a). Our analysis shows that the plateau values ξ depends both on S and on the mean degree of the network, *k* (Fig 1b).

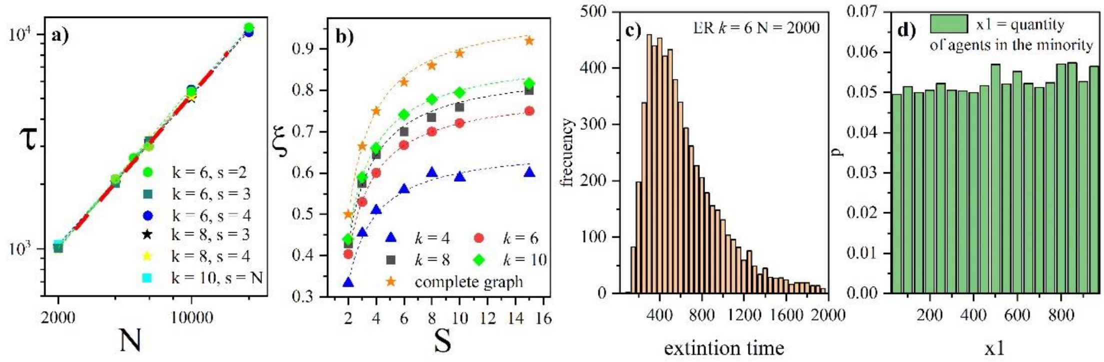


Figure 1: (a) Time scale τ of the exponential decay of <ρ>. The red line is our analytical prediction, symbols are from simulations on ER graphs. (b) Plateau value of average density of active links as function of S for different values of *k* as indicated. The dotted lines are our analytical prediction, symbols from simulations on ER graphs and for all-to-all interaction. (c) Distribution of time at which the first opinion disappears in. a model with S=3 states, (d) distribution of the proportion of agents in the minority surviving opinion when the first opinion disappears in a model on an ER graph with S=3 opinions initially.

For individual realizations of the MSVM, the density of active links ρ also initially decreases and then fluctuates around the average initial plateau ξ. Unlike the average density of active links, the dynamics in individual realizations show a sequence of intermediate plateaux, as opinions go extinct one by one. The different plateaux correspond to different partially ordered states with successively fewer surviving opinions. These intermediate plateaus are not seen in the ensemble average because the distribution of times for the extinction of opinions is relatively broad (Fig1c). We also find that the distribution of the number of agents in the surviving opinions is practically flat at the point where one opinion goes extinct (Fig1d).

References

[1] M.Starnini, et al, *Journal of Statistical Mechanics:Theory and Experiment*, (2012)P10027.

The consequences of hesitation: Axelrod model with intrinsic noise

R. Perrier, Y. Gandica, L. Hernández. Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS, CY Cergy Paris Université, France

Axelrod's model, initially introduced to study cultural dynamics, successfully explained how diversity arises from a dynamics based on homogenizing interactions [1]. In this model, agents are characterized by a discrete *cultural vector*, where each component, called *feature*, may take one of Q different values or *traits*. The original dynamics is such that when two agents had agreed in one particular feature they will never discuss about it again. Aware of the rigidity of such model, Axelrod himself suggested the interest of introducing noise into the system. Klemm et al [2] studied the model with an *external noise* controlled by an external parameter that gives the probability that an agent changes *spontaneously* the value of one of its traits. On the other hand, threshold models that control the homogenizing Axelrod's dynamics were also introduced. In particular, Gracia-Lázaro et al [3] consider the possibility that two agents whose disagreements is larger than the threshold, could disconnect and rewire to another neighbourhood.

In this work, we introduce a model that contemplates a more realistic situation in between the rigidity of the original model and the plasticity of the network led by disagreement. In our model agents are allowed to *hesitate*, this is to revert the agreement already achieved on a given feature but without breaking the link with their neighbors. The hesitation probability is *coupled to the dynamics* with the rationale that an agent is more likely revise the agreement already reached with one neighbor about one feature if they still disagree in many others. This disagreement is materialized by the active agent changing its mind and adopting a new random value of a feature that was initially in agreement with the said neighbor. We explore different intensities of this probability of disagreement by weighting it by an external, global parameter χ that can take value between 0 and 1, denoted *sensitivity*, which represents how doubtful agents can be.

Our results show a non-monotonous behavior of the freezing time as a function of the noise intensity and the variability of traits, as illustrated in Fig. 1, where the color map shows that the region where the system does not reach a frozen state correspond to values of Q where the original Axelrod reaches consensus. Compared to previous works, these results point out the fact that when noise is coupled with the Axelrod's dynamics, its outcomes strongly depend on whether it affects the agents' opinions or the agents' neighborhoods.

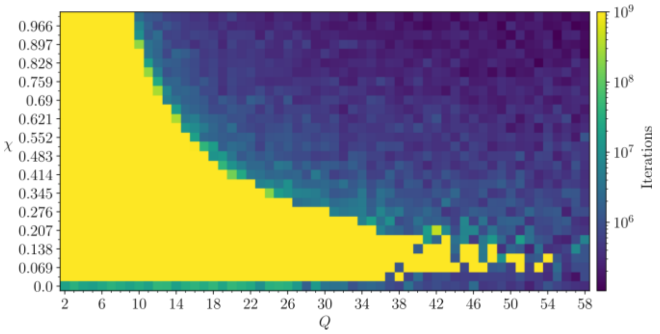


Figure 1: Phase plot: Influence of the noise sensitivity on the freezing time. In the solid yellow region, the system was not able to reach the frozen state and the dynamics was stopped at a threshold of 10^9 iterations. Due the high computation time, results are not averaged. N=40^2 agents, F=5, 1680 data points.

References

[1] R. Axelrod, Journal of Conflict Resolution 41, 203(1997).

[2] K.Klemm, V.M.Egíuluz, R.Toral, and M.S.Miguel, Phys. Rev. E **67** 045101(R) (2003).

[3] C. Gracia-Lázaro, F. Quijandría, L. Hernández, L. M. Floría, and Y. Moreno, Phys. Rev.E **84** 067101 (2011).

Life Worth Mentioning: Complexity in Life-Like Cellular Automata

***Eric Peña*[1] *and H. Sayama*[1,2]** *(1) Center for Collective Dynamics of Complex Systems, Binghamton University, Binghamton, NY, USA; (2) Waseda Innovation Lab, Waseda University, Tokyo, Japan; eric.pena@binghamton.edu.*

Cellular automata (CA) have been lauded for their ability to generate complex global patterns from simple local rules. The late English mathematician, John Horton Conway, developed his illustrious Game of Life (Life) CA in 1970, which has since remained the quintessential CA construction—capable of producing a myriad of complex dynamic patterns and computational universality. Life and several other Life-like rules were classified in the same group of aesthetically and dynamically interesting CA rules characterized by their complex behaviors. However, a rigorous quantitative comparison among similarly classified Life-like rules has not been fully established. Here we show that Life is capable of maintaining as much complexity as similar rules while remaining the most parsimonious rule [1]. In other words, Life contains a consistent amount of complexity throughout its evolution with the least number of rule conditions compared to other Life-like rules. We also found that the complexity of higher density Life-like rules which themselves contain the Life rule form a distinct concave density-complexity relationship whereby an optimal complexity candidate is proposed. Our results also support that Life functions as the basic ingredient for cultivating the balance between structure and randomness to maintain complexity in 2D CA for low and high density regimes, especially over many iterations. This work highlights the genius of John Horton Conway and serves as a testament to the timeless marvel he referred to simply as: Life.

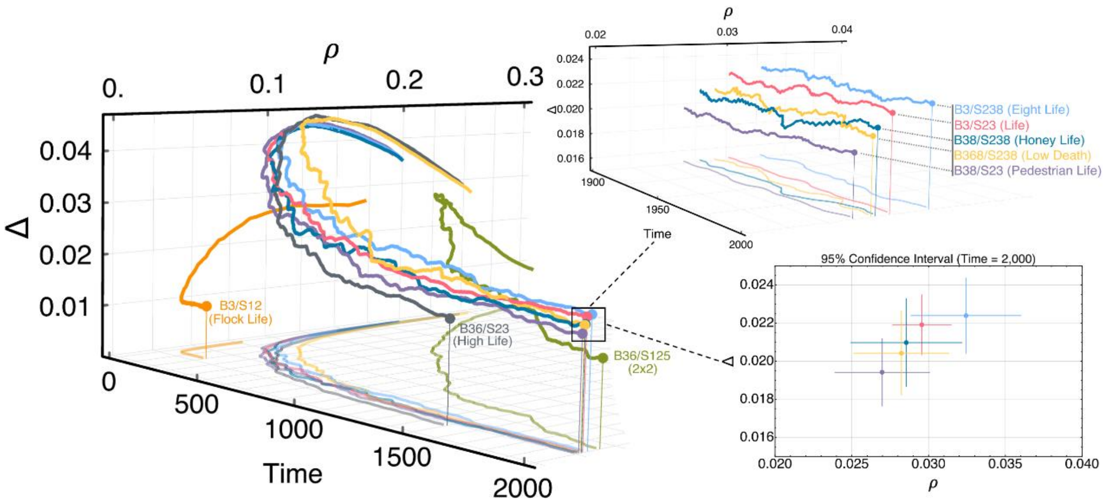


Figure 1: Temporal evolution of several low-density Life-like rules. Rules are specified in “B*x*/S*y*” notation, where *x* and *y* are the sets of numbers of live neighbor cells with which birth and survival occur, respectively. Density (ρ, ratio of live cells in space) and complexity (Δ, informational correlation between nearby cells [1]) are measured and averaged for each point in time over 2,000 time steps and 50 simulations for each rule. Life (B3/S23) is among the high complexity rules with the least number of rule conditions.

Acknowledgments

This work was supported in part by the Clifford D. Clark Diversity Fellowship at Binghamton University.

Reference

[1] Peña, E. & Sayama, H. (2021). *Artificial Life*, accepted pending revisions.

Generalized splay states: Dynamics on and in the vicinity of the splay manifold

R. Berner[1,2]*, S. Yanchuk*[1]*, E. Schöll*[2,3,4]*. (1) Institut für Mathematik, Technische Universität Berlin, Strasse des 17. Juni 136, 10623 Berlin, Germany, rico.berner@physik.tu-berlin.de. (2) Institut für Theoretische Physik, Technische Universität Berlin, Hardenbergstr. 36, 10623 Berlin, Germany. (3) Bernstein Center for Computational Neuroscience Berlin, Humboldt-Universität, Philippstrasse 13, 10115 Berlin, Germany. (4) Potsdam Institute for Climate Impact Research, Telegrafenberg, 14473 Potsdam, Germany.*

Networks of coupled phase oscillators play an important role in the analysis of emergent collective phenomena. We introduce generalized m-splay states constituting a special subclass of phase-locked states with vanishing mth order parameter. Such states typically manifest incoherent dynamics, and they often create high-dimensional families of solutions (splay manifolds). Moreover, these states have been found to be of importance for the phase space structure. For a general class of phase oscillator networks, we provide explicit linear stability conditions for splay states and exemplify our results with the well-known Kuramoto-Sakaguchi model. Importantly, our stability conditions are expressed in terms of just a few observables such as the order parameter. As a result, these conditions are simple and applicable to networks of arbitrary size. We generalize our findings to phase oscillators with inertia and adaptively coupled phase oscillator models and observe the emergence of additional complex dynamics within the splay manifold. The new degrees of freedom given by inertia or adaptivity give rise to multi(frequency)clusters (see Fig. 1) and other collective phenomena.

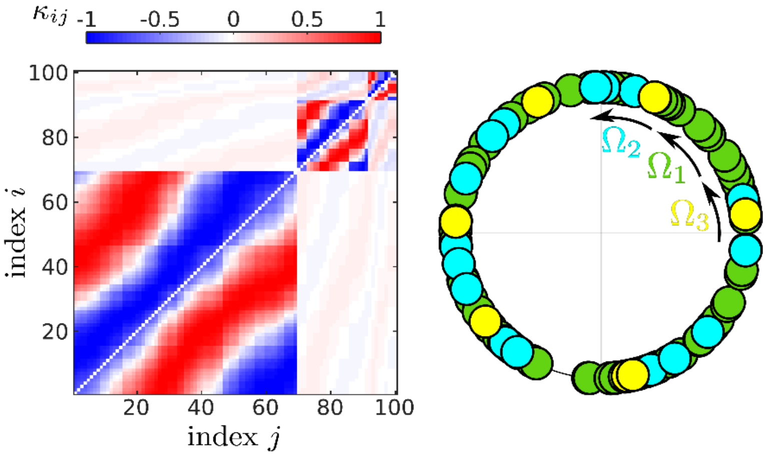


Figure 1: Multicluster that has formed with the generalized splay manifold. The figure shows (left) the adaptive network structure given by the coupling weights κ_{ij} and (right) the distribution of phase oscillators on the circle colored with respect to the cluster they belong to.

References

[1] R. Berner, S. Yanchuk, Y. Maistrenko, E. Schöll, Generalized splay states in phase oscillator networks, arXiv:2105.06694 (2021).

[2] R. Berner, E. Schöll, S. Yanchuk, Multiclusters in networks of adaptively coupled phase oscillators, SIAM J. Appl. Dyn. Syst. 18, 4, 2227 (2019).

[3] R. Berner, J. Fialkowski, D. V. Kasatkin, V. I. Nekorkin, S. Yanchuk, E. Schöll, Hierarchical frequency clusters in adaptive networks of phase oscillators, Chaos **29**, 103134 (2019).

Exploring Coupled Map Lattices with Physical Reservoir Computing

R. Shanaz[1]*, and P. Muruganandam*[1]*.*

(1) Department of Physics, Bharathidasan University, Tiruchirappalli, India, rasha@bdu.ac.in.

Reservoir computing (RC) is a state-of-the-art machine learning-based computing paradigm, that offers extensive capabilities to forecast spatiotemporally chaotic dynamics [1,2]. A typical reservoir computer consists of an input layer, a fixed reservoir, and a trainable readout layer. The state of the reservoir is the dynamical response to the encoded input. The reservoir output is mapped to the target output by the readout, which is the only layer that is trained. This allows for any physical dynamical system to be effectively used as a reservoir, instead of confining to recurrent neural networks [3]. This technique, called physical reservoir computing, paves the way for utilizing complex systems as reservoirs and has shown competent ability in various tasks [4]. Studying chaotic dynamics in spatiotemporally extended systems is crucial in understanding several systems occurring in nature, ranging from climate to cognition. The coupled map lattices (CMLs) serve as prototype models to explore spatiotemporal chaos and study complexity [5,6]. In this work, we aim to integrate the prediction ability of RC with the intrinsic computing capability of complex dynamical systems and try to improve the feasibility of CML for studying the naturally occurring spatially extended systems. We demonstrate the computation through a single node physical reservoir, made of a chaotic circuit and explore the effects of the system's dynamics on its performance in predicting spatiotemporal chaos. In this study, we implement the reservoir and train it for forecasting the spatiotemporal dynamics of CML. The performance of the reservoir is evaluated by the mean squared error between the test and target data.

Acknowledgements

This work was supported by Science and Engineering Research Board (SERB), India, under Grant No. CRG/2019/004059, DST-FIST under Grant No. SR/FST/PSI-204/2015(C), DST-PURSE, and MHRD-RUSA 2.0 (Physical Sciences).

References

[1] H. Jaeger, *GMD Report*, **148** (2001) 1.

[2] P. R. Vlachas, J. Pathak, B. R. Hunt, et al., *Neural Networks*, **126** (2020) 191.

[3] G. Tanaka, T. Yamane, J. B., Héroux, et al., *Neural Networks,* **115** (2019) 100.

[4] N. R. Shanaz, K. Murali, and P. Muruganadam, *Indian Academy of Sciences Conference Series* **3**(1), (2020) 45.

[5] K. Kaneko, *Chaos* **2**(3) (1992) 279.

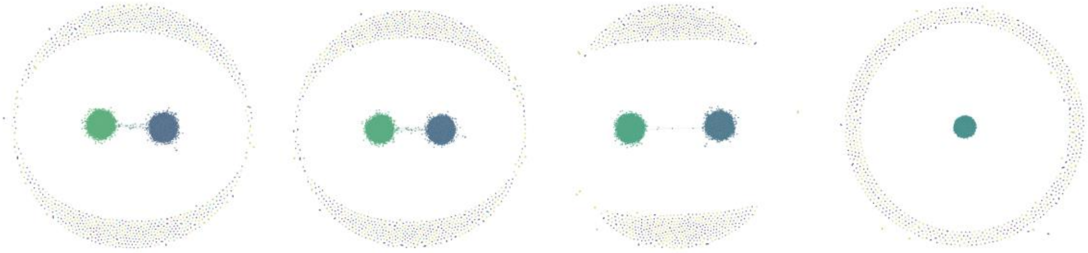
[6] P. Muruganandam, G. Francisco, M. Menezes and F. Ferreira, *Chaos Soliton and Fractals*, **41**(2), (2009) 997.

The bridges to consensus: Network effects in a bounded confidence opinion dynamics model.

H. Schawe, S. Fontaine, L. Hernández. Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS, CY Cergy Paris Université, France.

The role of heterogeneous confidences in the Hegselmann-Krause (HK) bounded confidence model has been extensively studied recently in a mixed population setting, revealing a complex phase diagram with a reentrant phase of consensus for confidence values where the homogeneous model leads to fragmentation [1]. However, another important source of heterogeneity comes from the interaction patterns. In a real society agents do not have the same a priori probability of interacting with any other agent, and it is well known that the outcomes of any diffusion process, may be different according to the fact that it takes place in networks or in a mixed population [2]. However, few studies of the HK model in networks exist, and given the richness found for the phase diagram of this model in a mixed population, the problem deserves more attention. One of these studies [3], focused on the condition the society to reach *unanimity* (all the agents having the same opinion) and concluded that it is more difficult to reach this state in a network of finite degree than in a mixed population. However, this situation is very unlikely in a society, where an opinion is considered to be consensual when it is shared by a large fraction of the population (but not necessarily by all the agents).

Here we present an extensive study of the HK model in different topologies, by computing the behavior of the order parameter compatible with consensus, the normalized size of the largest cluster of agents carrying the same opinion. We show that the HK model yields nontrivial results in network topologies. In particular, we show that randomness enhances consensus, which is at odds with the previous result. We show that the vanishing of the consensus threshold in the thermodynamic limit, observed in purely random networks is due to the existence of bridges, containing several nodes, between clusters of very different opinions that could not interact otherwise. The nodes in the bridge, through successive interactions allow the clusters to merge, hence the dramatic increase of the convergence time. This mechanism is absent from lattices either pure or perturbed by a Watts Strogatz-like algorithm with low rewiring probability.



Evolution towards consensus in a scale free network. Each panel shows the evolution of two clusters of agents holding different opinions (i.e., not all edges of the underlying graph are shown, but only those who can be "active" given that the agents at each extreme have opinions that differ in less than the confidence $\varepsilon = 0.14$) at $t = 100$, $t = 1000$, $t = 9000$, $t = t\to\infty$ (from left to right), for a system with $N = 16384$ agents. The colour of the points representing the agents indicate their current opinion. The bridge of the mediating agents is very well visible pulling the two large clusters together.

References:

[1] H. Schawe, L. Hernández, Scientific Reports volume 10, Article number: 8273 (2020)
[2] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. 81, 591(2009).
[3] S. Fortunato, International Journal of Modern Physics C16, 259 (2005)

Flow Stability for Dynamic Community Detection

Alexandre Bovet[1,2] *, Jean-Charles Delvenne*[1]*, Renaud Lambiotte*[2] *(1) ICTEAM, Université catholique de Louvain, 1348 Louvain-la-Neuve, Belgium. (2) Mathematical Institute, University of Oxford, Oxford, UK*
alexandre.bovet@maths.ox.ac.uk

Many systems respesented as temporal networks exhibit complex temporal dynamics due to the presence of different processes taking place simultaneously. Several works have generalized existing community detection methods for static networks to temporal networks (see [1]), but they usually rely on temporal aggregation over time windows, the assumption of an underlying stationary process. Here [2], we propose a novel method that consider random walks (RW) evolving on the network and restricted by the activation times of the edges. We derive quality functions that allows one to find partitions that best cluster the flow of random walker and provide a coarse-grained description of the temporal network evolution. We show that the temporal evolution of networks leads to potentially asymmetrical relations between vertices that can be captured by using two network partitions for a given time interval: the *forward* partition and the *backward* partition. We demonstrate the abilities of our method on several synthetic and real-world temporal networks [2], such as a contact network of free-living wild mice (Fig. 1).

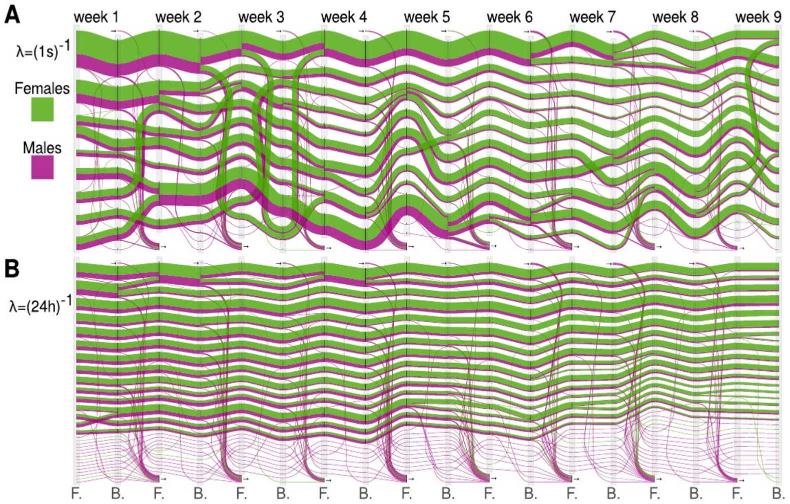


Figure 1: Flow clustering of a contact network of free-living wild mice. The community dynamics at a rate of (1s)−1 (**A**) reveals the existence of large communities during the first weeks corresponding to the end of February and beginning of March that split in smaller communities as spring arrives. At a RW rate of (24h)−1 (**B**) a finer description of the dynamics is revealed with the presence of smaller social groups with compositions and sizes that are very stable over the entire observation period.

References (optional)

[1] Rossetti, G. & Cazabet, R. Community Discovery in Dynamic Networks. ACM Computing Surveys 51, 1–37 (2018).

[2] Bovet, A., Delvenne, J.-C. & Lambiotte, R. Flow stability for dynamic community detection. Preprint https://arxiv.org/abs/2101.06131 (2021).

A Study of Collective Behaviour on Simplicial Complexes Using Higher-order Interactions

M. Chutani[1], *B. Tadic*[2,3], *N. Gupte*[1,4]. *(1) Indian Institute of Technology Madras, Chennai 600036, India, malayajac@physics.iitm.ac.in. (2) Jožef Stefan Institute, Jamova 39, Ljubljana, Slovenia. (3) Complexity Science Hub Vienna - Josephstadter Strasse 39, Vienna, Austria. (4) Complex Systems and Dynamics Group, Indian Institute of Technology Madras, Chennai 600036, India*

Studying collective dynamics in coupled dynamical systems has been an important field of study [1, 2]. The dynamics of the whole system depends on how the constituent oscillators are coupled with each other, which is represented by the structure of the underlying network. Most studies of collective dynamics consider, as part of their model, only pairwise interactions between oscillators, for which the Kuramoto model is most widely used. Works like [3, 4] have highlighted the importance of considering higher-order connections in networked systems. In the same vein, works like [5, 6] have used higher-order interactions to study synchronization in systems. Higher-order interactions give more information about aggregated local interactions. In our work, we study the effects on synchronization due to the structure of the underlying simplicial complex as well as higher-order interactions between oscillators. We find that pairwise interactions are most important in attaining synchronization. And, the introduction of higher-order Kuramoto interactions makes the transition between desynchronized state and synchronized state irreversible and leads to presence of hysteresis.

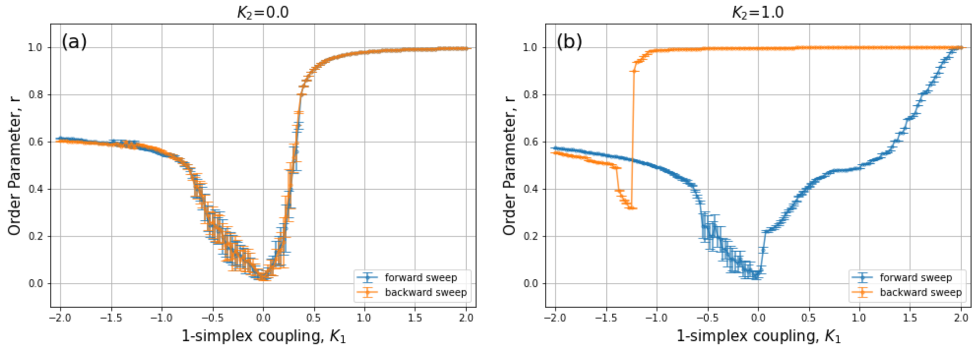


Figure 1: Order parameter $r = \left\langle \left| \frac{1}{N}\sum_{j=1}^{N} e^{i\theta_j} \right| \right\rangle$ as a function of pairwise coupling K_1 with (a) only pairwise interactions and (b) addition of 3-node interaction with coupling strength K_2=1.0. These calculations have been carried out on a 4-dimensional simplicial complex of size N=1000.

References

[1] A. Arenas, A. Diaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou, *Physics Reports*, 469, 93 (2008).

[2] F. A. Rodrigues, T. K. D. Peron, P. Ji, and J. Kurths, *Physics Reports* 610, 1, (2016).

[3] D. Mulder and G. Bianconi, *Journal of Statistical Physics*, 173, 783 (2018).

[4] A. Patania, G. Petri, and F. Vaccarino, *EPJ Data Science* 6, 18, (2017).

[5] P. S. Skardal and A. Arenas, *Phys. Rev. Lett.* 122, 248301 (2019).

[6] P. S. Skardal and A. Arenas, *Communications Physics* 3, 218 (2020).

Universal Relation for Electorate Involvement Disorder Amplitude in Second Round of Brazilian Presidential Election

M. Cardoso[1]*; L. M. C. Silva* [12]*; R. R. Neli*[1]*, W. E. Souza. (1) Programa de Pós-graduação em Inovações Tecnológicas, Universidade Tecnológica Federal do Paraná, 87301 - 899, Campo Mourão - Paraná, Brasil, magdacm@utfpr.edu.br. (2) Instituto Federal do Paraná, 86730-000, Astorga-Paraná, Brasil.*

The voter's opinion may be implicit among abstentions or blank and null votes and not just in the votes for one of the candidates. In this study, we investigated the electorate involvement, which we consider here, as the list of the proportions of the three possible forms of voter participation in the election result (abstention, blank or null vote, and valid vote), by using data from the second rounds of the last five Brazilian presidential elections. In order to obtain the main result of our work, we introduced a measure of the "disorder" of electorate involvement through of the concept of entropy used in Statistical Mechanics, relating the proportions that make up the involvement, per ballot box. Our approach was motivated by [1]. We found a general and parameter-free model for the relationship between the amplitude of the electorate involvement disorder and the electorate size per city,

$$S_A \approx 0.06\, log_{10} V\xi(V), \qquad (1)$$

where ς is a Gaussian noise with mean μ = 0 and standard deviation σ = 0.04, taken approximately as the mean of the means and the mean of the standard deviations of the fluctuations over the 5 elections, respectively, (Fig.1). Despite this result, we observed that the disorder in ballot boxes was "more intense" in 2018.

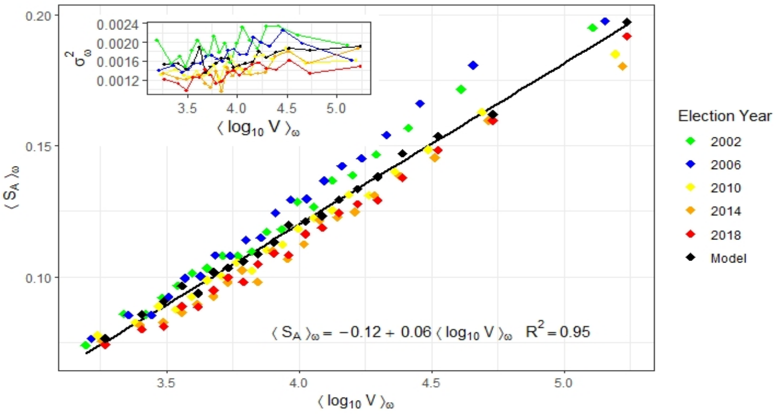


***Figure 1**: **Linear fit of the average values of the electorate involvement disorder amplitude.** Mean values of the variable S_A (electorate involvement disorder amplitude per city), denoted by $\langle S_A \rangle_\omega$, in 20 ω windows of the variable $log_{10} V$ with the same city numbers (V number of registered voters per city), versus the mean values, $\langle log_{10} V \rangle_\omega$, in ω. The black line represents the single linear fit, by using the least squares method, for all the years of election studied here, the equation of the line and the coefficient of determination are shown in the lower inset. Graph at the top inset shows the variance σ_2 of S_A in the 20 windows ω as a function of $\langle log_{10} V \rangle$. Black dots represent the average values in the same 20 windows ω for the data simulated by the model* (1).

Reference

[1] C. Borghesi, J. Chiche, J.-P. Nadal, Between order and disorder: a "weak law" on recent electoral behavior among urban voters, PLoS ONE 7 (2012) e 39916.

Emergent Dichotomy of Oscillations and Scale-free Avalanches in a Minimal Adaptive Neural Model near Criticality

__F. Lombardi__[1], G. Tkačik[2], D. De Martino[3]. (1) Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria, fabrizio.lombardi@ist.ac.at. (2) Institute of Science and Technology Austria, Am Campus 1, 3400 Klosterneuburg, Austria. (3) Ikerbasque, 48009 Bilbao,Spain.

Rhythmical cortical activity has long been recognized as a pillar in the architecture of brain functions. More recently, empirical observations of neuronal avalanches have shown that spontaneous activity is organized in sequences of scale-invariant cascades in which neuronal groups fire in patterns with no characteristic scales — a typical feature of systems at criticality. The coexistence of scale-free and scale-specific collective behaviors constitutes an intriguing dichotomy whose underlying mechanisms, dynamics, and functional role are poorly understood. A key question concerns the dynamical regime in which critical-like avalanche dynamics and oscillations can coexist in brain activity. Here we disucss a minimal, microscopic, and analytically tractable model that captures a wide spectrum of emergent phenomena in brain dynamics, from neural oscillations to extreme events statistics, and scale-free neuronal avalanches. The model is a non-equilibrium Ising-like network inspired by the self-adapting principles governing biological neural networks, and combines the possibility of microscopic network dynamics investigations with analytical mean-field solution. The model exhibits intermittent oscillations in the proximity of the critical point. Crucially, the critical point of the model retains the features of the standard Ising critical point, and scale-free neuronal avalanches simultaneously emerge. Importantly, our model is analytically tractable, and its parameters (only two) can be fully inferred from brain recordings. By exploiting this feature, we study the model in connection with the resting state activity of the human brain. We show that the inferred model dynamics closely captures brain dynamics across scales, from single sensor signal characteristics to collective behavior of extreme events and neuronal avalanches. Remarkably, we demonstrate that our inferred model parameters correlate with model-independent signatures of "closeness to criticality", suggesting that the coexistence of scale-specific (neural oscillations) and scale-free (neuronal avalanches) dynamics in brain activity occurs slightly below a non-equilibrium critical point at the onset of self-sustained oscillations.

Acknowledgements (optional)

FL acknowledges support from the European Union’s Horizon 2020 research and innovation program under the Marie Sklodowska-Curie Grant Agreement No. 75441.

An Equity-Oriented Rethink Of Global Rankings With Complex Networks Mapping Development

L. Bellantuono[1]*, A. Monaco*[2]*, S. Tangaro*[3,2]*, N. Amoroso*[4,2]*, V. Aquaro*[5]*, R. Bellotti*[6,2]*.*

(1) Dipartimento di Scienze mediche di base, neuroscienze e organi di senso, Università degli Studi di Bari Aldo Moro, 70124 Bari, Italy, loredana.bellantuono@ba.infn.it. (2) Istituto Nazionale di Fisica Nucleare, Sezione di Bari, 70126 Bari, Italy. (3) Dipartimento di Scienze del Suolo, della Pianta e degli Alimenti, Università degli Studi di Bari Aldo Moro, 70125 Bari, Italy. (4) Dipartimento di Farmacia-Scienze del Farmaco, Università degli Studi di Bari Aldo Moro, 70125 Bari, Italy. (5) United Nations Department of Economic and Social Affairs (DESA), Division for Public Institutions and Digital Government, 10017 New York NY, United States. (6) Dipartimento Interateneo di Fisica "M. Merlin", Università degli Studi di Bari Aldo Moro, 70125 Bari, Italy.

The development degree of a country can be viewed as the result of a plurality of social, economic, cultural, environmental and geopolitical factors, combined with different modalities. If, on one hand, it is appropriate to characterize in a clear and synthetic form the status of a country through general development indexes, on the other hand it is necessary to keep track of the complexity behind. Rankings constitute a popular tool to rate the performance of countries with respect to a specific indicator or a group of indicators [1]. However, rankings often represent a mere snapshot of the status quo of the considered countries, and do not take into account the development degree behind a given result and the heterogeneity of possible starting conditions. Novel instruments of complexity science can integrate conventional approaches in assessing the performance of a country or territory, and unveil hidden information. In this work, we propose a robust and insightful strategy to evaluate the performance of a country in rankings, starting from a novel frame of reference that mirrors both development status and similarities with other nations [2]. We adopt the machinery of complex network theory to partition world countries in network communities [3] characterized by a homogeneous development level. The complex network model, in which each node corresponds to one of the 193 UN Member States, is based on the similarity of World Development Indicator (WDI) values for different states. The communities identified from the network analysis provide not only a tool to group countries by development similarity, but also a way to reinterpret their positions in international rankings. After a quantitative check of the correlation between community membership and score in a given ranking, we identify countries whose performance goes beyond the expectations based on their development status, and countries that have the potential to reach, by increasing their efforts, the score of their community peers. We analyze the cases of five different rankings, focusing on different domains. The proposed methodology highlights the different facets of a country's condition, which are often the result of the actions undertaken and priorities established by its government. Development communities and rating criteria define a rigorous, transparent and reproducible procedure, that can have a pivotal value in development strategy planning, as well as to assess the effects of measures already adopted. The results provide the basis of a new equity-oriented interpretation of international rankings. *Disclaimer*: the views expressed in this abstract are the author's and do not necessarily reflect the opinion of UN DESA.

References

[1] A. Cooley, J. Snyder. *Ranking the World: Grading States as a Tool of Global Governance*, Cambridge University Press, Cambridge, 2015.

[2] L. Bellantuono et al. *Scientific Reports* 10 (2020) 18046.

[3] M. E. J. Newman. *Physical Review E* 69 (2004) 066133.

2 Psychology, Neurosciences, Linguistics

Contents

Predictive Processing as a Cybernetic Model of Cognition

***R. Aybar**. University of Vienna, Universitätsring 1, 1010 Vienna, Austria, raphael.aybar@univie.ac.at*

The status of the processing approach to cognition is currently under dispute. Scholars disagree on whether it constitutes a framework, a theory or a model of cognition. This paper argues that the approach is indeed a systems theory of cognition by tracing the historical influence of cybernetics in predictive processing. It suggests that Ashby's view of the brain as a regulatory system is in many respects equivalent to predictive processing' view of the brain as a predictive inferential engine. It shows that, for both accounts, the brain deals with complexity by exerting a form of predictive control. Optimal control, observable in adaptive systems, requires regulatory systems to generate models of themselves improved via error-driven learning. Crucially, if these constructs are components of a cybernetic theory, they should not be conflated with natural kinds.

Persistence of information flow: a multiscale characterization of human brain

B. Benigni,[1, 2] A. Ghavasieh,[1, 3] A. Corso,[1] V. D'Andrea,[1] and M. De Domenico[1]
[1] *CoMuNe Lab, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo (TN), Italy*
[2] *University of Trento, Department of Information Engineering and*
Computer Science, Via Sommarive, 9, 38123 Povo,Trento TN
[3] *University of Trento, Department of Physics, Via Sommarive, 9, 38123 Povo,Trento TN*

Information exchange in the human brain is crucial for a myriad of vital tasks, from the most automatic actions or involuntary reaction to memory retrieval of concepts, from how we experience the world to how we make decisions. In the last decades, network science reveals to be a powerful tool for modeling the structure and the function of human brain when regarded as a complex system. Specifically, neuroimaging tecniques are widely used for identifying the exchange of information between different areas of the brain, thus allowing for network reconstruction, either structural or functional, of the human brain. Commonly, the characterization of such reconstructed networks is relegated to compute a small set of network indicators (e.g., centralities, clustering, modularity,...) and their joint probability distribution. Recently, cutting-edge models such as multilayer networks has been proposed to accounting for multimodal measurements or correlations in the temporal and frequency domanis.

Here, we propose an innovative information-theoretic approach for the analysis of synthetic (based on some classes of generative models) and empirical brain networks and for the assessment of information capacity at different stages of dementia, namely Mild Cognitive Impairment and Alzheimer's disease. Specifically, capitalizing on the contribution of the whole network state encoded into a density matrix [1, 2], we define the spectral entropy of the brain structural network at different temporal scales by tuning a temporal parameter τ. To this aim, we use two distinct diffusive processes exchanging information among units: i) a classical random walk (CRW), where the walker has no global knowledge of the connectome and performs decisions based only on local knowledge of the connectivity, while keeping a uniform probability of choosing a connection for jumping; ii) a max-entropy random walk (MERW), where the walker has global knowledge of the connectome and jumps through a connection while keeping a uniform probability of choosing any trajectory on the network. Thus, our analysis is performed from two different perspectives: one where only local knowledge is used to explore the connectome (CRW) and one where global knowledge is used instead (MERW). Interestingly, we find that the spectral entropy of empirical connectomes lies between two generative models, indicating the possible mixed origin interpolating between Stochastic Block Model and Hyperbolic Model. We show that the mesoscale ($\sqrt{N} \leq \tau \leq N$, where τ is the Markov Time) is a suitable scale for distinguishing the differences (and the similarities) between empirical brains and the corresponding generative models. Furthermore, from the analysis of connectomes obtained from healthy and unhealthy subjects, we demonstrate that significant differences between healthy individuals and the ones affected by Alzheimer's diseases arise at the microscale (adjusted p-value $< 0.1\%$; maximum posterior probability $< 1\%$) and at the mesoscale (adjusted p-value $< 1\%$; maximum posterior probability $< 10\%$) [3].

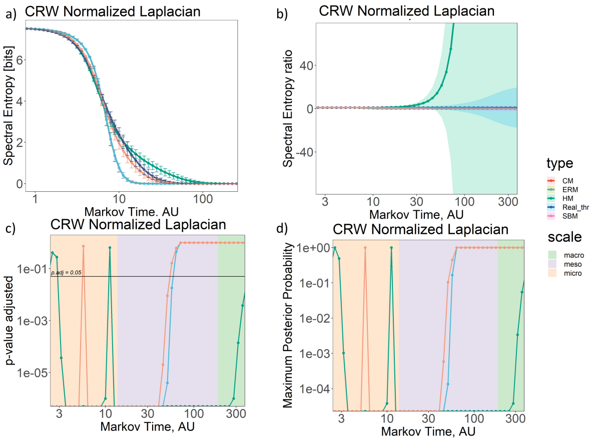


Figure 1. **Persistence of information flow in human brain and in generative models, CRW dynamics.** In a) we report the average value of spectral entropy varying the Markov time, τ, for the real data (blue line) and all the considered generative models (encoded with colored lines), i.e. Erdös-Rényi model (ERM), configuration model (CM), hyperbolic model (HM) and the stochastic block model (SBM), while in b) we show the ratio between spectral entropies of generative models and real data. In c) and d) we display the p-values adjusted and the maximum posterior probability recalibrated from the adjusted p-values, respectively. In these two cases, log-log scales are considered. Shaded areas represent in the order the micro ($\tau \leq \sqrt{N}$), meso ($\sqrt{N} \leq \tau \leq N$) and macro ($\tau \geq N$) scales. Note that ERM and CM overlap.

---

[1] M. De Domenico and J. Biamonte, Physical Review X **6**, 041062 (2016).
[2] A. Ghavasieh, C. Nicolini, and M. De Domenico, Physical Review E **102**, 052304 (2020).
[3] B. Benigni, A. Ghavasieh, A. Corso, V. D'Andrea, and M. De Domenico, Submitted (2021).

Simlandr: Potential landscape methods for psychological formal models

J. Cui[1]*, M. Olthof*[1]*, A. Lichtwarck-Aschoff*[2]*, F. Hasselman*[1]*, T. Li*[3]*. (1) Behavioural Science Institute, Radboud University, Montessorilaan 3, 6525HR Nijmegen, the Netherlands, jingmeng.cui@outlook.com. (2) Faculty of Behavioural and Social Sciences, University of Groningen, Grote Rozenstraat 38, 9712TJ Groningen, the Netherlands (3) LMAM and School of Mathematical Sciences, Peking University, Beijing 100871, China.*

Generalized potential landscapes have been used for investigating multivariate dynamic models in many fields. This method provides clear, quantitative information on the stability of states, thereby enriching our understanding of a specific dynamical system. The stability of states and phases are also important properties of formal models in psychology. However, the application of this method for psychological models is still in its infancy.

In this work, we developed an R package, *simlandr*, to provide a set of methods that are particularly useful for psychological modelers. In this package, the potential landscape is defined base on the steady-state distribution of the system and calculated from Monte-Carlo simulation. The tools for performing single or batch simulations, large data storage, checking simulation convergence, constructing 2D to 4D potential landscapes, and calculate paths for phase transitions will be introduced. The model of panic disorder by Robinaugh et al. (2019). will be used to illustrate the use of potential landscapes on showing the stability of states and phases, investigating the influence of parameters, and guiding model modification.

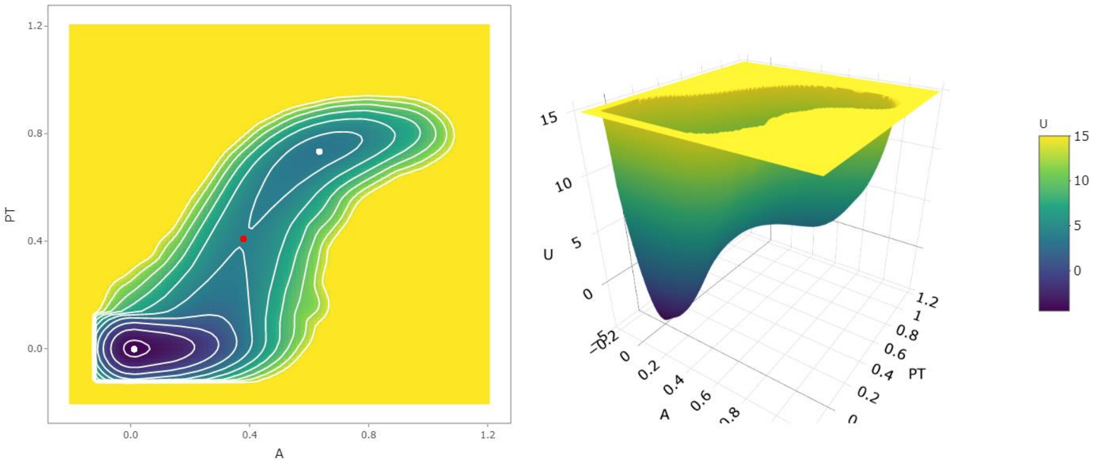


Figure 1: The potential landscape of the variables *physical arousal* (*A*) and *perceived threat* (*PT*) for the model of panic disorder. The white dots represent the local minima of the stable phases in the landscape, and the red dots represent the saddle point. This landscape is constructed using the simlandr package.

References

[1] J. Cui, *simlandr: Simulation-based Landscape Construction for Dynamical Systems, R package version 0.1.0*, https://github.com/Sciurus365/simlandr/.

[2] D. Robinaugh *et al.*, PsyArXiv, (2019) https://psyarxiv.com/km37w/.

[3] P. Zhou, T. Li, *The Journal of Chemical Physics*, 144, 094109 (2016).

Capturing aesthetic complexity in art using compression ensembles

A. Karjus[1]*, T. Ohm*[1]*, M. Canet Solà*[1]*, S. Ahnert*[2,3]*, M. Schich*[1]*. (1) Tallinn University, Tallinn, Estonia, andres.karjus@tlu.ee. (2) University of Cambridge, Cambridge, UK. (3) The Alan Turing Institute, London, UK.*

The quantification of visual aesthetics, including artistic expression, has a long history [1-3]. Previous research drawing on information theory has shown that visual complexity can be estimated with some accuracy using compression algorithms, but results based on human complexity ratings diverge as to which single approach is optimal [4]. We propose a novel method – compression ensembles – which consists of multiple compression lengths not only of the image itself but also its various transformations [5], embedded in a latent vector space.

We show that our method is cognitively plausible by demonstrating strong correlations with multiple visual complexity norms, and good performance on downstream tasks like artist identification. We then apply this approach to tens of thousands of paintings spanning half a millennium, probing established narratives in the historiography of Western art.

The proposed approach allows for capturing the "algorithmic fingerprints" of artworks and by extension artists. This allows us to compare artworks simultaneously in multiple dimensions of complexity and more generally, capture polymorphic family resemblance [6]. We argue that a generalized variant of this method would be also applicable to domains of cumulative culture beyond visual art. In particular, we see synergy with the growing body of literature in linguistics concerned with what is referred to as the complexity-expressivity trade-off [7,8].

References

[1] Birkhoff, G.D. (1933). Aesthetic Measure. Harvard University Press.

[2] Rigau, J., Miquel F., and M. Sbert (2007). “Conceptualizing Birkhoff’s Aesthetic Measure Using Shannon Entropy and Kolmogorov Complexity”. In: Proceedings of the Third Eurographics Conference on Computational Aesthetics in Graphics, Visualization and Imaging. pp. 105–112.

[3] Lee, B., M.K. Seo, D. Kim, I. Shin, M. Schich, H. Jeong, and S.K. Han (2020). “Dissecting landscape art history with information theory”. In: PNAS 117.43, pp. 26580–26590.

[4] Forsythe, A., M. Nadal, N. Sheehy, C. J. Cela-Conde, and M. Sawey (2011). “Predicting Beauty: Fractal Dimension and Visual Complexity in Art”. Br. J. Psychol 102.1.

[5] Tran, L.T., K Park, P. Lee, J. West, and M. Schich (2018). “Chromatic Structure and Family Resemblance in Large Art Collections – Exemplary Quantification and Visualizations”. In: 13th Annual International Conference of the Alliance of Digital Humanities Organizations, DH 2018, p. 679.

[6] Rosch, E. and C.B. Mervis (1975). “Family resemblances: Studies in the internal structure of categories”. In: Cognitive psychology 7.4, pp. 573–605.

[7] Kirby, S., M. Tamariz, H. Cornish, and K. Smith (2015). “Compression and Communication in the Cultural Evolution of Linguistic Structure”. In: Cognition 141.

[8] Kemp, C., Y. Xu, and T. Regier (2018). “Semantic Typology and Efficient Communication”. In: Annual Review of Linguistics 4.1.

Typical and Atypical Development in a Morphospace of Syntax Networks

L.F. Seoane[1,2,3], *V.M. Eguiluz*[3,4], *L Barceló-Coblijn*[5]. *(1) Centro Nacional de Biotecnología CSIC, C/ Darwin 3, Madrid, Spain. brigan@gmail.com. (2) Grupo Interdisciplinar de Sistemas Complejos, Madrid, Spain. (3) Instituto de Física Interdisciplinar y Sistemas Complejos, CSIC-UIB, Campus UIB, 07122, Palma de Mallorca, Spain. (4) Red Sea Research Centre (RSRC) and Computational Bioscience Research Center (CBRC), King Abdullah University of Science and Technology, Thuwal, 23955, Saudi Arabia. (5) Laboratori d'Investigació en Complexitat i de Lingüística Experimental (LICLE), Universitat de les Illes Balears, Spain.*

In human language, words in a sentence appear linked by syntax, forming a tree. Aggregating connections between words from several sentences, a “syntactic network” emerges. While it is invasive to access the brain circuits that produce language, syntactic networks capture the surface form of this faculty. We expect that this surface form inherits some of the traits and constraints of the corresponding neural circuitry. We derive such networks from transcripts of spontaneous speech of Dutch children. As these children mature, how does the developmental path of typical networks look like? How does it look like for cognitively impaired children (affected by Down’s syndrome, hearing impairment, or Specific Language Impairment)? Studying the largest cohort to date (~100 networks), we show how both typical and atypical development fall along a same path, with atypical cases stuck at intermediate stages (panel **a**). This strongly supports a unique developmental path for language that children follow as far as their cognitive abilities allow. Children with Down’s syndrome stop earlier – around an important transition in network topology. Other impaired children reach a next stage – past the transition, just short of the endpoint for normal development. We confirm earlier studies with much smaller cohorts (<5 participants) that had found the phase-like transition in typical language maturation (panel **b**) from small trees to cyclic networks with a solid core [1,2]. We relate topological measurements of our networks to computational properties of language. Our study hints us at mechanisms behind language maturation, suggesting aspects that should be reinforced to improve its development in cognitively impaired children.

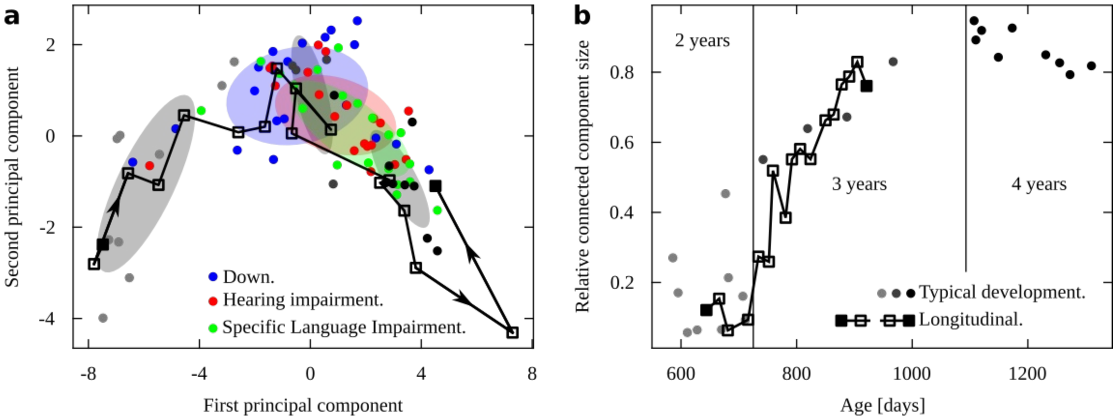


Figure 1: **a** Syntax networks in the Principal Component space of their topological properties. **b** Fraction of nodes in the largest connected component for typically developing chlidren.

References:

[1] B. Corominas-Murtra, S. Valverde, R. Solé. The ontogeny of scale-free syntax networks: phase transitions in early language acquisition. *Advances in Complex Systems* **12**(03), 371-392 (2009).

[2] B. Corominas-Murtra,M. Sànchez Fibla, S. Valverde, R. Solé. Chromatic transition in the emergence of syntax networks. *Royal Society open science* **5**(12), 181286 (2018).

Patterns of Cooperative Rhythm Production Between One Leader and Two Followers Through Auditory and Visual Information

H. Misumi[1], *T. Ogata*[2], *T. Kito*[1]. *(1) Waseda University, 3-4-1 Okubo, Shinjuku-ku, Tokyo, Japan, trijoc8833@akane.waseda.jp. (2) Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku Yokohama, Kanagawa, Japan*

In a musical ensemble performance such as an orchestra, people synchronize tempo and timing with others using visual and auditory information. The synchronization/rhythmic pattern creation problems have long been studied in cognitive psychology[1], and more recently, also in complexity science. However, cooperative rhythm production (i.e., multiple people produce rhythm together), especially among three people (and more), has thus far remain insufficiently studied. Three-persons synchronization is obviously not just a sum of two-persons ones. We conducted alternate tapping experiments, in which each group of three participants was instructed to tap a pressure sensor alternately to keep a constant rhythm. A group consisted of one leader and two followers, and the leader was asked to maintain the pre-indicated tempo, whereas the two followers simultaneously synchronized with the leader. A Leader's tap was presented to the followers as a visual or auditory stimulus, and a followers' tap was presented to the leader and the other follower as an auditory stimulus.

Our specific foci were on (i) how important the other follower's information for a follower is when syncing to a leader, and (ii) which kind of stimuli, visual or auditory, may be better for leaders to maintain the tempo.

Through the experiments, we found that the groups that managed to maintain tempo had mainly two particular patterns of interdependency between the leader and the follower (see Figure). Here, we consider that A was dependent on B when A was influenced by the previous taps of B and performed the next tapping to correct the timing shift with B.

Pattern (a), in which two followers are strongly dependent on a leader (= what was supposed to happen), was observed in both cases of when the leader's taps were presented as visual stimuli and when presented as audio stimuli. Interestingly, Pattern (b) (two followers are strongly dependent on each other) was also observed when visual stimuli were used. That is, even though two followers were not syncing so well with the leader, they could still manage to maintain the rhythm by syncing "locally" between the followers. This result suggests that, in an orchestra, it can be useful to match the timing first on a part-by-part basis, and then to match the overall timing by looking at the conductor.

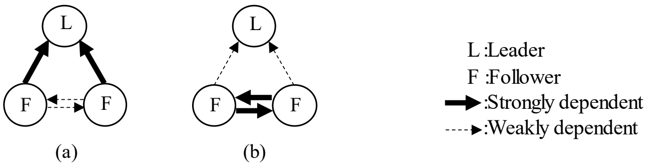


Figure: Dependency Patterns of Leader and Followers

Notes: (a) Two followers are strongly dependent on the leader. (b) Two followers are weakly dependent on the leader, but strongly dependent on each other.

Acknowledgement

This work was supported by JSPS KAKENHI Grant Number JP20K11903.

Reference

[1] B. H. Repp, *Psychonomic Bulletin & Review*, 12, 969--992, 2005.

Correlation Structure Between Brain Regions in Working-memory Tasks: fMRI Fractal and Spectral Analysis

__P. Oświęcimka__[1,2], A. Ceglarek[3], J.K. Ochab[1,4], M. Wątorek[1].

(1) Institute of Theoretical Physics, Jagiellonian University, Kraków, 30-348, Poland, pawel.oswiecimka@uj.edu.pl
(2) Complex Systems Theory Department, Institute of Nuclear Physics Polish Academy of Sciences, 31-342 Kraków, Poland
(3) Department of Cognitive Neuroscience and Neuroergonomics, Jagiellonian University, Kraków, 30-348, Poland
(4) M. Kac Complex Systems Research Center, Jagiellonian University, Kraków, 30-348, Poland

False memories are a topic that has enjoyed decades of fascinating research. With the advance of experimental techniques, insight into cognitive processes of short-term memory distortions has become possible [1]. One such technique allowing to measure neural activity while a person is memorising and retrieving information is functional magnetic resonance imaging (fMRI). In our study, we applied detrended fluctuation analysis to investigate fMRI data representing a diurnal variation of working memory [2] in four types of experimental tasks: two visual-verbal (based on lists of semantically or phonetically associated words) and two non-verbal (pictures of similar objects). The regional brain activity was quantified with the Hurst exponent and detrended cross-correlation coefficients [3].

Our analyses clearly show that the fMRI data obtained from most brain areas within a small-scale range can be regarded as 1/f type process identified in many physical, biological or even economic systems. However, the obtained characteristics of the signals in specific occipital lobe areas depend not only on the type of experimental tasks but also on the stage of the experiment, i.e. memorising the stimuli or information retrieval. A particularly apparent difference is visible between memorisation in verbal and non-verbal tasks. In the former case, for some brain regions in the Visual II resting-state network, the Hurst exponents assume values very close to 0.5, indicating a lack of linear temporal correlations in the signals. In contrast, we observe more persistent behaviour in the latter. The cross-correlations between brain areas are, too, indicative of differences in the processing of tasks and experimental stages. The detrended correlations turn out to be more sensitive than Pearson correlations, showing the greatest differences between the resting state and other tasks, between memorisation and retrieval and between verbal and non-verbal tasks, as well as other subtler results.

Acknowledgements

The study was supported by the Foundation for Polish Science co-financed by the European Union under the European Regional Development Fund in the POIR.04.04.00-00-14DE/18-00 project is carried out within the Team-Net programme.

References

[1] A.S. Atkins, P.A. Reuter-Lorenz, *NeuroImage* 56 (2011) 1726.

[2] K. Lewandowska, B. Wachowicz, T. Marek, H. Oginska, M. Fafrowicz, *Chronobiology International* 35 (2018) 80.

[3] J. Kwapień, P. Oświęcimka, S. Drożdż, *Physical Review E* 92 (2015) 052815.

Assessing the emotional structure of suicide letters and healthy recollections with cognitive network science

***M Stella**[1], T. Swanson[2], Y. Li[3], S. Talaga[4], T.T. Hills[5] and A.S. Teixeira[6], (1) CogNosco Lab, Department of Computer Science, University of Exeter, UK, massimo.stella@inbox.com. (2) Department of Psychology, University of Kansas, USA. (3) Max Planck Institute for Human Development, Germany. (4) Faculty of Psychology, University of Warsaw, Poland. (5) Department of Psychology, University of Warwick, UK. (6) Hospital da Luz Learning Health and INESC-ID, Portugal.*

Communicating one's mindset means transmitting complex relationships between concepts and emotions. Using cognitive network science, we reconstruct the mindset around suicide as communicated in 139 genuine suicide notes (cf. Teixeira et al., ArXiv, 2021).

Despite their negative context, suicide notes are surprisingly positively valenced and their ending statements are markedly more emotional, i.e. elicit deeper fear/sadness but also stronger joy/trust and anticipation, than their main body. By using emotional states from the Emotional Recall Task (Li et al., Journ. Exp. Psych., 2020), we "open the lid" of suicidal narratives and compare their emotional backbone against emotion recall in mentally healthy individuals.

Supported by psychological literature, we introduce emotional complexity as an affective analogue of structural balance theory, measuring how elementary cycles (closed triads) of emotion co-occurrences mix positive, negative and neutral states in narratives and recollections. As reported in Figure 1, the co-occurrence networks from both suicide authors and healthy individuals exhibit a higher emotional coherence than expected by chance. However, suicide narratives display higher complexity, i.e. a lower level of coherently valenced triads, than healthy individuals recalling the same states. Entropy measures identified a similar tendency for suicide letters to shift more frequently than recalls between contrasting emotional states.

Our results demonstrate that suicide notes possess highly contrastive narratives of emotions, more complex than expected by null models and healthy populations.

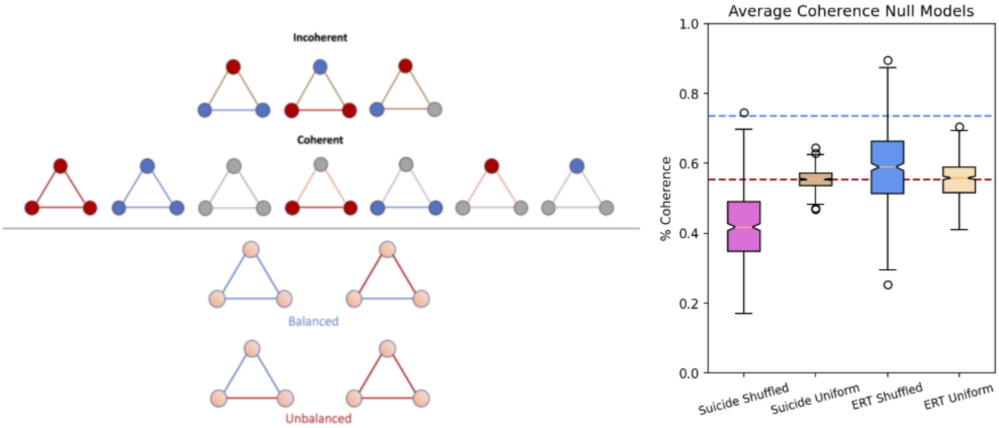


Figure 1: **Left:** The ratio of incoherent triangles defines emotional complexity, i.e. how many elementary cycles mix positive (blue), negative (red) and neutral (gray) emotional states (nodes) linked through co-occurrence patterns (links). This is different from structural balance theory (see the triangles below the horizonal line). **Right:** Ratio of coherent triangles in the suicide network (red dashed line) against recall in mentally healthy individuals (ERT, blue dashed line). Box plots are relative to random configuration models (Shuffled) and to uniformly random allocations of valence in emotional states (Uniform).

Automated Method for Detection of Chronic Insomnia with Motion Data

***M. Angelova*[1]*,* *S. Kusmakar*[1]*, Y. Zhu*[1]*, C. Karmakar*[1]*, S. Shelyag*[1]

(1) Data Analytics Research Lab, Deakin University, Melbourne Burwood Campus, Burwood, Victoria, Australia

maia.a@deakin.edu.au

Insomnia is characterised by the inability to fall asleep or stay asleep and/or waking too early and being unable to fall back asleep (DSM-5). It is associated with increased risk of diverse health problems and reduced quality of life. It affects more than 20% of the human population, has a strong negative impact on the productivity and efficiency at work, as well as on the cognitive and physical performance. It is a sleep disorder that remains under-detected and under-diagnosed. Early diagnosis of insomnia is essential for successful treatment and positive future outcomes. Depending on its duration, insomnia can be acute or chronic, with the transition point being three months. Clinical assessment of insomnia is based on sleep diaries [1], which are subjective and can increase the burden for the individual. Actigraphy is a non-invasive method for monitoring human rest/activity patterns. It is increasingly used to evaluate sleep in a natural home environment and complement the clinical assessment.

In this work we present a data driven method for classification of nocturnal awakening of individuals with chronic insomnia and healthy sleep controls. The method is a further development of our machine learning model for classification of acute insomnia [2] and is designed for actigraphy time series data. For this study we used actigraphy data collected continuously with wrist actiwatch (Respitronics Actiwatch Spectrum Pro) for one week from 40 individuals with chronic insomnia and 40 healthy bed partners. The data collection and protocol are described in [3]. We derived from the actigraphy time series data statistical and dynamical nonlinear features and features equivalent to sleep parameters, in total 48 features. We developed a two-level machine learning model using these features. The model distinguished chronic insomnia from normal sleep with accuracy above 80%. We also evaluated the minimal number of nights of actigraphy data collection for accurate classification. Our method does not require sleep diaries or other subjective information from the individuals. The model provides the first step towards the development of screening tool for low cost detection of chronic insomnia in home living conditions.

References

[1] D. Riemann, C. Baglioni, C. Bassetti, et al, *J Sleep Research*, 26 (2017) 675.

[2] M. Angelova, C. Karmakar, Y. Zhu, S.P. Drummond, J. Ellis, *IEEE Access*, 8 (2020) 74413.

[3] E. Walters, A. Phillips, A. Mellor, K. Hamill, M. Jenkins, P. Norton, D.H. Baucom, S.P. Drummond, *Sleep*, 43(1) (2020) zsz206.

Recent results in the Bayesian naming game

Gionni Marchetti, Marco Patriarca, Els Heinsalu.
National Institute of Chemical Physics and Biophysics, Rävala 10, 10143 Tallinn, Estonia
gionni.marchetti@kbfi.ee.

We study the basic naming game model [1] and the recently introduced Bayesian naming game model [2], in which the name learning process is described more realistically within a Bayesian learning framework. We focus on the dynamics of the bilinguals population and show that in the Bayesian model the number of bilinguals is always lower with respect to the basic naming game model. We provide some analytical estimates of the upper bound for the number of bilinguals in both models and validate the estimate through extensive numerical simulations.

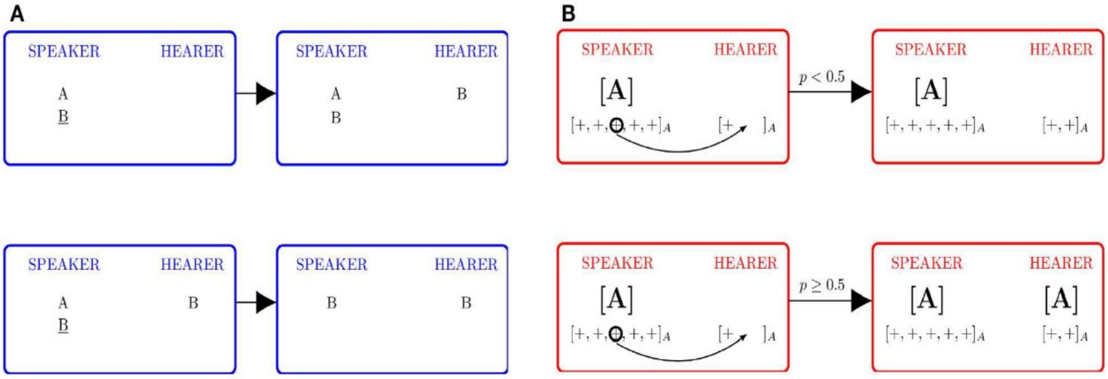


Figure 1: Comparison of the interaction scheme between the speaker and the hearer in the basic (A) and Bayesian (B) Naming Game Model.

Acknowledgements

The authors acknowledge support from the Estonian Ministry of Education and Research through Institutional Research Funding IUT39-1, the Estonian Research Council through Grants PUT1356 and PRG1059, and the ERDF (European Development Research Fund) CoE (Center of Excellence) program through Grant TK133.

References

[1] G. Marchetti, M. Patriarca, E. Heinsalu. *A bird's-eye view of naming game dynamics: From trait competition to Bayesian inference,* Chaos 30, 063119 (2020) https://doi.org/10.1063/5.0009569

[2] G. Marchetti, M. Patriarca, E. Heinsalu. *A Bayesian Approach to the Naming Game Model*, Front. Phys. 13 February 2020. https://doi.org/10.3389/fphy.2020.00010

Intentionality of emotions enables to detect the 'emotions of the crowd'

O. Mokryn[1]*, D. Boddof*[2]*, J. Lanir*[1]*. (1) Information Systems, University of Haifa, Israel, ossimo@gmail.com. (2) Management, University of Haifa, Israel.*

In literary and philosophy studies, emotions are generally defined as having a target, or "intentionality". The common wisdom in the computational social sciences is that the sentiment and emotions expressed in online reviews are targeted at the reviewed item, may it be a book, camera, or a film. The review reflects the consumer's opinion towards the reviewed item. Focusing on the film industry, we show that reviews also reflect the emotions elicited by a movie. That is, emotions that are targeted at the (fictional or non-fictional) reality that is being depicted in the film, e.g., trust in the hero, or fear of the monster, or grief in response to a character's death. Films are considered "emotion machines" and are made to create an emotional experience. Research shows that emotional experiences make us want to share them with others [1]. We then hypothesize that online opinions in the form of reviews for films reflect the emotions that were evoked by the film.

We systematically conduct a set of analyses. We show that a film's emotional signature, calculated in the manner we propose below, (a) makes intuitive sense; (b) is similar between sequels; (c) correlates with the emotions that subjects in an MTurk (Mechanical Turk) experiment report experiencing; (d) is similar to other movies of similar genre; (e) can predict what genre a movie belongs to and (f) can predict two film success indicators - rating and gross.

The method for constructing an emotional signature takes a bag-of-words approach and utilizes keywords in conjunction with the NRC lexicon of emotional words [2] to construct a per-movie emotional vector from all the reviews written for that film. All the experiments were conducted over a very large dataset of reviews obtained from IMDb (2937 films with at least 30 reviews each to a total of 717498 reviews).

The full set of experiments and results is available in [3].

We demonstrate that online reviews express the shared emotional experience. Our results indicate that a collective signal of the 'experience of the crowd' can be assembled from online opinions. While we have successfully demonstrated it for the case of reviews for films, it is reasonable to assume that the result can be extended to other forms and venues for online sharing of opinions.

References

[1] Rime´ B, Philippot P, Boca S, Mesquita B (1992) Long-lasting cognitive and social consequences of emotion: Social sharing and rumination. European Review of Social Psychology 3(1):225–258

[2] Mohammad SM, Turney PD (2013) Crowdsourcing a word–emotion association lexicon. Computational Intelligence 29(3):436–465

[3] Mokryn, O., Bodoff, D., Bader, N., Albo, Y., Lanir, J. (2020). Sharing emotions: determining films' evoked emotional experience from their online reviews. Information Retrieval Journal, 23, 475-501.

Quantitative methods for the analysis of judicial decisions

L. Font-Pomarol[1], *A. Piga*[1,3], *M. Sales-Pardo*[1,] , *R. Guimerà*[1,2] , *S. Nasarre-Aznar*[3,4]*(1) Dept. Chemical Engineering, Universitat Rovira i Virgili, 43007 Tarragona, lluc.font@urv.cat. (2) Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, (3) Dept. Private, Procedural and Financial Law, Universitat Rovira i Virgili, 43002 Tarragona (4) Càtedra UNESCO d'Habitatge, Universitat Rovira i Virgili, 43002 Tarragona*

Systematic and quantitative analysis are key features of any evidence-based research. In the scope of legal studies, it has been stressed the need for such analysis so to be able to make objective claims about the case law [1] and thus contribute to improve the legal practice [2]. In the present work, we present a series of quantitative methodologies aiming to study the case law by scrutinizing data from judicial decisions so to reveal systemic patterns and pillar decisions, trying to shed light on the different aspects behind the legal reasoning and the judiciary process.

First, we focus on the use that judges make of legislation to support their rulings. This legislation appears cited in decisions and can be seen as a trace of the legal reasoning behind each ruling. As observed in studies related to science of science [3,4], where an unconventional use of previous knowledge tends to lead to science success and to boost knowledge, we expect uncommon or innovative use of legislation to be associated with the decisions relevance in the case law. Using a network science approach, we evaluate law citations in a decision by quantifying the degree of innovation they carry out. Aiming to characterize several facets of innovation, we define it considering either the propensity to cite rare combinations of laws or pioneering a specific citation strategy for the first time. This methodology allows us to obtain singular decisions in the case law and to assess which aspects of judicial innovation are related to.

Second, we focus on the decisions discourse. Using a network approach to infer topics in a corpus of documents [5], we measure time variations of them, analyzing the historical discursive evolution, finding trends and important changes. Specifically, by means of information-theory based measures of surprise [6], we are able to quantify the dynamism of a particular corpus of judicial decisions. When applied to decisions in the scope of housing issues in the Spanish judiciary system, our approach is able to detect the extent to which the judiciary responded to the consequences derived from the 2007 financial crisis, showing our method as very suitable to measure the impact of the most sensitive social issues in the judicial system.

References

[1] W. Baude, A. S. Chilton, A. Malani, *The University of Chicago Law Review*, 84, 37, 2017.

[2] A. R. Pah, D. L. Schwartz, S. Sanga, Z. D. Clopton, P. DiCola, R. D. Mersey, C. S. Alexander, K. J. Hammond, L. A. Nunes Amaral, *Science* 134-136, 2020.

[3] B. Uzzi, S. Mukherjee, M. Stringer, B. Jones, *Science*, 342(6157):468–472, 2013.

[4] A. Rzhetsky, J. G. Foster, I. T. Foster, J. A. Evans, *Proceedings of the National Academy of Sciences,* 112(47):14569–14574, 2015.

[5] M. Gerlach, T. P. Peixoto, and E. G. Altmann, *Science Advances*, vol. 4, no. 7, 2018.

[6] A. T. J. Barron, J. Huang, R. L. Spang, S. DeDeo, *Proceedings of the National Academy of Sciences*, 115 (18) 4607-4612, 2018

TMS-induced brain connectivity modulation in Genetic Generalized Epilepsy

I. Vlachos[1], ***D. Kugiumtzis***[1],*V.K. Kimiskidis*·[1]. *(1) Aristotle University of Thessaloniki, Thessaloniki 54124, Greece,* *dkugiu@auth.gr*

Transcranial magnetic stimulation (TMS) combined with electroencephalography (EEG) is an effective stimulation approach used for therapeutic and diagnostic purposes in neurological disorders [1]. In epilepsy patients, TMS may result in the induction and modulation of epileptiform discharges [2]. We hereby investigate the modulatory effects of TMS on brain connectivity in Genetic Generalized Epilepsy (GGE) and explore their potential as a diagnostic biomarker in GGE. Patients with GGE (n=18) and healthy controls (n=11) were investigated with a paired-pulse TMS-EEG protocol. The brain network was studied at local and at global levels using Coherence as an EEG connectivity measure. Comparison of patients vs controls was performed in a time-resolved manner by analyzing comparatively pre- vs post-TMS brain networks. There was statistically significant TMS-induced modulation of connectivity at specific frequency bands within groups and difference in TMS-induced modulation between the two groups. The most significant difference between patients and controls related to connectivity modulation in the γ band at 1-3 sec post-TMS (p=0.004) (see Fig.1). These findings suggest that: a) TMS modulates the healthy and epileptic brain connectivity in different ways, b) TMS-EEG connectivity analysis can be a basis for a diagnostic biomarker of epilepsy. The analysis identifies specific time periods and frequency bands of interest of TMS-induced connectivity modulation and elucidates the effect of TMS on the healthy and epileptic brain connectivity.

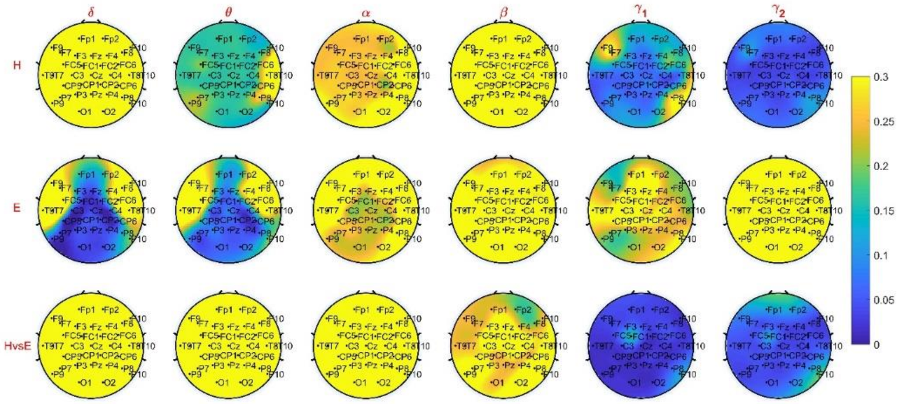


Figure 1: The p-values of the statistical significance of the TMS-induced modulation in local connectivity, in healthy (top row), epileptic (middle row) and of the difference between healthy and epileptic (bottom row), for the different frequency bands (columns) for TPOI2 1-3 s post-TMS.

Acknowledgements (optional)

This work was supported by the Action "Bilateral and Multilateral RTD cooperation between Greece and Israel" supported by GSRT and IIA under grant agreement number 122994/i4.

References (optional)

[1] R. J. Ilmoniemi, D. Kičić, *Brain Topography*, 22 (2010) 233–48I.

[2] D. Kugiumtzis, Ch. Koutlis, A. Tsimpiris, V. Kimiskidis. *International Journal of Neural Systems*, 7 (2017) 1750037.

A quantitative approach to track the evolution of topics in text corpora.

A. Piga[1,3], L. Font-Pomarol[1], M. Sales-Pardo[1], R. Guimerà[1,2], S. Nasarre-Aznar[3,4]

1. Dept. Chemical Engineering, Universitat Rovira i Virgili, 43007 Tarragona, 2. Institució Catalana de Recerca i Estudis Avançats, 08010 Barcelona, 3. Dept. Private, Procedural and Financial Law, Universitat Rovira i Virgili, 43002 Tarragona. 4. Càtedra UNESCO d'Habitatge, Universitat Rovira i Virgili, 43002 Tarragona

Angelo.piga@urv.cat

Characterizing how innovations are spread, stabilized and transmitted along years in a social context is a key inquiry in social science [1]. The evolution of language is a concrete instance of such a phenomenon and, nowadays, the increasing availability of corpus of digitalized texts allows quantitative approaches to the problem.

Here, we exploit a methodology capable to reveal patterns in the topics of different corpus of texts and their evolution in time.

Our study is based on a complex-networks approach to topic models [2]: a corpus of texts is encoded in a bipartite network, where the two parts are the documents and their words, respectively. A stochastic block model is inferred from this network and topic inference results analogous to a community-detection analysis, where documents belong to each topic with some probabilities. The historical evolution of discourse is then analyzed comparing the topics distributions between different epochs, by using the Kullback-Leibler divergence [3].

Concretely, our analysis is two-fold: on one hand, we analyze two different real datasets of judiciary decisions: the house-related decisions in Spain and the European Court of Human Rights decisions, both spanning over the last 20 years. On the other hand, we benchmark our techniques over a synthetic corpus of texts emerging from a precise generative model. In this way we are in particular able to test a clear procedure that allow to well interpret data, also in cases where corpus of documents is non-uniform, for example justifying extrapolations and scaling procedures when only few documents are available.

References

[1] B. Uzzi, S. Mukherjee, M. Stringer, B. Jones, *Science*, 342(6157):468–472, 2013

[2] M. Gerlach, T. P. Peixoto, and E. G. Altmann, *Science Advances*, vol. 4, no. 7, 2018.

[3] A. T. J. Barron, J. Huang, R. L. Spang, S. DeDeo, *Proceedings of the National Academy of Sciences*, 115 (18) 4607-4612, 2018

Linguistic laws in speech

Ivan G Torre[1], *B. Luque*[2], *L. Lacasa*[3], *C.T.Kello*[4,], *A. Hernandez-Fernandez*[5,6] *(1) Vicomtech, San Sebastian, Spain, ivangonzaleztorre@gmail.com. (2) ETSIAE, UPM, Madrid, Spain. (3) Queen Mary University of London, UK. (4) Cogsci Lab, UC Merced, USA. (6) IEC, Barcelona, Spain. (7) Institut de Ciencies de l'Educació, UPC, Spain.*

Linguistic laws constitute one of the cornerstones, quantitatively measurable, of modern cognitive sciences and linguistics, and have been intensively researched during last century, mainly in written corpora. Physical manifestations of linguistic units include sources of variability due to factors of speech production which are by definition excluded from counts of linguistic symbols. In this work, we examine whether linguistic laws hold with respect to the physical manifestations of linguistic units using phonetically transcribed and aligned databases of acoustic recordings of spontaneous speech in spoken English (Buckeye corpus).

First, we verify with unprecedented accuracy that acoustically transcribed durations of linguistic units at several scales comply with a lognormal distribution, and we quantitatively justify this 'lognormality law' using a stochastic generative model. Second, we explore the four classical linguistic laws (Zipf's Law, Herdan-Heap's Law, Brevity Law and Menzerath–Altmann's Law) in oral communication, both in physical units and in symbolic units measured in the speech transcriptions, and find that the validity of these laws is typically stronger when using physical units than in their symbolic counterpart. Additional results include (i) coining a Herdan-Heap's Law in physical units, (ii) a precise mathematical formulation of Brevity Law, which we show to be connected to optimal compression principles in information theory and allows to formulate and validate yet another law which we call the size-rank law or (iii) a mathematical derivation of MAL which highlights a regime where the law is inverted.

We are able to measure and compare the agreement of each of these laws when measured in both physical and symbolic units. Our results show that these laws are recovered in the three languages under study but considerably more emphatically so when these are examined in physical units, hence reinforcing the so-called 'physical hypothesis' according to which linguistic laws might indeed have a physical origin and the patterns recovered in written texts would, therefore, be just a byproduct of the regularities already present in the acoustic signals of oral communication. Altogether, these results support the hypothesis that statistical laws in language have a physical origin.

References

[1] Torre, I. G., Luque, B., Lacasa, L., Kello, C. T., and Hernández- Fernández, A, On the physical origin of linguistic laws and log-normality in speech, Royal Society open science, 6(8), 191023

[3] Torre, I. G., Luque, B., Lacasa, L., Luque, J., and Hernández-Fernández, A. (2017). Emergence of linguistic laws in human voice. *Scientific reports*, *7*(1), 1-10.

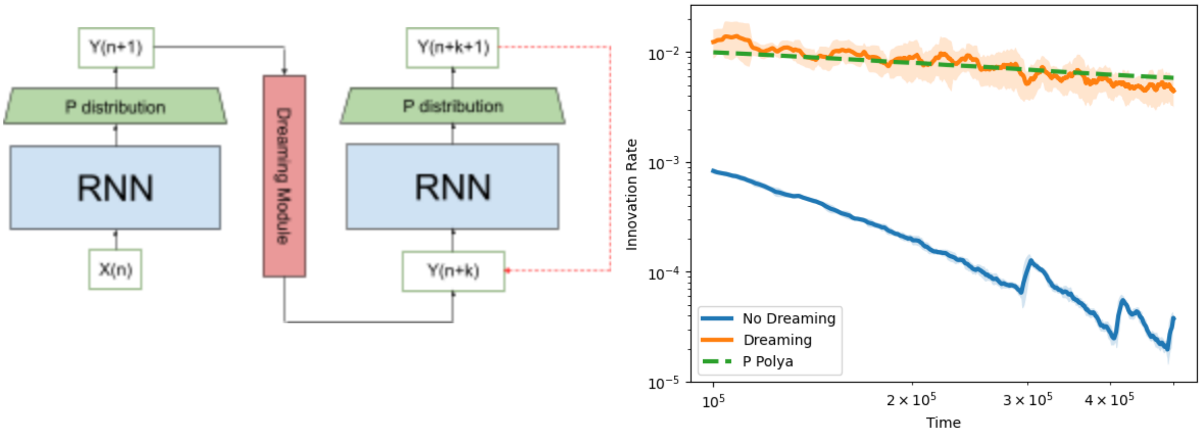
[4] Torre, I. G., Łukasz Dębowsk and Hernández-Fernández, A. Can Menzerath's law be a criterion of complexity in communication? Under review.

Dreaming Neural Networks

Alessandro Londei[1] and Vittorio Loreto[2]

[1]Sony Computer Science Laboratories, Paris, 6, rue Amyot, 75005, Paris, France alessandro.londei@sony.com
[2]Sapienza University of Rome, Physics Department, Piazzale Aldo Moro 2, 00185 Rome, Italy
[3]Complexity Science Hub, Josefstädter Strasse 39, A 1080 Vienna, Austria

Knowledge in cognitive systems extends when some novelty, or unexpected stimulus, occurs at the input [1]. These systems can explore the space of the possibilities autonomously, allowing for better predicting the unusual and adapting to environmental changes. However, though novelties can constantly appear as inputs of an artificial learning system, supervised and unsupervised approaches hardly cope with upcoming unexpected information. This feature often leads to catastrophic forgetting [2], i.e., the overwriting or the modification of previously stored information while learning new elements. Here, we introduce the idea of Dreaming Learning, an approach aimed at including novelties in the ongoing training of an artificial system. Dreaming Learning relies on the metaphor of dreaming, i.e., a generation of artificial realities while (almost) disconnected by environmental inputs. In the same way, we propose a mechanism through which the artificial neural system gets temporarily disconnected from the external world and let free of generating artificial training sets based on its current knowledge. This artificial information is then fed back in input as an additional training set. Learning on such artificially generated data allows the network to reinforce the system's preparedness for new information in the space of possibilities. We demonstrate the approach in the case of non-stationary, open-ended systems. We show how systems subject to Dreaming Learning can better explore the Adjacent Possible space [3] and be better prepared for the unexpected. This approach significantly improves the performances of the standard neural networks, both in terms of improved stability of the previously stored information and a better predisposition to forecast future events.



Left: Dreaming Learning system. Right: Innovation rate over time with and without Dreaming.

[1] Sutton, R. S. & Barto, A. G. *Reinforcement Learning: An Introduction* (The MIT Press, 2018), second edn.
[2] Kirkpatrick, J. *et al. Overcoming catastrophic forgetting in neural networks*. Proceedings of the National Academy of Sciences 114, 3521–3526 (2017).
[3] Kauffman, S. *Investigations* (Oxford University Press, 2000).

Domain-based Latent Personal Analysis and its uses for online Impersonation Detection

O. Mokryn[1]*, H. Ben-Shoshan*[2]*. (1) Information Systems, University of Haifa, Israel, ossimo@gmail.com. (2) Management, University of Haifa, Israel.*

Zipf's law defines an inverse proportion between a word's ranking in a given corpus and its frequency in it, roughly dividing the vocabulary into frequent words and infrequent ones. Here, we stipulate that within a domain, an author's signature can be derived from, in loose terms, the author's missing popular words and frequently used infrequent words. The missing, not conveyed, is an integral part of the whole. "Music is the silence between the notes", described Debussy; Freud coined negation as the dual process to affirmation, an expulsion from one's perception of reality. The part that is present, yet not recognized. In language, we claim, style is the silence between the words. Domain-popular words that are not used should contribute to personal characterization. We maintain that considering the words users choose to use and those they do not use would help identify their style. A choice to omit a domain-popular term is just as characteristic of personal style as a choice to include a domain-infrequent one.

We devise a method, termed Latent Personal Analysis (LPA), for finding domain-based attributes for entities in a domain: their distance from the domain and their signature, which determines how they most differ from a domain. We identify the most suitable distance metric for the method among several and construct the distances and personal signatures for authors, the domain's entities. The signature consists of both over-used terms (compared to the average) and missing popular terms. We validate the correctness and power of the signatures in identifying users and set existence conditions.

We then show uses for the method in explainable authorship attribution: we define algorithms that utilize LPA to identify two types of impersonation in social media: (1) authors with sockpuppets (multiple) accounts; (2) front user accounts, operated by several authors. We first validate the algorithms using labeled information, and then employ them over a large-scale dataset obtained from a social media site with over 4000 anonymous users. We corroborate these results using both qualitative analysis and temporal rate analysis. LPA can further be used to devise personal attributes in a wide range of scientific domains in which the constituents have a long-tail distribution of elements, ranging from music listening habits to the per-tissue immune system reaction.

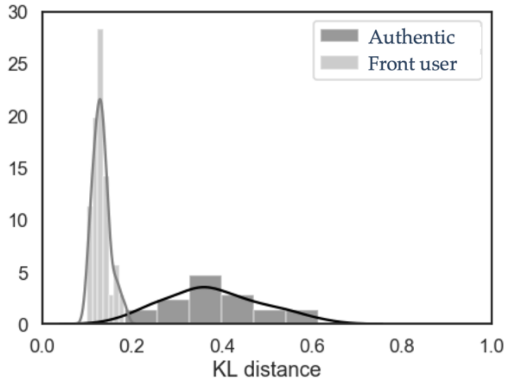


Figure 1: Utilizing LPA-distance to identify online *front-user* accounts

Effects Of Noise On Ageing Transitions In Network Of Rulkov Neurons

***D. Biswas**[1,2], Samana. P[2], S. Gupta[2,3].(1) Department of Physics, IIT Madras, Chennai, dhrubajyoti98@gmail.com (2) The Uncertainty Lab, Department of Applied Mechanics, IIT Madras, Chennai (3) Complex Systems and Dynamics Group, IIT Madras, Chennai.*

In this work, we study ageing transitions in a network of coupled Rulkov neurons[1]. Ageing Transitions are important, specially in biological and mechanical systems, where there is a natural tendency of the system to “degrade” - lose it’s natural dynamical behavior due to age or malfunctions. In such systems, it is desirable that they continue to function, even if sub-optimally, under such malfunctions and ageing effects. In our work, we consider a neuronal network, which is modeled as a complex system with an underlying network structure. The nodes in this network represent the neurons and the edges of the network represent the interactions/couplings between the neurons in the network. The dynamics of the neurons is chosen to be governed by the 2D Rulkov map model, which has distinct “active” and “inactive” states in it’s parameter space. We have proposed a symmetric mean-field coupling term for the interactions between the neurons and have further defined an order parameter to investigate the dynamical behavior of the system and to distinguish between qualitatively distinct dynamical states of the coupled system. We have introduced the effects of aging in our system by having a certain number of “inactive” neurons, which directly corresponds to the “age” of the network. By varying the probability of having “inactive” neurons in our networked system, denoted by P_d, we observe the existence of both *smooth* and *explosive/sudden* transitions for certain values of system parameters like the coupling strength and the network connectivity. We further look into the regions of the parameter space where the aforementioned transitions are seen and estimate numerically the boundaries of such regions. We further investigate how the presence of noise, modeled as uncertainties in our system, affects the transitions in terms of it’s critical onset value and the behavior of the system post the transition.

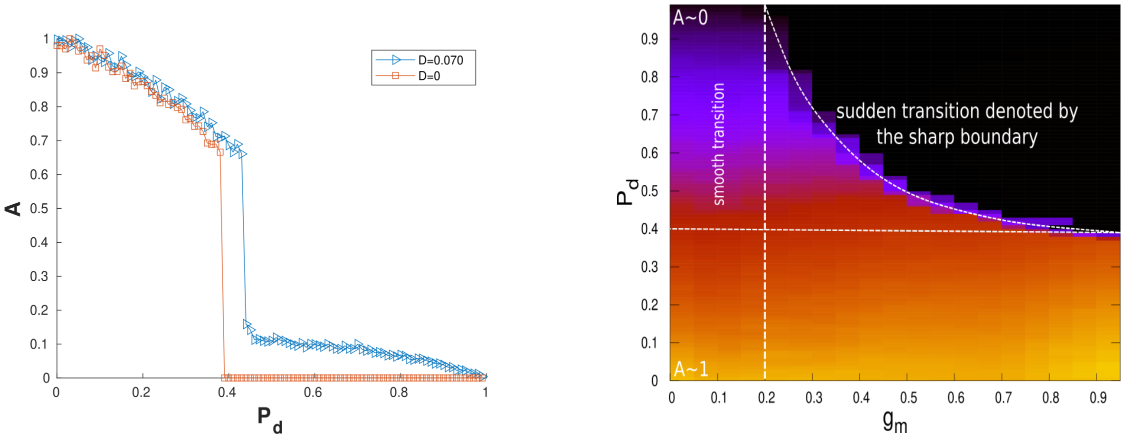


Figure 1: (Left) Variation of order parameter **A** as a function of $\mathbf{P_d}$ where **D** is the noise intensity; (Right) Regions in $\mathbf{P_d}$-$\mathbf{g_m}$ parameter space showing the sudden and smooth transition boundaries.

References:[1] Rulkov, N. F. Modeling of spiking-bursting neural behavior using two-dimensional map. Phys. Rev. E 65, 041922 (2002).

Directed Persistent Homology: testing sensitivity to brain network disorders

J. Hlinka[1,2], *L. Caputi*[1,3], *A. Pidnebesna*[1,2]. *(1) Institute of Computer Science, Czech Academy of Sciences, Pod vodarenskou vezi 271/2, 182 07, Prague, Czech Republic, hlinka@cs.cas.cz (2) National Institute of Mental Health, Topolová 748, 250 67, Klecany, Czech Republic (3) University of Aberdeen, Aberdeen AB24 3UE, Scotland, UK.*

Topological Data Analysis (TDA [1]), despite its relative novelty, has already been applied to study network connectivity structure across fields. We propose that its prominent tool of persistent homology (PH) may apart from the more common dependence networks (functional connectivity – FC) be applied also to directed, causal, networks – known as effective connectivity (EC) in neuroscience. We test the PH discriminatory power in two archetypal examples of disease-related brain connectivity alterations: during epilepsy seizures (captured by electrophysiology – EEG) and in schizophrenia patients (using functional magnetic resonance imaging - fMRI). We employ a range of PH-based features and quantify ability to distinguish healthy from diseased brain states by applying a support vector machine (SVM). We compare this novel approach to using standard undirected PH and raw EC/FC [2]. The fMRI data included the signal from 90 anatomical atlas regions, capturing 10 minute resting state brain activity from 90 schizophrenia patients and 90 healthy controls; the data were obtained using standard processing steps [2]. The scalp EEG data included signal from 23 electrodes, fs=256Hz) from 18 pediatric subjects with intractable seizures [https://physionet.org/content/chbmit/1.0.0/], providing 102 epileptic events of length 20 to 120 seconds and non-epileptic segments of the same length chosen randomly from each record. The intracranial EEG data from 16 patients with pharmacoresistant epilepsies [http://ieeg-swez.ethz.ch/] included recordings of the first 30 seconds of the first seizure ('ictal' data), and the 30 seconds segment starting 1 min before the first seizure ('preictal' data). The first minute of interictal data provides estimate of baseline connectivity subtracted from both ictal and preictal connectivities. Pearson correlation and Granger Causality (GC) provided estimates of FC and EC; selection of common PH features extracted and fed to linear SVM classification (healthy/patient from fMRI, ictal/interictal from (i)EEG). In the schizophrenia task TDA performed close to random, while raw connectivity reached 80% accuracy; likely due to topographical (rather than topological) specificity of the effects. In scalp EEG seizure task, PH worked comparably to raw connectivity. Specific niche for PH opens when direct comparison of connectivity matrices is unsuitable - such as for intracranial EEG with individual number and location of electrodes. Standard PH performed overall better than directed PH, likely due to imperfect estimation of EC by GC. To summarize, we show that directed PH can detect altered brain connectivity, however comparison against standard PH and the use of the raw FC/EC matrices reveals challenges to be overcome to fully utilize the theoretical promise of TDA and establish it as a common tool in neuroimaging.

Acknowledgements

The research was supported by the Czech Science Foundation project No. 21-17211S.

References

[1] G. Carlsson, Topology and data. *Bull. Amer. Math. Soc.*, 46 (2009) 255-308. doi:10.1090/S0273-0979-09-01249-X.

[2] L. Caputi, A. Pidnebesna, J. Hlinka. Promises and pitfalls of Topological Data Analysis for brain connectivity analysis. *BioRxiv*. 2021; doi:10.1101/2021.02.10.430469

A bounded confidence model of emotionally aroused integrate and fire oscillators.

I. Ferri[1]*, E. Cozzo*[1,2]*, A. Díaz-Guilera*[1] *L. Prignano*[1] *(1)UBICS irene.ferri@ub.edu (2) IN3-UOC*

The bounded confidence model introduced by Deffuant et al. [1] is a popular model of opinion dynamics in which actors have a [0,1]-valued opinion and interact only if their opinions differ by at most a deviation threshold, then both move closer in barycentric fashion governed by a confidence factor. A number of extensions have been proposed in order to go beyond mean-field approximation [2] or to include emotional dynamics [3,4]. Here, we propose a framework in which actors are prevented from interacting with those that differ in opinion (O) more than a deviation threshold determined by a decreasing function of their emotional arousal (EA). The higher the emotional arousal, the smaller the deviation threshold. Additionally, by interacting, they also influence each other's activity's timing. We model a population of integrate-and-fire oscillators in an O-EA plane. The internal phases of the oscillators increase uniformly with period T until they reach a maximum value of 1, when a firing event - here understood as a communication action - occurs and the phase is reset. Upon such an event, the firing oscillator influences those agents that have it within their interaction band (see Fig. 1, panel A), updating their phases by a factor (1+ε) and making them move in the O-EA plane towards the firing oscillator. The system evolves until all the interacting oscillators reach consensus, collapsing into single points in the plane and synchronizing their phases. We find that, in general, the stationary state is characterized by an increase of the average emotional arousal and the final number of consensus groups is larger than the initial number of connected components of the interaction graph. Moreover, depending on how easily the oscillators synchronize, i.e., on the parameter ε, both aspects may be reinforced or smoothed (see Fig. 1, panel B and C).

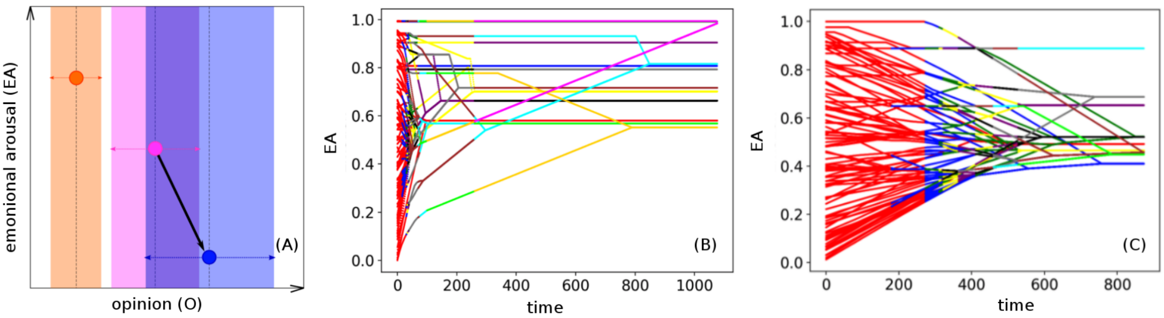


Figure 1. Panel A. Interaction rule for a minimal setting of 3 oscillators. Panel B. Evolution of the EA coordinates in a system of 200 oscillators with random initial conditions and ε = 0. Panel C: The same as in Panel B, but in the opposite scenario of synchronized initial phases.

References

[1] G Deffuant, D Neau, F Amblard, G Weisbuch - Advances in Complex Systems, 2000.

[2] G Weisbuch *The European Physical Journal B* 38.2 (2004): 339-343.

[3] P Sobkowicz - *Frontiers in Physics* 3 (2015): 17.

[4] Schweitzer, T Krivachy, D Garcia - Complexity, 2020

3 Economics and Finance

Contents

Games in Rigged Economies

L.F. Seoane[1,2,3]*. (1) Departamento de Biología de Sistemas, Centro Nacional de Biotecnología (CSIC), C/ Darwin 3, Madrid, Spain. brigan@gmail.com. (2) Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain. (3) Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Campus UIB, 07122, Palma de Mallorca, Spain.*

Modern economies evolved from simpler human exchanges into very convoluted systems. Today, a multitude of aspects can be regulated, tampered with, or left to chance; these are economic *degrees of freedom* (DOF) which together shape the flow of wealth. Economic actors can exploit them, at a cost, and bend that flow in their favor [1]. If intervention becomes widespread, microeconomic strategies of different actors can collide or resonate, building into macroeconomic effects. How viable is a ‘rigged’ economy? How do growing economic complexity and wealth affect it? Here we capture essential elements of ‘rigged’ economies with a simple model [2]. Nash equilibria of payoff matrices in tractable cases show how increased intervention turns economic DOF from minority into majority games through a dynamical phase. These stages are reproduced by agent-based simulations of our model (Figure 1), which allow us to explore scenarios out of reach for payoff matrices. Increasing economic complexity is then revealed as a mechanism that spontaneously defuses cartels or consensus situations. But excessive complexity enters abruptly into a regime of large fluctuations that threaten the system’s viability. This regime results from non-competitive efforts to intervene the economy coupled across DOF, becoming unpredictable. Thus high economic complexity can result in negative spillover from non-competitive actions. Simulations suggest that wealth must grow faster than linearly with complexity to avoid this large fluctuations regime and keep economies viable in the long run. Our model provides testable conclusions and phenomenological charts to guide policing of ‘rigged’ economies. We discuss the recent, real-world case of the Game Stop short-squeeze, in which multiple economic actors got coordinated through social media to invest in allegedly undervalued stocks. This resulted in an emergent upset of the stock market and still-ongoing investigations of market manipulation.

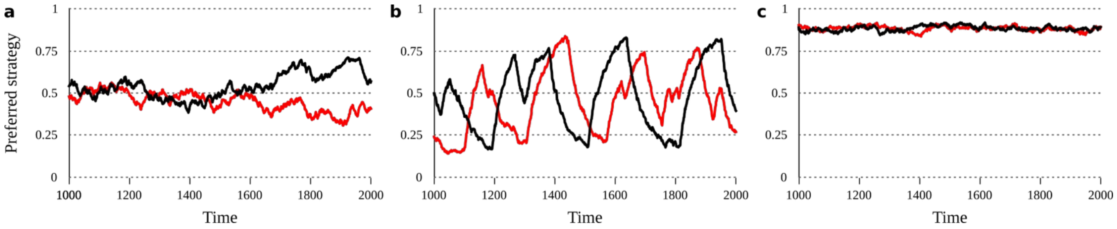


Figure 1: Three regimes with increased ‘rigging’ of 2 economic games. **a** With little intervention, economic DOF are minority games. As agents attempt to play the minority option, an evolving population splits itself equally (with random fluctuations) between the two strategies available. **b** Intermediate ‘rigging’ prompts structured fluctuations: each game is temporarily manipulated in a direction or the other. **c** Large ‘rigging’ turns economic DOF into majority games, which the population agrees to play and rig in the same direction.

References:

[1] D. Wolpert, J. Grana. How Much Would You Pay to Change a Game before Playing It? *Entropy*, **21**(7), p.686 (2019).

[2] L.F. Seoane. Games in rigged economies. https://arxiv.org/abs/2007.03886 (2020).

The effect of a public health crisis on economic outputs: The case of Covid-19 distress propagation and recovery

Jonathan M. Barlow[1,3] and Irena Vodenska[1,2]

[1] Laboratory for Interdisciplinary Finance and Economics (LIFE) Research, Boston University, Boston, MA, 02155

[2] Administrative Sciences Department, Metropolitan College, Boston University

[3] Department of Physics, Graduate School of Arts and Sciences, Boston University

Abstract

Wassily Leontief formulated the well-known input-output (I-O) model to characterize the interrelationships between sectors of an economy - the first matrix representation of an economy. In his 1977 paper *Understanding Business Cycles (1)*, Robert Lucas argues that as the representation of an economy becomes increasingly disaggregated, independent shocks to each sector average out, and aggregate volatility becomes increasingly negligible. However, this analysis did not take into account network dependencies. As shown by Acemoglu *et al. (2),* economic networks exhibit a much slower convergence of aggregate volatility than Lucas suggests. This work proposes a dynamic cascade model to investigate the systemic risk posed by sector-level industries within the U.S. inter-industry network. We then use this model to study the effect of the disruption presented by COVID-19 on the U.S. economy. We construct a weighted digraph $G = (V, E, W)$ using the industry-by-industry total requirements table for 2018, provided by the Bureau of Economic Analysis. We impose an initial shock that disrupts one or more industries' production capacity, and we calculate the propagation of production shortage with a modified Cobb-Douglas production function. We model the COVID-19 initial shock based on the unemployment between March and April 2020 as reported by the Bureau of Labor Statistics (BLS). King and Rebelo (3) show that the total hours worked in an industry are highly correlated with industries' total output. We use the total number of employees as a proxy for total hours worked. The fractional loss in inputs determines the propagation of unemployment shocks. The industries within the network are assigned a resilience that determines the ability of an industry to absorb input losses, such that if the rate of input loss exceeds the resilience, the industry fails, and its outputs go to zero. We observe a critical resilience, such that below this critical value, the network experiences a catastrophic cascade resulting in total network collapse. Lastly, we model the economic recovery from June 2020 through March 2021 using BLS data.

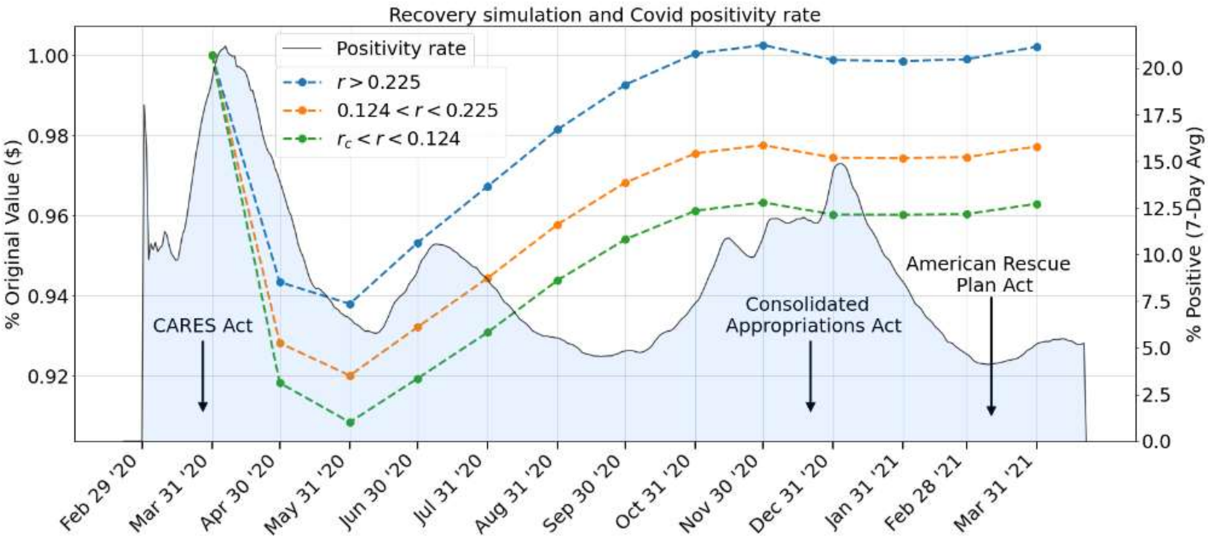


Figure 1: Comparing network response (dashed lines, left vertical axis) with Covid positivity rate in the U.S. (solid line, right vertical axis). Network response is indexed relative to March 2020 output levels. The dates of the signing of the three U.S. stimulus packages indicated: CARES Act - March 27, 2020, Consolidated Appropriations Act - December 21, 2020, and American Rescue Plan Act - March 11, 2021. Positivity data is obtained from the Center for Disease Control and Prevention (CDC).

References:

(1) Lucas, R. E. (1977): "Understanding Business Cycles," *Carnegie–Rochester Conference Series on Public Policy*, 5, 7–29. [1977,1996]

(2) Acemoglu, Daron, Vasco M. Carvalho, Asuman Ozdaglar, and Alireza Tahbaz-Salehi(2012). "The network origins of aggregate fluctuations." Econometrica 80, 1977–2016.

(3) King, Robert G. & Rebelo, Sergio T., 1999. "Resuscitating real business cycles," Handbook of Macroeconomics, in: J. B. Taylor & M. Woodford (ed.), Handbook of Macroeconomics, edition 1, volume 1, chapter 14, pages 927-1007 Elsevier.

Extracting complements and substitutes from sales data: a network perspective

Y. Tian[1], *S. Lautz*[2], *A. Wallis*[2], *R. Lambiotte*[1]. *(1) Mathematical Institute, University of Oxford, Oxford, UK, yu.tian@maths.ox.ac.uk. (2) Tesco PLC, Tesco House, Shire Park, Welwyn Garden City, UK*

We take a network perspective to help automatically identify two central concepts in economics and marketing research, complementarity and substitutability, from easily accessible sales transaction data. Specifically, we model the structure in the data as a bipartite network, and characterise complements and substitutes as product nodes exhibiting particular connectivity patterns, i.e. specific roles in the bipartite network. To extract such roles, we develop customised methods: firstly, we propose null models on bipartite networks to infer the complement and substitute incidence; secondly, we propose measures to quantify their importance, based on random walks on networks; thirdly, we extract roles in the bipartite network by uncovering communities in the projections onto the product side by both the incidence and the importance of the two types of relationships. This step can be done via standard community detection tools, e.g. the map equation. Promisingly, our methods retrieve the ground truth complements and substitutes in the simulated transaction data, with known product relationships and purchase rules; see the figure. Furthermore, we apply our methods to real-world sales data provided by Tesco, and validate the results with a large-scale flavour compound and recipe dataset, where we find features of complements to co-appear in more recipes and of substitutes to share more flavour compounds.

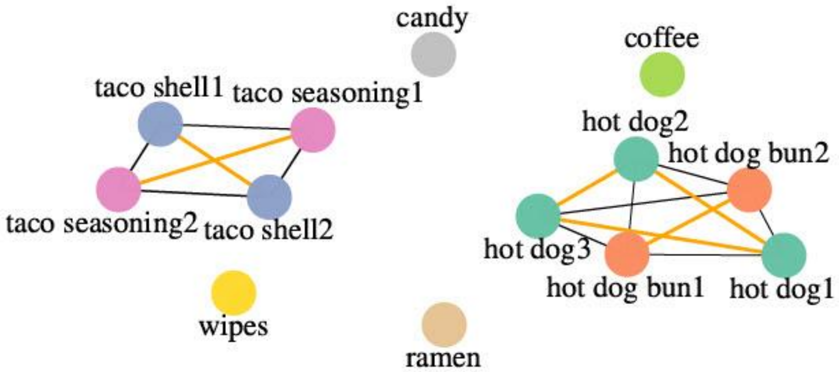


Figure 1: The product network from the simulated data, connected by both the complementarity (in black) and the substitutability (in orange), with products in the same extracted substitute group in the same colour. The true product relationships: (i) hot-dog-and-hot-dog-bun and taco-shell-and-taco-seasoning are complementary pairs; (ii) products of the same names (ignoring numbers) are substitutes.

Acknowledgements

Y. Tian is supported by the EPSRC Centre for Doctoral Training in Industrially Focused Mathematical Modelling (EP/L015803/1) in collaboration with Tesco PLC. We also thank Yong-Yeol Ahn for providing the flavour compound and recipe data, and for his assistance.

References

[1] Y. Tian, S. Lautz, A. Wallis, R. Lambiotte, *arXiv:2103.02042* (2021).

An Experimental Approach to Hierarchies: High Ranked Individuals Exploit Low Ranked Ones But Act Altruistically

***P. Lozano**[1,2], A. Antonioni[1,2], A. Sánchez[1,2]. (1) Universidad Carlos III de Madrid, Leganés, Spain, pablo.lozano@uc3m.es. (2) Grupo Interdisciplinar de Sistemas Complejos (GISC) Leganés, Spain.*

Hierarchies are present in many animal and human societies [1], and most of these societies also show cooperative behavior. This is intriguing because increasing evidence supports the idea that rigidly enforced hierarchies lead to a decrease of cooperation in non-human primates as well as in human groups [2]. We define cooperation as working together for a shared benefit, while we speak of hierarchy as a ranking according to a relative status that allows different access to resources. Thus, there has to be a mechanism that allows both a hierarchical organization and a high level of cooperation. We propose a mechanism that can support cooperation in hierarchical societies based on the definition of social norms, which is particularly suited to the human case. We specify this mechanism through an agent-based model [3] including collective work to produce a public good, fights as social interactions governing the hierarchy, and norm internalization. Hierarchy is dynamic as it arises from the results of fights and is modified by them. Norm internalization reflects the individual's perception of how much she is going to retain from the public good due to her position in the hierarchy. Evolution is also included in the model by considering both individual and group selection. We tested the model against real-world data [4] using an experimental approach where the participants were involved in several experiments resembling the model. The experimental data validates the results from the model, recovering the principal characteristics of human societies. The emergence of hierarchies is a robust feature of the model and is also observed in the experimental data. Individuals react differently to other people's behavior: defectors are more frequently chosen to fight against, therefore explaining different heuristics for conflicts depending not only on one's hierarchy position but also on the others', which could be thought of as a social norm. Individuals in the top part of the hierarchy fight more than those in the bottom, as predicted by the model. There is correlation between an individual's position in the hierarchy and the frequencies of cooperation and fighting: the higher (respectively, lower) the rank in the hierarchy, the more (respectively, less) the individuals fight and cooperate. We thus see that egalitarian societies can transition to hierarchical ones and still maintain high levels of cooperation, contrary to [2] where hierarchy was fixed and cooperation did not emerge, so our result could be possibly explained by the dynamic hierarchy acting as indirect punishment.

References

[1] Cronin, K. and Sánchez, A. Social dynamics and cooperation. Advances in Complex Systems (2012).

[2] Cronin, K. et al. Hierarchy is Detrimental for Human Cooperation. Scientific Reports (2016).

[3] Lozano, P., Gavrilets, S., and Sánchez, A. Cooperation, social norms, and the emergence of hierarchical societies. Scientific Reports (2020).

[4] Lozano, P. Antonioni, A., and Sánchez, A. Hierarchical and cooperation: An experimental approach. Under review.

Quantifying firm-level economic systemic risk from nation-wide supply networks

Christian Diem[1,2], András Borsos[1,3,4] Tobias Reisch[5,1], János Kertész[4,1], and Stefan Thurner[5,1,6]

[1] Complexity Science Hub Vienna, Josefstädter Strasse 39, A-1080 Vienna, Austria
[2] Institute for Finance, Banking and Insurance, Vienna University of Economics and Business, Welthandelsplatz 1, A-1080 Vienna, Austria
[3] Financial Systems Analysis, Central Bank of Hungary, Szabadsag ter 9, H-1054 Budapest
[4] Department of Network and Data Science, Central European University, Quellenstrasse 51, A-1080 Vienna, Austria
[5] Section for Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse 23, A-1090, Austria
[6] Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 85701, USA

Crises like COVID-19 or the Japanese earthquake in 2011 exposed the fragility of corporate supply networks. The production of goods and services is a highly interdependent process and can be severely impacted by the default of critical suppliers or customers. While knowing the impact of individual companies on national economies is a prerequisite for efficient risk management, the quantitative assessment of the involved economic systemic risks (ESR) is hitherto practically non-existent, mainly because of a lack of fine-grained data in combination with coherent methods. Based on a unique value added tax dataset we derive the detailed production network of an entire country and present a novel approach for computing the ESR of all individual firms. We demonstrate that a tiny fraction (0.035%) of companies has extraordinarily high systemic risk impacting about 23% of the national economic production should any of them default. Firm size alone cannot explain the ESR of individual companies; their position in the production networks does matter substantially. If companies are ranked according to their economic systemic risk index (ESRI), firms with a rank above a characteristic value have very similar ESRI values, while for the rest the rank distribution of ESRI decays slowly as a power-law; 99.8% of all companies have an impact on less than 1% of the economy. We show that the assessment of ESR is impossible with aggregate data as used in traditional Input-Output Economics. We discuss how simple policies of introducing supply chain redundancies can reduce ESR of some extremely risky companies.

Does Size Matter?

K. Malarz,[1] *M. J. Krawczyk*,[1] *(1) AGH University of Science and Technology, al. Mickiewicza 30, Kraków, Poland,* malarz@agh.edu.pl.

Publishing a list of the richest people in a given country has become an annual practice for many journals and magazines. The length of this list is country/magazine-dependent. For example, in Poland, lists of 100-richest people have been published every year by Wprost magazine since 1990 [1], i.e. starting only one year after the political and economic transformation. In this presentation we show how the list length influences some statistical features of the richest people ranking. In Fig. 1 we present the dependence of the overlap ratio of top n richest lists at times t_1 and t_2 on the separation time $k=t_2-t_1$. The overlap ratio Ω_n is measured as the percentage of people that are on both lists. The data points are obtained by averaging ($\overline{\cdots}$) over pairs of t_2, t_1 such that $t_2-t_1=k$. In the left panel of Fig. 1 the real-world data for Poland for various lengths n is shown. The data comes from the ranking lists published by the Wprost magazine. In the right panel of Fig. 1 we show results of Monte Carlo simulations of an agent-based model [2, 3] describing evolution of wealth of N interacting agents in a closed macro-economic system. Both the empirical overlap ratio and the overlap ratio obtained in the Monte Carlo simulations appear to be n-independent, except for very short list lengths ($n < 20$) and relatively long time separations ($k > 15$ years) where we can see a weak size dependence in real world data.

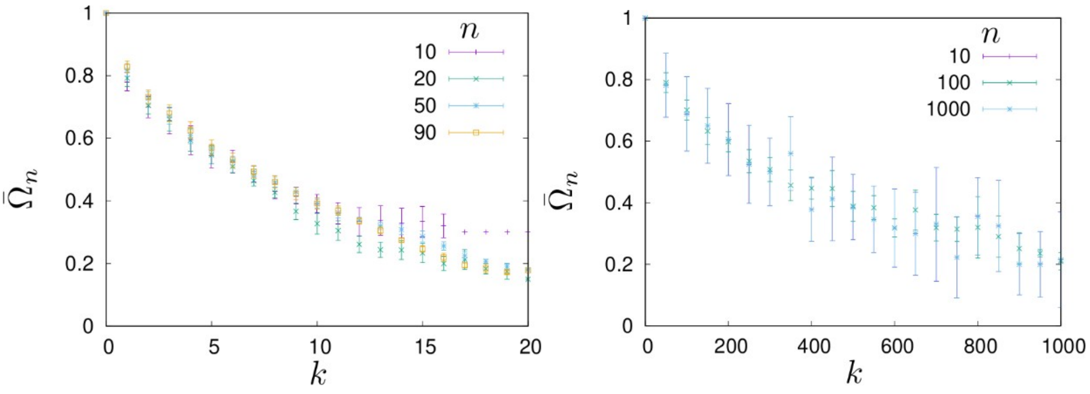


Figure 1: Time evolution of top-n richest people overlap in Poland [1] (left panel, n=10, 20, 50, 90 among N=360 people appearing at least one time on Wprost list between 2000 and 2020) and in agent-based model [2] simulation (right panel, n=10, 100, 1000 among $N=10^4$ agents).

References

[1] Wprost Magazine, *The top-100 richest people in Poland*.

[2] Z. Burda, M. J. Krawczyk, K. Malarz, M. Snarska, *arXiv:2105.08048 [econ.TH]*, (2021).

[3] J. P. Bouchaud, M. Mézard, Physica A, 282 (2000) 536.

Intra-Industry-Sector Heterogeneity and its Implication for the Spreading of Shocks in Production Networks

C. Diem^{1}*, A. Borsos*^{1,2,3}*, T. Reisch*^{,4,1}*, J. Kertész*^{3,1}*, S. Thurner*^{4,1,5} *(1) Complexity Science Hub Vienna, Josefstädter Strasse 39, A-1080 Vienna, Austria, diem@csh.ac.at (2) Central Bank of Hungary, Szabadság tér 9, 1054 Budapest, Hungary (3) Central European University, Quellenstrasse 51, A-1100 Vienna, Austria (4) Medical University of Vienna, Spitalgasse 23, A-1090, Austria (5) Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 85701, USA*

Models for the spreading of shocks in production networks are one of the key tools to assess the economic and consequently societal effects of crises like natural disasters and climate change [1], or the COVID-19 pandemic [2]. A crucial element of these models are the highly aggregated industry sector level production networks (based on Input-Output tables), on which the spreading dynamics of the direct initial shocks -- stemming from crisis scenarios -- to other parts of the economy are simulated. Based on the firm-level production network of Hungary, containing 91,595 firms, we show that firms within a given industry exhibit vastly heterogenous interlinkages in the production network. By using a shock propagation model on the firm level [3], we show that crisis scenarios with the same initial shock size for each sector, but affecting different companies within it, lead to substantially different shock cascades. These would be indistinguishable when considering aggregated sector level production networks. The red box plots in Fig. 1 show for each of 86 sectors how strong it was affected by the shock cascades triggered by 1,000 different agricultural crisis scenarios and blue '+'-symbols show how the sectors are affected by the corresponding sector level cascade. The large interquartile ranges indicate that aggregated models can produce sizable under- or overestimation of shock cascades triggered by crisis scenarios that can lead to fundamentally wrong assessments of systemic risk.

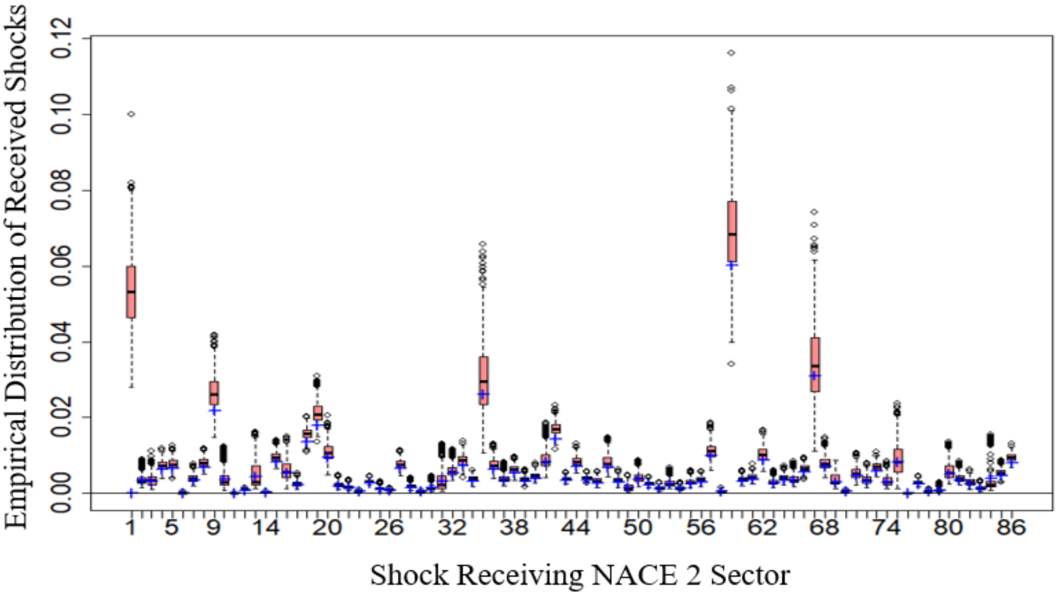


Figure 1: Distribution of shock cascades across NACE 2 sectors in response to 1,000 initial crisis scenarios each destroying 20 % of the sector crop and animal production.

References

[1] S. Hallegatte, C. Green, R.J. Nicholls, J. Corfee-Morlot, Nat. Clim. Change 3, (2013) 802

[2] A. Pichler, M. Pangallo, M. d. Rio-Chanona, F. Lafond, D. Farmer, SSRN3788494 (2021)

[3] C. Diem, A. Borsos, T. Reisch, J. Kertész, S. Thurner, arXiv:2104.07260 (2021)

The Marriage Game: New Insights from a Dynamic Model of Marriage, Domestic Labour, and Divorce

Stefano Benati[1]*, Eugenio Dacrema*[2]*, Paolo Eugenio Petit Valdes*[3]*. (1) Department of Sociology and Social Research, University of Trento, Via Verdi 26,Trento, Italy, stefano.benati@unitn.it. (2) World Food Programme, Headquarters, Rome, Italy (3) Department of Mathematics, University of Trento, Trento, Italy*

The Marriage Game [1] is a Bayesian Game in which women and men must decide whether to marry, to cooperate in the domestic labor, and ultimately to divorce in the case that the husband does not cooperate. The game is played by three archetypes of women: the Traditional, the Transitional and the Autonomous, and by three types of men: the Hardliner, the Adjuster and the Cooperator. Types of men and women differ in their preferences and decisions along the game tree. At the beginning, women make the first decision about whether to marry or not, depending on their perception regarding the possibility that the husband will cooperate or not cooperate in the domestic labor. Next, the wife can use divorce as a threat against a lazy husband, but not all women are really willing to give effect to the threat. Some will, who are endowed with sufficient resources to afford a divorce, However, others will ultimately allow the husband to keep dismissing his share of domestic labor. The static analysis of the game reveals the existence of different equilibria, characterized by the decisions of each player regarding to marry, to cooperate and to divorce, depending on the static assumption of the frequencies of archetypes within the population. In one equilibrium, all types of men and women will marry, while in other equilibria only a fraction of players decides so. A crucial (and unrealistic) element of the static analysis is that frequencies are players' common knowledge.

We extend the analysis of the Marriage Game using an Agent Based Model accounting for the imperfect information that players have about the frequencies of the different archetypes within the population. Their only source of information is the divorce rate. The higher the rate, the higher the possibility of marrying a non-cooperator man (for a woman), or an autonomous woman willing to divorce in case of non-cooperation (for a man). Using this only source of information, the dynamic game reveals periodical fluctuations of marriages and divorces. The higher the divorce rate, the higher the probability that a man cooperates, enhancing the probability for women to marry and at the same time decreasing the rate of divorces. However, the diminishing rate of divorces leads gradually men to cooperate less, enhancing once again the number of divorces.

References

[1] Richard Breen and Lynn Prince, The Persistence of the Gendered Division of Domestic Labor, *European Sociological Review*, 21 (2005) 43-57.

[

From Code to Market: Network of Developers and Correlated Returns of Cryptocurrencies

__L. Lucchini__[1], L. Alessandretti[2], B. Lepri[1], A. Gallo[3], A. Baronchelli[3]. (1) FBK, Sommarive 18, Trento, Italy, llucchini@fbk.eu. (2) University of Copenhagen, Copenhagen, Denmark (3) City University of London, London, UK.

Cryptocurrencies are a paradigmatic environment in which “code is law” is a founding principle [1]. In the case of crypto-assets, their security, transferability, and availability are solely determined by the code they are built on. In this work, we identify and show how one simple event in the code-development space anticipates a corresponding behaviour in the financial market. If code is open source, as it is customary for most of the cryptocurrencies, it would in principle prevent asset manipulations and grant transparency to users and traders. However, this principle considers cryptocurrencies as isolated entities, neglecting possible connections between projects. By leveraging on code developing events (on *GitHub*), we discover that 4% of the cryptocurrency code-developers contributes to more than one cryptocurrency project. This reveals that the cryptocurrency ecosystem has an underlying networked structure, where projects are connected one another when they share a common developer. By defining as ``connection time" the time at which each link was activated (see Fig.1(A)), we discover that, at the turn of this event, two linked crypto-assets increase their return correlation with respect to a randomized sample of pairs (see Fig.1(B-C)). This result is reported in Fig.1(D), where we show the average standardized correlation for linked crypto-assets (in red) and for randomized pairs (in blue). We repeated the analyses to study the correlations also between daily changes in trading volume and market capitalization time-series. We found no significant difference at the turn of the connection time for those measures. Robustness tests, to ensure the stability of the analysis include project age-based randomization and linked pairs sub-sampling. Our results establishes a first and direct link between two realms often-regarded as separated. In particular, it reveals that the temporal dynamics of cryptocurrency co-coding provides insights on market behaviours which could not be deduced solely based on the combined knowledge of a cryptocurrency code and its market state. This informs us that transparency (being the availability of relevant market information to market participants) is a property of the system if it is considered as a whole.

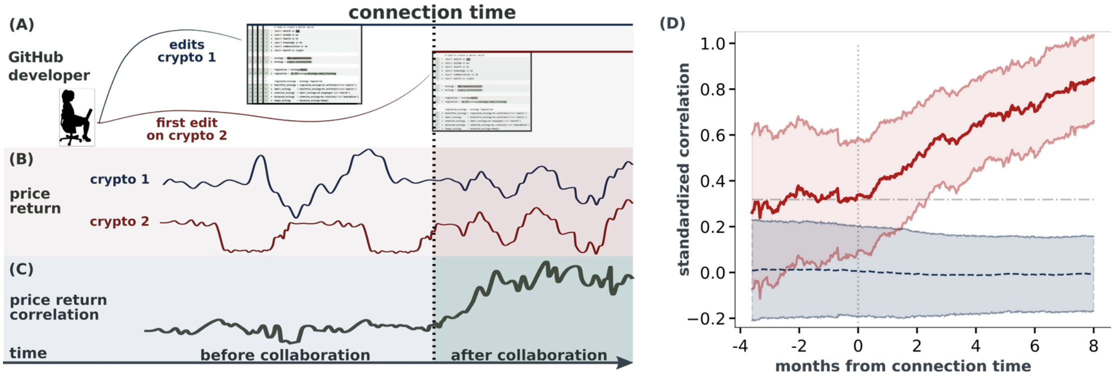


Figure 1: Collaboration identification and consequent market synchronization

References

[1] Lawrence Lessig. Code is law: On liberty in cyberspace.Harvard Magazine, 01 2000.

Higher-order hierarchical spectral clustering for multidimensional data

G. Brandi[1], *T. Di Matteo*[1,2,3]. *(1) Department of Mathematics, King's College London, The Strand, London, WC2R 2LS, UK, giuseppe.brandi@kcl.ac.uk, (2) Complexity Science Hub Vienna, Josefstaedter Strasse 39, A 1080 Vienna, Austria, (3) entro Ricerche Enrico Fermi, Via Panisperna 89 A, 00184 Rome, Italy.*

Understanding the community structure of countries in the international food network is of great importance for policymakers. Indeed, clusters might be the key for the understanding of the geopolitical and economic interconnectedness between countries. Their detection and analysis might lead to a bona fide evaluation of the impact of spillover effects between countries in situations of distress. In this paper, we introduce a clustering methodology that we name Higher-order Hierarchical Spectral Clustering (HHSC), which combines a higher-order tensor factorization and a hierarchical clustering algorithm. We apply this methodology to a multidimensional system of countries and products involved in the import-export trade network (FAO dataset). We find a structural proxy of countries interconnectedness that is not only valid for a specific product but for the whole trade system. We retrieve clusters that are linked to economic activity and geographical proximity. We also show how to use this type of modeling approach to make predictions.

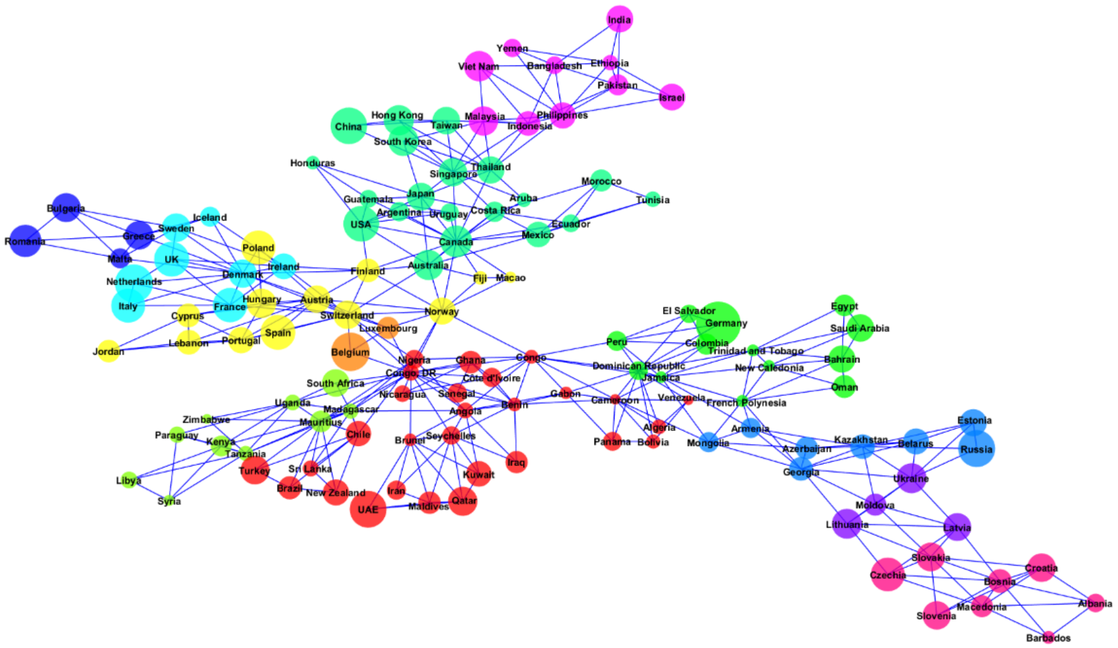


Figure 1: HHSC clustering for imports. The size of the node is proportional to the total rate of growth of products the country imported between 1986 and 2018.

References

[1] G. Brandi and T. Di Matteo. Predicting multidimensional data via tensor learning. Journal of Computational Science, 2021.

[2] G. Brandi and T. Di Matteo. Higher-order hierarchical spectral clustering for multidimensional data. In ICCS 2021 proceedings. Springer International Publishing.

Market-Share Structures of Competing Firms: A Simple Physics Model

Joseph Hickey[1]*. (1) Bank of Canada, 234 Wellington St., Ottawa, Canada, jhickey@bank-banque-canada.ca*

In many markets, firms compete continuously with one another for clients. This competition creates market-share structures that range from domination by one or a few giant companies (high market concentration, e.g. monopoly or oligopoly) to a market that is split among many small competing firms with similar numbers of clients (low concentration). Although market-shares have been studied extensively [1,2], the core mechanisms influencing their structures and evolution remain unknown. I propose a simple model in which there are N individual clients, each of which is a customer of one of N firms (every individual thus has the potential to act as a firm and attract other clients). At every step in the simulation, a single client is chosen at random and "reconsiders" his choice of firm, choosing a destination firm $i \in N$ with probability $p_i = f(n_i, \alpha, \beta)$, where n_i is the number of clients currently with firm i, α is a parameter that controls how the advantage in attracting clients scales with firm size and β is a parameter controlling the probability of attracting clients regardless of firm size (reflecting the ease of entering and participating in the market). Fig. 1 shows a phase diagram with regions of high and low market concentration and a region of transition between them. Markets with higher ease of entry (high β) have a more gradual transition to monopoly than low β markets, and can remain at relatively low concentrations over a larger span of α. This suggests that policies to promote competition (e.g. antitrust regulations) may be more critical in markets with low β, which can experience rapid shifts to monopoly for small changes in α. We validate the model using data from real markets including mobile phone service providers in Canada and discuss how the model could be used to predict the stability of an existing market vs. potential for disruption by a new entrant.

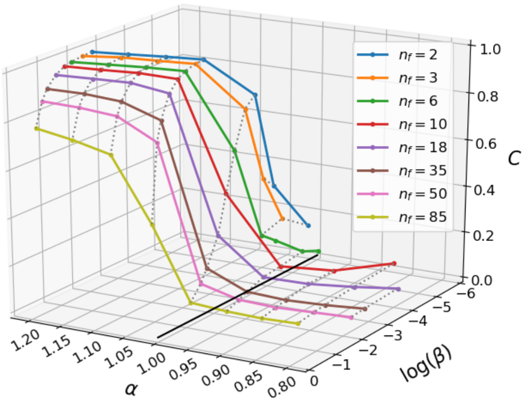


Figure 1: Phase diagram with concentration, C, as function of α and β. $C = (\langle m_1 \rangle - 1/n_f)/(1 - 1/n_f)$, where $\langle m_1 \rangle$ is the average (steady-state) market-share of the largest firm and n_f is the average number of firms with at least one client. Dashed grey lines: guide to the eye connecting points with same value of α. System size N = 500.

References

[1] J. Sutton, *Handbook of Industrial Organization*, Vol. 3, Ch. 35 (2007) 2301-2368.

[2] S. Klepper, J.L. Simons, *Int. J. Industrial Organization*, 23 (2005) 23.

Wavelets in clustering time series: an application in portfolio selection using price and volume

__E. Alvarez__[1], J. G. Brida[1], L. Moreno[1,], F. Negreira[1,], (1) Universidad de la República, Gonzalo Ramírez 1926,Montevideo, Uruguay, ealvarez@ccee.edu.uy.

There are multiple problems in Economics and Social Sciences where it is important to compare the dynamics of different time series. However, in the academic literature, there are few techniques for the clustering of multidimensional time series. This work presents a clustering methodology for multidimensional time series based on wavelets. The multivariate time series of the companies are considered as observations of a functional random variable, that is, they take values in an infinite-dimensional space. This methodology is used to analyze an important case in finance, such as the construction of portfolios when the information from the prices and volumes of the assets is used. Using daily price and volume data from Nasdaq-100 companies during the period 2018-03 to 2020-04, this article shows the existence of three clusters of companies. The results of this work show that there are differences in the way in which monetary policy and trade war affect companies, depending on the type of industry and its linkages in the global economy.

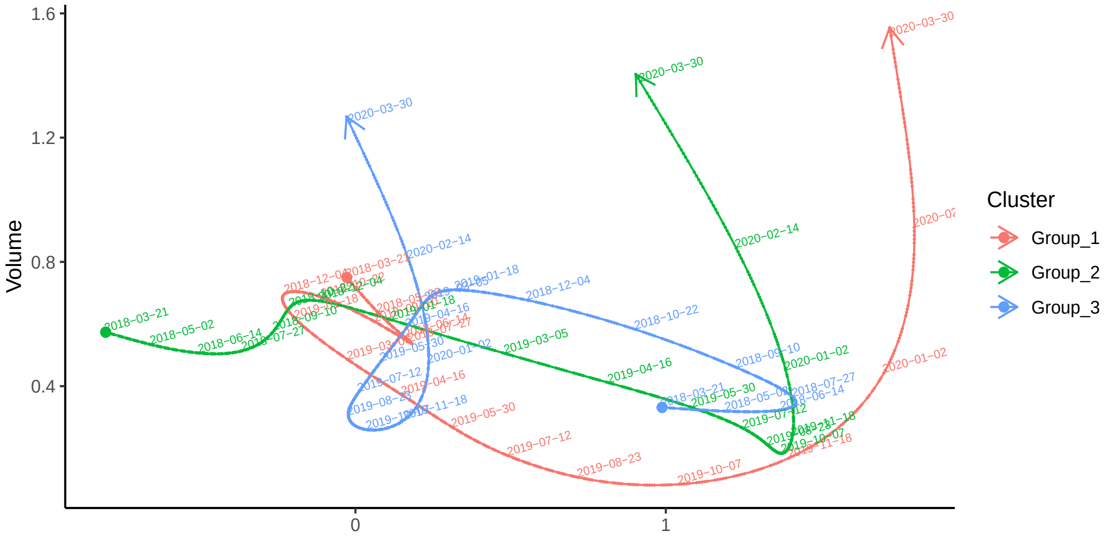


Figure 1: Dynamics of clusters. Source: own construction based on data from https://www.nasdaq.com/

Inferring supply networks from communication data

T. Reisch[1,2]*, G. Heiler*[2,3]*, C. Diem*[2]*, S. Thurner*[1,2,4]*. (1) Medical University Vienna, Spitalgasse 23, 1090 Vienna, Austria, reisch@csh.ac.at. (2) Complexity Science Hub Vienna, Josefstädter Straße 39, 1080 Vienna, Austria, (3) Technical University Vienna, 1040 Vienna, Austria (4) Santa Fe Institute, Santa Fe, 1399 Hyde Park Rd., NM 87501, USA*

The underlying structure of an economy is the network of customer-supplier relationships. Economic shocks are propagated as demand-side shocks from buyers to suppliers or as supply-side shocks from supplier to buyer. Mostly due to a lack of appropriately granular data, current literature is typically focussing on sector level economic shocks and shock propagation [1]. Only recently has data on firm-level trade relationships become accessible - restricted, however, to a small number of countries, such as Japan, Belgium, Brasil or Hungary. Such data is hard to obtain and needs extensive surveys, tax or payment system data that is hard to get. Here we present telecommunication data as an inexpensive and alternative way to get an estimate of a national customer-supplier network. within a cooperation with an Austrian mobile phone operator, we are able to analyse mobile phone communication between anonymized companies. We compare the communication links with a survey on critical suppliers conducted in April 2020. From this ground truth we estimate the probability of correctly identifying a customer-supplier relation, given a communication link. The probability increases with the length of the communication between companies and reaches values above 90%. We estimate the direction of the flow of goods from input-output data and finally compute the economic systemic risk of individual firms as the fraction of the economy that is potentially affected by cascading effects subsequent to the default of a given firm [2]. We discuss the limitations, in particular the effects resulting from our assumptions.

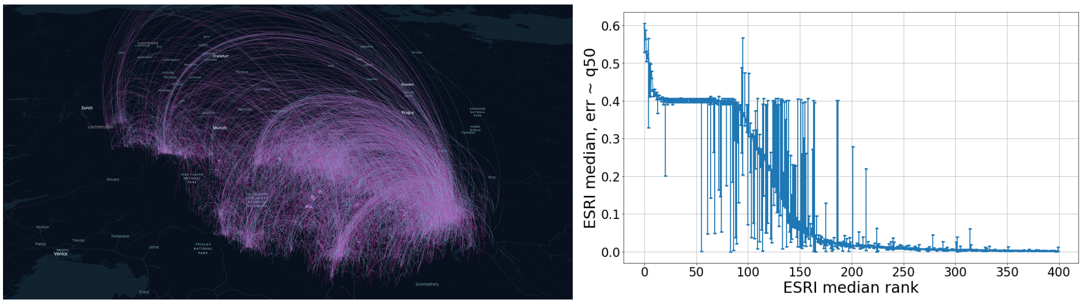


Figure 1: *Left panel:* Anonymized firm-firm communication network on a map of Austria. Arcs represent communication events between two firms. The color and thickness of a link indicates its direction and strength, respectively. For the purpose of anonymizing origin and destination of links, positions of the arcs are randomly shifted within a region. *Right panel:* Preliminary Economic Systemic Risk Profile, calculated for simulated buyer-supplier relationships. Some firms can affect significant fractions of the economy.

References

[1] Klimek, P., Poledna, S., & Thurner, S. (2019). Quantifying economic resilience from input–output susceptibility to improve predictions of economic growth and recovery. Nature communications, 10(1), 1-9.

[2] Diem, C. et al. (2021). Quantifying firm-level economic systemic risk from nation-wide supply networks. Arxiv preprint arXiv:2104.07260

Socio-Spatial Complexity: Cities As Neural Networks
A Case Study of Housing Prices Estimations

Sarah Soleiman[1,2], *J. Randon-Furling*[1,3,4]. *(1) SAMM, Univ. Paris 1 Panthéon-Sorbonne – FP2M, CNRS, Paris, France, ssoleiman@meilleursagents.com. (2) MeilleursAgents, Paris, France. (3) Department of Mathematics, Columbia University, New York, USA. (4) MSDA, UM6P, Ben Guerir, Morocco.*

The intertwining of socio-spatial complexity with that of price formation leads to challenging questions in spatial statistics when modeling urban real estate markets and their dynamics. The exact same apartment typically will not have the same value depending on its location in the city – due to the specifics of the local neighborhood and other qualitative features. Traditional methods rely on so-called hedonic prices and implicit markets [1], modified in order to incorporate spatial effects via geographically weighted regressions (GWR, [2]). However, the availability of large data sets pertaining to the socio-economic characteristics of cities, at a very fine-grained level, together with the advent of computational and algorithmic power allows one to capture in much finer detail the complex relationship between space and price in the real estate market. We present here a two-fold approach.

First, using a Self-Organizing Map algorithm [3] on sets of demographical, economical and infrastructural data, we bring out socio-spatial structures for a heterogeneous sample of French cities. We then combine geographical distance and distance on the Self-Organized map into a geo-statistical estimation model of real estate prices – and we show that this approach out-performs the standard GWR models (Fig. 1). Second, we introduce an altogether new paradigm, where the city itself is thought of as a network of locations, that functions as a neural network. Each "neuron" has a value that corresponds to the price per square meter at the location. Data from past transactions allow us to train this network, so that it learns the housing market in the city. We apply this paradigm to data from the Paris housing market between 2010 and 2020.

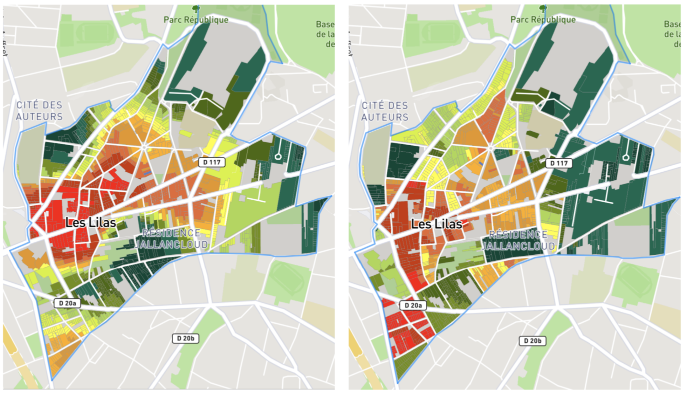


Figure 1: Price estimates in Les Lilas (Paris): standard GWR model (left) *vs* ours (right).

References

[1] S. Rosen, *Journal of Political Economy*, 82 (1) (1974) 34.

[2] A.S. Fotheringham, C. Brunsdon, M. Chalton, *GWR*, Wiley & Sons, 2002.

[3] M. Olteanu, A. Hazan, M. Cottrell, J. Randon-Furling, *Neur. Comp. & Appl.*, 31 (2019) 1.

Systemic Risk in Populations' Basic Supply Networks

W. Schueller^{1,2}*, C. Diem*^{2,3}*, J. Stangl*^{2}*, B. Conrady*^{4} *,M. Gerschberger*^{5}*, S. Thurner*^{1,2,6}*. (1) Medical University of Vienna, Spitalgasse 23, A-1090, Austria (2) Complexity Science Hub Vienna, Josefstädter Strasse 39, A-1080 Vienna, Austria, schueller@csh.ac.at, diem@csh.ac.at (3) Vienna University of Economics and Business, Welthandelsplatz 1, A-1080 Vienna, Austria (4) University of Copenhagen, Grønnegaardsvej 8, 1870 Denmark, (5) Josef Ressel Centre for Real-Time Value Network Visibility, Logistikum, University of Applied Sciences Upper Austria, (6) Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 85701, USA*

The COVID-19 crisis has drastically emphasized the fragility of national and international supply networks, leading to significant shortages of peoples' supply with basic goods and services --- like food, medical equipment etc. One reason why these shortages occur is when critical nodes (firms, business facilities) in the supply network stop functioning and the failures propagate along the corporate supplier-buyer links until they reach the population (final consumer). Models for shock propagation and systemic risk assessment in supply- and production networks usually focus on the effects on firms [2] or the overall economy [3], but neglect the systemic risks threatening the basic supply of the population. We address this question by developing a new network-based systemic risk measure that enables us to identify the critical nodes in the supply network and estimate the effects on basic supply security. To empirically validate our methodology we construct a highly granular section of the Austrian pork meat supply network, containing all approx. 20,000 relevant firms and business facilities (nodes) --- ranging from farms, slaughter houses, distribution centers, supermarkets, down to the final consumer. Figure 1 illustrates the distribution of the calculated systemic risk measure of all nodes. Further, we simulate COVID-19 related stress-scenarios like border closures and local lockdowns to assess the effects of systemically relevant events for the basic supply of the population. The developed risk measure is immediately policy relevant, as a fact-driven crisis management tool for administration and governments.

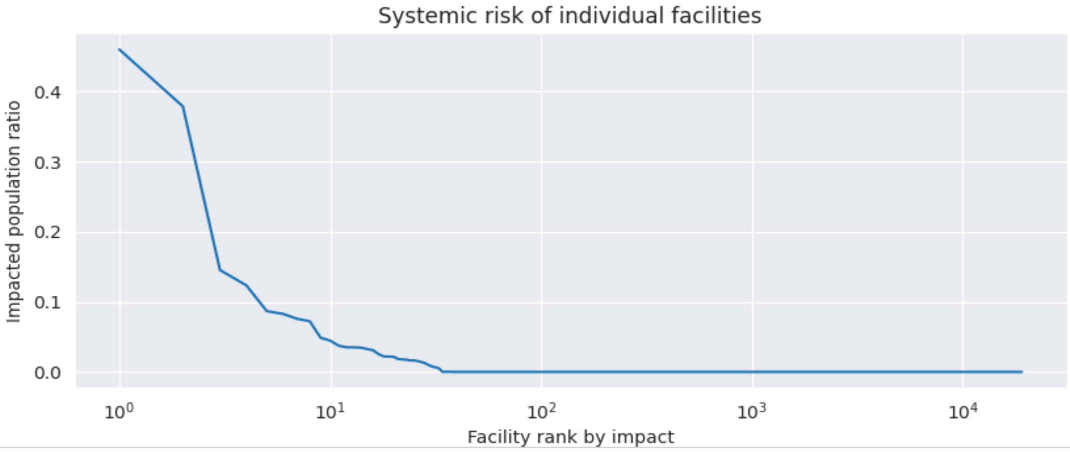


Figure 1: Distribution of individual nodes (business facilities) systemic risk to the basic supply of the population in the Austrian pork meat supply network.

References

[1] Bloomberg, Meat-Shortage Risk Climbs with 25% of U.S. Pork Capacity Offline, (2020) 04-22.

[2] B. Shao, Z. Shi, T. Choi, S. Chae, *Decision Support Systems*, 114 (2018) 37.

[3] C. Diem, A. Borsos, T. Reisch, J. Kertész, S. Thurner, arXiv:2104.07260 (2021).

Deep Hawkes Process for High Frequency Market Making

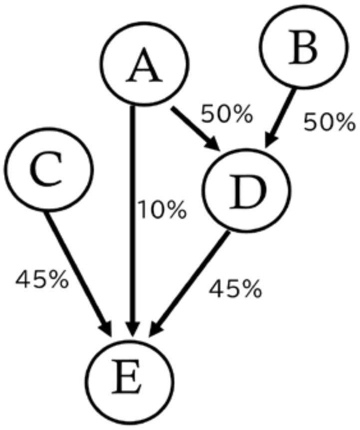
***Pankaj Kumar**, Copenhagen Business School, Solbjerg Plads 3, 2000 Frederiksberg, Denmark pk.mpp@cbs.dk.*

A high frequency market making is liquidity providing trading strategy that simultaneously generates many bids and asks for a security at ultra low latency while keeping a relatively neutral position. The strategy makes a profit in the form of bid-ask spread for every buy and sell transaction against the risk of adverse selection, uncertain execution, and inventory risk. We design realistic simulations of limit order markets and develop a high frequency market making strategy where agents process order book information to post the optimal price, order type, and execution time. By introducing the Deep Hawkes process in the market making strategy, we allow a feedback loop between order arrival and state of limit order book together with self- and cross-excitation effect. Our market making strategy accounts for the cancellation of orders that influence order queue position, profitability, bid-ask spread, and value of the order. The experimental results show that our trading agent outperforms the baseline strategy which uses a probability density estimate of the fundamental price. We investigate the effect of cancellations on the market quality and agent's profitability. We validate the closeness of the simulation framework to realism by reproducing stylized facts from the empirical analysis of the simulated order book data.

Flow of corporate control power in the global ownership network

Takayuki Mizuno[1], Shouhei Doi[2], Takahiro Tsuchiya[3], Shuhei Kurizaki[4]
[1] National Institute of Informatics, Tokyo, Japan
[2] Hokkaido University, Sapporo, Japan
[3] Kyoto University of Advanced Science, Kyoto, Japan
[4] Waseda University, Tokyo, Japan

In recent years, many shareholders have vowed to optimize their investment strategy with the environmental, social, and governance (ESG) benchmarks. However, because the structure of the global ownership network is more complex than before, unfortunately, most shareholders have linkages to non-ESG firms [1]. To investigate the social responsibility of ultimate/intermediate shareholders to control non-ESG firms through the ownership linkages, we introduce a new measure, *Network Power Index* (NPI), that explicitly takes into account both the possibility that centralized (proxy) voting behavior among dispersed shareholders converts fragmented ownership into corporate control and the sequence of consolidated control that may be channeled through the ownership network [2,3].



In the Figure, the circles indicate shareholders and companies, and the arrows with the percentages present the share of equity stakes each shareholder owns. We follow the Shapley and Shubik power index to assume that they may randomly form a coalition out of a permutation of possible combinations. If A prevails in D, it would indirectly control D's 45% share of E's equity on top of its direct ownership of the 10% share. Consequently, A obtains the power to control E's decision-making fully. If B prevails in D, on the other hand, A can only control E with probability $\frac{1}{3}$ because C, A, and B, needs one additional shareholder to build a majority coalition. In short, A's NPI to E is $\frac{1}{2}\times 1+\frac{1}{2}\times\frac{1}{3}=\frac{2}{3}$. The intermediate shareholder D mediates the control of E by the ultimate owners, A and B. This mediating force of D to E, so-called D's *Network Power Flow* (NPF) to E, is $\frac{1}{6}+\frac{1}{4}=\frac{5}{12}$.

Applying our algorithm to the data on global ownership from Bureau van Dijk in 2020, we quantify the connection between the ultimate owners and non-ESG firms. The results show that investment management companies play important roles in channeling the investment money into non-ESG firms.

References

[1] Mizuno, T., Doi, S., Tsuchiya, T., Kurizaki, S. (2021) Socially Responsible Investing through the Equity Funds in the Global Ownership Network. Submitted to PLOS ONE.

[2] Mizuno T, Doi S, Kurizaki S. (2020) The Power of Corporate Control in the Global Ownership Network. PLOS ONE15(8), e0237862.

[3] Mizuno, T., Doi, S., Kurizaki, S. (2019). Network Shapley-Shubik Power Index: Measuring Indirect Influence in Shareholding Networks. Studies in Computational Intelligence 882, 611-619.

Under Which Conditions Would a Post-Growth Transition Trigger a Financial Market Crash? Analysis Via an Heterogeneous Agent Model.

***A. Janischewski**, Faculty of Economics and Business Administration, Chemnitz University of Technology, Thüringer Weg 7, 09126 Chemnitz, Germany, anja.janischewski@wiwi.tu-chemnitz.de.*

Economic growth has been a main goal in economic systems for decades. Financial markets, too, have experienced much growth in this time. Thus, some system functions in financial markets, or financial systems in general may have become adapted to this growth trend.

However, preventing a climate crisis and its ecological tipping points requires large scale political measures, which would certainly lead to the shrinking of CO2 intensive sectors. Therefore, future returns on investment in CO2 intensive sectors would fall and thus reducing the fundamental value of the respective financial products. Furthermore, such large scale environmental measures might also have the side effect of reducing overall economic growth rates and overall returns on investment in the case the returns on environmentally sustainable investments do not increase substantially.

How would a complex system such as the financial market react to this scenario? Does a transition to significantly lower financial returns necessarily lead to a financial crash, as claimed by [1]? Financial asset prices are prone to sharp price drops when triggered by some event as for example in spring 2020. On the other hand, could a transition to lower economic growth rates even lead to more financial stability as found in [2]? One might state that the foreseeability of political measures with respect to climate change would smooth any asset price drops in shrinking sectors and thus, no crash would occur. On the other hand, positive feedback mechanisms such as trend following investors or fire sales can cause price movements that do not strictly follow fundamental values. Such feedback mechanisms could lead to accelerating dynamics causing a crash instead of a gradual adjustment of asset prices.

We propose to analyse the question whether a transition to lower rates of return due to environmental measures leads to a financial crash, with an heterogeneous agent model of a stylized financial market. We consider two main scenarios, both starting with the information of shrinking returns on investment. In the first scenario, agents believe that the shock is temporary and that the economy will catch up after a while (as during the Covid-19 pandemic). In the second one, agents believe that the transition to lower growth rates is of a more permanent nature. Furthermore, in the two scenarios a setup with heterogeneous behaviour is compared with a setup with only fundamental investors.

Are the emerging dynamics qualitatively different between a temporary shock and the prospect of long term low growth rates? Which mechanisms are responsible for this difference? Which role would a "run towards more liquid assets" play? Does the occurrence of the financial crash critically depend on how much returns on investment decline, i.e. is there a tipping point?

References

[1] D. Tokic, *The economic and financial dimensions of degrowth, Ecological Economics, 84, pp. 49-56* (2012).

[2] K. Safarzyńska, J.C.J.M. van den Bergh, *Financial stability at risk due to investing rapidly in renewable energy, Energy Policy, 108, pp.12-20 (2017)*

The Flocking Behaviour of Firms: an Agent-based Simulation of Organizational Sensemaking

F.P Zasa *[1], S. Magnanini[1], P.Bellis[1], T.Buganza[1], D.Trabucchi[1], R.Verganti[2] (1) School fo Management, Politecnico di Milano, Via Lambruschini 4/C, Milano, Italy, federicopaolo.zasa@polimi.it. (2) House of Innovation, Stockholm School of Economics, Saltmätargatan 13, 113 59 Stockholm, Sweden*

In today's turbulent environment, organizations need to make sense of continuous change [17]. Organizations represent complex systems, and their behavior is the result of interactions among individuals [1; 4]. Hence, also the organizational sensemaking process represents an emergent reaction based on the sensemaking of all individuals [3]. Indeed, when change disrupts the current functioning of a firm, every individual needs to make sense of the change [5; 8]. Building on previous studies on complexity theory of organizations, we adopt an agent-based perspective on sensemaking.

Organizational sensemaking represents the process through which a vision to react to the change is formed [5; 17]. Past studies have focused predominantly on understanding how to foster interactions to achieve collaborative sensemaking [e.g., 13; 7]. Today, the environment is increasingly turbulent, with large changes and disruption which continuously affect every individual in the firm [11]. Hence, we study how individual reactions to change give rise to a collective behavior: we build on the analogy of organizations as a flock to simulate the organizational sensemaking process [20]. According to this analogy, the behavior of groups of agents is governed by a limited set of heuristics – namely their ability to sense their peers, and their reactions to their behaviors. Our study thus aims to answer *how* different individual interaction norms in an organization generate different collective behaviors.

We extend Wilensky's [18] flocking model using NetLogo [19], building on key constructs in organizational theory to deduce a model of the organization as a flock [20]. We leverage the constructs of Cognitive Flexibility as peer exposure [16], Uniqueness [12], Synthesis capabilities as consensus [14] and Flexibility for change as communication [2].

Data analysis will be based on data stemming from five multinational organizations part of the research platform IDeaLs – Innovation and Design as Leadership. For each organization, approximately 200 employees are sampled to create an overview on the flocking behavior of the organization.

Various studies, particularly in the field of leadership, have shown the possibility and value of analysing the organization as a flock [20; 6]. The contribution of this research will be evaluating how an agent-based simulation based on flock rules can predict organizational outcome variables as engagement [9] and the development of a Shared Vision [15]. Gioia and Chittipeddi [5] argue that the development of a vision in the light of change is the outcome of the organizational sensemaking process. Our study contributes to expanding extant managerial theories through the flocking analogy, expanding the value which agent-based simulations can bring [10].

References

[1] Anderson, P. Complexity theory and organization science. *Organization Science, 10*(3) (1999) 216-232. doi:10.1287/orsc.10.3.216

[2] Di Fabio, A., & Gori, A. Developing a new instrument for assessing acceptance of change. *Frontiers in Psychology***, *7:802*** (2016) doi:10.3389/fpsyg.2016.00802

[3] Fulmer, C. A., & Ostroff, C. Convergence and emergence in organizations: An integrative framework and review. *Journal of Organizational Behavior,* ***37*****(S1)** (2016) S122-S145. doi:https://doi.org/10.1002/job.1987

[4] Garcia, R. Uses of agent-based modeling in innovation/new product development research. *Journal of Product Innovation Management,* ***22*****(5)** (2005) 380-398. doi:10.1111/j.1540-5885.2005.00136.x

[5] Gioia, D. A., & Chittipeddi, K. Sensemaking and sensegiving in strategic change initiation. *Strategic Management Journal,* ***12*****(6)** (1991) 433-448. doi:10.1002/smj.4250120604

[6] Jahmani, K., Fadiya, S. O., Abubakar, A. M., & Elrehail, H. Knowledge content quality, perceived usefulness, KMS use for sharing and retrieval: A flock leadership application. *VINE Journal of Information and Knowledge Management Systems,* ***48*****(4)** (2018) 470-490. doi:10.1108/VJIKMS-08-2017-0054

[7] Maitlis, S., & Christianson, M. Sensemaking in organizations: Taking stock and moving forward. *Academy of Management Annals,* ***8*****(1)** (2014) 57-125. doi:10.1080/19416520.2014.873177

[8] Maitlis, S., & Sonenshein, S. Sensemaking in crisis and change: Inspiration and insights from weick (1988) *Journal of Management Studies,* ***47*****(3)** (2010) 551-580.

[9] Martinez, M. G. Solver engagement in knowledge sharing in crowdsourcing communities: Exploring the link to creativity. *Research Policy,* ***44*****(8)** (2015) 1419-1430.

[10] Miller, K. D. Agent-based modeling and organization studies: A critical realist perspective. *Organization Studies,* ***36*****(2)** (2015) 175-196. doi:10.1177/0170840614556921

[11] Raveendran, M., Silvestri, L., & Gulati, R. The role of interdependence in the microfoundations of organization design: Task, goal, and knowledge interdependence. *Academy of Management Annals,* ***14*****(2)** (2020) 828-868. doi:10.5465/annals.2018.0015

[12] Schumpe, B., Herzberg, P., & Erb, H. Assessing the need for uniqueness: Validation of the german NfU-G scale. *Personality and Individual Differences,* ***90*** (2016) 231-237. doi:10.1016/j.paid.2015.11.012

[13] Stigliani, I., & Ravasi, D. Organizing thoughts and connecting brains: Material practices and the transition from individual to group-level prospective sensemaking. *Academy of Management Journal,* ***55*****(5)** (2012) 1232-1259.

[14] Taggar, S. Individual creativity and group ability to utilize individual creative resources: A multilevel model. *The Academy of Management Journal,* ***45*****(2)** (2002) 315-330. doi:10.2307/3069349

[15] Tsai, W., & Ghoshal, S. Social capital and value creation: The role of intrafirm networks. *Academy of Management Journal,* ***41*****(4)** (1998) 464-476. doi:10.2307/257085

[16] Webster, D. M., & Kruglanski, A. W. Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology,* ***67*****(6)** (1994) 1049–1062. Retrieved from https://doi.org/10.1037/0022-3514.67.6.1049

[17] Weick, K. E., Sutcliffe, K. M., & Obstfeld, D. Organizing and the process of sensemaking. *Organization Science,* ***16*****(4)** (2005) 409-421. doi:10.1287/orsc.1050.0133

[18] Wilensky, U. NetLogo Flocking model. http://ccl.northwestern.edu/netlogo/models/Flocking. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1998).

[19] Wilensky, U. NetLogo. http://ccl.northwestern.edu/netlogo/. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL (1999).

[20] Will, T. E. Flock leadership: Understanding and influencing emergent collective behavior. *Leadership Quarterly,* ***27*****(2)** (2016) 261-279. doi:10.1016/j.leaqua.2016.01.002

Economic Interplay Forecasting Business Success

N. Amoroso^{1,2}*, L. Bellantuono*^{3}*, A. Monaco*^{2,*}*, F. De Nicolò*^{3} *, E. Somma*^{4,5} *, R. Bellotti*^{2,3}*. (1) Università di Bari, Dipartimento di Farmacia-Scienze del Farmaco, Bari, Italy, nicola.amoroso@ba.infn.it. (2) Istituto Nazionale di Fisica Nucleare, Sez. di Bari, Bari, Italy. (3) Università di Bari, Dipartimento Interateneo di Fisica, Bari, Italy. (4) Università di Bari, Dipartimento di Economia, Management e Diritto dell'Impresa, Bari, Italy. (5) Invitalia-Agenzia Nazionale per l'attrazione Degli Investimenti e lo sviluppo d'impresa S.p.A, Roma, Italy.*

The "startup ecosystem" (the set of startups, funders and their relationships) has witnessed growing attention in recent years, both from an academic and a political point of view. In fact, startups are universally recognized as the "boosts" of technological innovation and, therefore, of economic growth. Despite this, it is, at the best of our knowledge, only rarely studied from a quantitative point of view. Our work primarily aims at assessing a mathematical model to investigate the complex relationships in the "startup ecosystem" and to give quantitative measures about the "importance" of the various economic agents. Secondly, we analyze how the elements' importance is related to the amount of money they will be able to collect in future years. According to these targets, we consider one of the most used dataset (for research purposes) about startups, funders and their relationships: "Crunchbase" (updated to 2017). The only quantitative variable in the dataset is the amount of money an element receives in every funding round it is involved in. We propose to model the "startup ecosystem" represented in Crunchbase as a complex network, where the nodes represent the elements of the ecosystem (startups and investors) and two nodes are connected by a directed link if one element (source node) funds another one (target node). Moreover, every node has been endowed with its own atttributes (economic category, investor type, headquarter's country and the overall-funds received in all the funding rounds it is involved in). In our work, we consider three different "centrality measures" to quantify the importance of nodes: "indegree" (representing the number of investors funding an element); "outdegree" (expressing the number of funding events a node is involved in); "betweenness" (describing the ability of an element to convey money within the startup ecosystem). Then we analyze differences between the only quantitative information given by the dataset and that represented by these network centralities. We find these differences applying the Kolmogorov-Smirnov test to the corresponding distributions and a novel information content conveyed by the network is observed. Then, we focus our attention on detecting the differences in the attributes characterizing the "strategic elements". They are defined using the "boxplot method" as the right outliers of the corresponding importance measure distribution. We define four classes of strategic elements: one for the overall-funding received and one for each of the network metrics. For the second target of our work, we find how the present role of an element in the startup ecosystem, as measured by its network centralities, influences its probability of being a funding-outlier in future years. Accordingly, we consider the Crunchbase data in the time period 2000 – 2017, create a complex network for each year in this interval (following the same method as in the previous part of this work) and find the annual funding-outliers. Then, we use a logistic regression model to relate the network centralities of an element at time T (=2000, …, 2017) with its probability of being a funding outlier at time $T+K$ (we consider K=1, …, 9). The proposed methodology sheds light on various aspects of the startup ecosystems that are rarely studied quantitatively. First, the network model naturally endows us with a set of measures representing various kinds of "importance" of the nodes in the considered system, taking into account the relationships between them. Secondly, this model gives us novel insights about the system, allowing us to get information otherwise unobtainable. Finally, the novel quantitative information provided by the network model can be used in machine learning models to forecast future properties of the elements.

Macroscopic properties of buyer-seller networks in online marketplaces

A. Bracci[1]*, J. Boehnke*[2]*, A. ElBahrawy*[3]*, N. Perra*[4]*, A. Teytelboym*[5]*, A. Baronchelli*[1,6,7]*. (1) Department of Mathematics, City, University of London, London EC1V 0HB, UK alberto.bracci@city.ac.uk, (2) Graduate School of Management, University of California Davis, 1 Shields Ave Davis, CA, USA, (3) Chainalysis Inc, NY, USA, (4) Networks and Urban Systems Centre, University of Greenwich, London SE10 9LS, UK, (5) Department of Economics, University of Oxford, Oxford OX1 3UQ, UK (6) UCL Centre for Blockchain Technologies, University College London, London WC1E 6BT, UK (7) The Alan Turing Institute, London NW1 2DB, UK*

Online marketplaces are a key driver of legal and illegal e-commerce. But the aggregate properties of buyer-seller networks in online marketplaces are very poorly understood. We analyse two datasets containing 245M transactions (16B USD) between 2010 and 2021 involving online marketplaces: 28 dark web marketplaces (DWM), unregulated markets whose main currency is bitcoin, and 144 product markets of one regulated e-commerce platform. We show how transactions in online marketplaces exhibit striking patterns of aggregate behavior despite significant differences in language, products, time, regulation, oversight, and technology. We find remarkable regularities in the distributions of (i) money amounts, (ii) number of transactions, (iii) inter-event times, (iv) time between first and last transactions. We then show how buyer behavior is affected by the memory of past interactions, and draw on these observations to propose a model of network formation able to explain the main stylised facts of the data. Our findings could have important implications for understanding competition, market power, and seller survival on online marketplaces.

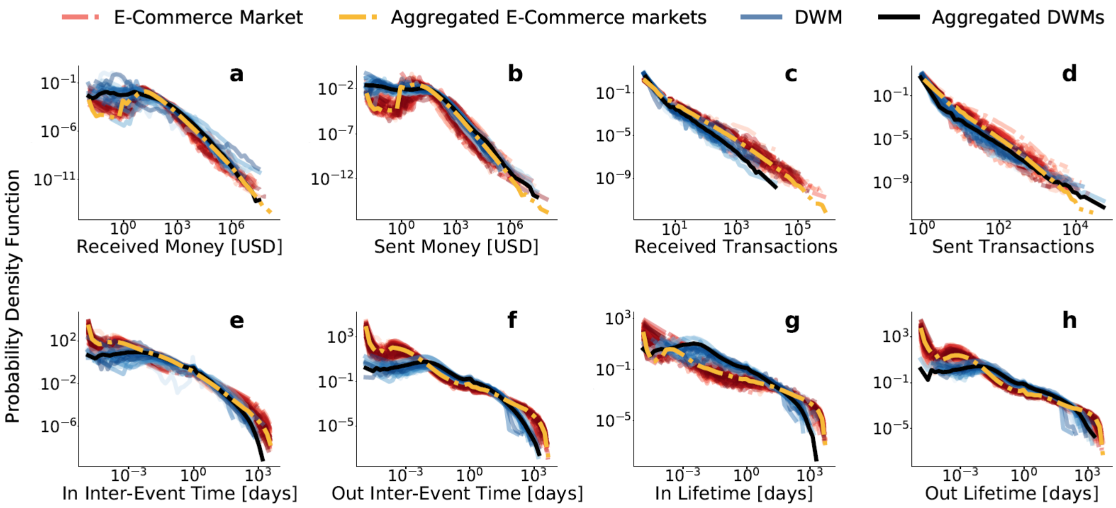


Figure 1: Distributions of static and temporal user quantities Top (a to d): Distributions of 4 analyzed users' aggregated quantities: money and number of transactions, both sent and received. Bottom (e to h): Distributions of 4 analyzed users' temporal quantities: the lifetime, time between first and last transaction, and the inter-event time, time between successive transactions, both measured in days. Each blue line represents one market, the black line is the distribution built aggregating all markets together. In dashed red each e-commerce platform market, while in dashed yellow the distribution aggregating all e-commerce markets. Striking patterns are observed across different DWMs or e-commerce categories, but also between regulated and unregulated online marketplaces, without any rescaling of the data.

Emergence of Risk Sensitivity in a First-principles Agent-based model

F. Bertolotti[1]*, A. Locoro*[1]*, L. Mari*[1]*. (1) Università Cattaneo - LIUC, Castellanza (VA), Italy, fbertolotti@liuc.it*

Risk aversion and risk-seeking are acknowledged features of human and non-human behaviour. The risk sensitivity of different entities seems to be shaped by the same underlying principles. Thus, there could be a regularity concerning wherewith living entities originate risk-sensible behaviour. Previous works employed theoretical work and simulations to examine if risk preferences could arise from various adaptation styles (cultural and genetic).

The purpose of this work is to understand how risk preferences emerge in a complex system, identifying existing relationships between the emergence, the adaptation styles and the environmental variables. This abstract presents an agent-based model in which a single variety of agents can move on a toroidal surface. Agents survive by consuming energy. When an agent terminates energy, it dies. Hence, agents move to collect energy, and they have to ways to do it. First, an agent can head in the direction of a passive energy source and supply from it. Second, it can move towards a different agent and attack it with a given probability of success. One alternative is risky, and the other is safe. Agents choose between these two possibilities computing a payoff. The decision-making weighs the distance from an option and its preference, which can be genetic or learned from experience.

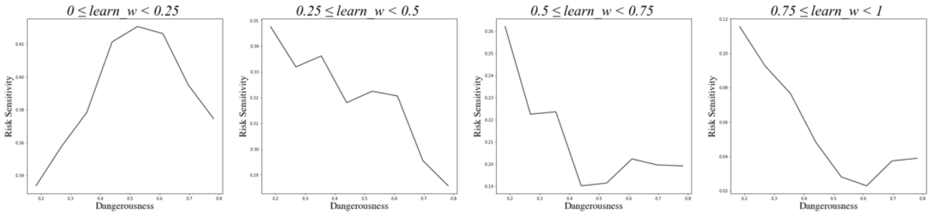


Figure 1: Risk sensitivity of population plotted on the dangerousness of the system with respect to the different style of adaptation, from evolution (left) to learning (right)

A sensitivity analysis was performed on the simulation model. It allowed identifying the influence of each parameter on the resulting risk preferences, such as the probability of success of aggression or the weight of adaptation styles on the decision-making. The analysis of the simulation results confirmed that both risk aversion and risk-seeking behaviours could emerge from the interaction of agents. Besides, we built a metric of environmental dangerousness to estimate the effect of every variable on the surviving rate of the population of agents. It seemed to exist a non-monotonic non-linear relationship between the dangerousness and the emergence of risk sensitivity, and that this relationship varies with the adaptation style (Figure 1). We propose the shape of the curve derives from the combination of two elements. One is the existence of incentives to develop risk preferences, which changes with the dangerousness of the surroundings. The second is a direct relationship between the harshness of the environment and the expected benefit of developing risk preferences. The simplicity of this model indicates that these findings could apply to different application fields. Future studies include the generalization of the results and the validation of the findings in real-world scenarios.

A networked global economy: the role of social capital in economic growth

J. Oliver[12*], *A. Antonioni (1)*[1], *F. Lipari*[1]*., I. Tamarit*[2], *(1) Carlos III University, Getafe, Madrid, Spain. (2) Clarity AI, Madrid, Madrid, Spain.* jaiolive@est-econ.uc3m.es.*

Understanding the drivers of macroeconomic output is one of the central questions in economics. Whereas the role of the first two factors of production--capital and labor--is well understood, the drivers of the Total Factor Productivity (TFP) are still open to discussion. Among the main signals studied in the literature, there is transmission and absorption of knowledge, factor supply, and economic integration (among others), but we lack intuitive, quantitative proxies for them.

Nowadays capital and labor flow across the globe through Global Value Chains and international migration, creating a network structure for the factors of production. Modern economic complexity measures suggest that the network structure of trade across countries plays an important role in economic growth [1] and a link between network centrality and different social capital measures has also been established [2].

In this work we benefit from the EORA Supply Chain Database [3] and the OECD International Migration Database to build the networks for the factors of production of 36 economies over 18 years. We compile a panel dataset with different social capital measures--proxied with network centrality--and we connect them to disparate factors affecting the TFP. Our results indicate that the *social capital* of countries in the factors of production network significantly drives growth through the TFP.

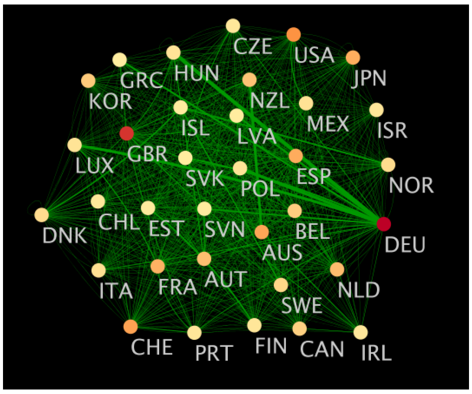


Figure 1. Snapshot of the 2015 weighted directed network structure of cross-country migratory flows. Edge weights represent migratory volume from one country to another relative to the home's country population.

References

[1] Hidalgo, C.A. and Hausmann, R., 2009. The building blocks of economic complexity. Proceedings of the national academy of sciences, 106(26), pp.10570-10575.

[2] Jackson, M.O., 2020. A typology of social capital and associated network measures. Social Choice and Welfare, 54(2), pp.311-336.

[3] Lenzen, M., Kanemoto, K., Moran, D. and Geschke, A., 2012. Mapping the structure of the world economy. Environmental science & technology, 46(15), pp.8374-8381.

MicroVelocity: Rethinking the Velocity of Money for Digital Currency Systems

C. Campajola[1]*, M. D'Errico*[2]*, C. J. Tessone*[1]*. (1) URPP Social Networks and UZH Blockchain Center, University of Zurich, Andreasstrasse 15, 8050 Zurich, Switzerland, carlo.campajola@uzh.ch. (2) European Central Bank, 60640 Frankfurt am Main, Germany*

In the quantity theory of money, the equation of exchange $MV=PQ$ states the proportionality between the amount of circulating money M and the price level P, the coefficient being the ratio between the velocity of money V and the volume of economic activity Q. In particular the velocity is relevant to monitor the demand for money and the correct functioning of the economy, and is typically measured by inverting the above identity using aggregated macroeconomic data. As many economic institutions make plans to develop digital versions of cash money such as Central Bank Digital Currencies, the ability to tune the measurement of V over micro-level, fine-grained data can be a game changer for regulators and policymakers in the implementation of effective monetary policy. Leveraging on the wealth of data made available by cryptocurrencies and adapting a theoretical framework developed in [1], in this work we propose a microeconomical velocity of money – the **MicroVelocity** - measuring for each economic agent their contribution to the average velocity at any point in time. We find that the distribution of MicroVelocities across cryptocurrency users is extremely heterogeneous and well approximated by a power law with exponent $\alpha\approx1.5$, and that MicroVelocity is in general negatively correlated with wealth and positively correlated with the number of addresses belonging to the user. This metric provides an entirely new angle to the analysis of monetary systems, highlighting the complexity and heterogeneity of agents which is otherwise neglected by classical economic theory, potentially endowing regulators and policymakers with a powerful tool to monitor the functioning of present and future digital currency systems.

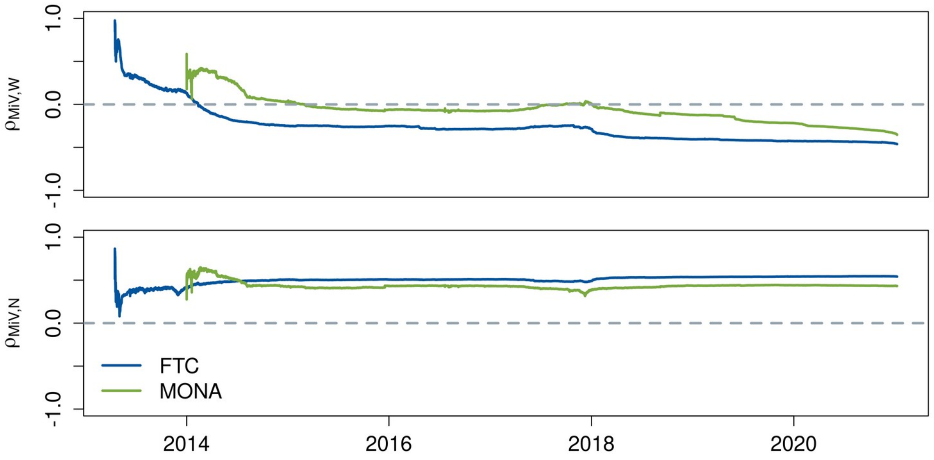


Figure: MicroVelocity in Feathercoin (FTC) and Monacoin (MONA). (top) Spearman's rank correlation ρ between user wealth and MicroVelocity as a function of time; (bottom) Spearman's ρ between the number of addresses belonging to a user and their MicroVelocity.

References

[1] Y. Wang, N. Ding, L. Zhang. "The circulation of money and holding time distribution" *Physica A: Statistical Mechanics and its Applications*, 324 (2003) 665-677.

Life Cycle Assesment-based Networks and Multiscale Representations

__A. Hazan__[1]. (1) Université Paris-Est Créteil Val-de-Marne, IUT Sénart-Fontainebleau, 36-37 rue Georges Charpak, 77567 Lieusaint, France, aurelien.hazan@u-pec.fr.

In spite of rising climatic and environmental concern, modeling of Input-Output theory, Industrial Ecology, Life Cycle Assessment, and Material Flow Analysis are conspicuously under-represented in complex networks research. To fill that gap we compare various network representations of interactions occurring in the aforementioned domains, following [1], and extending the scope (e.g. to chemical reactions networks, and the Von-Neuman's growth problem).

First the different methods to construct such networks are summarized, starting from available data. Then, a classical empirical study of network's properties is conducted and shows a high heterogeneity. More precisely, networks stemming from Life Cycle Inventory databases are compared to networks built using mappings with actual industrial process operated by economic actors, such as hybrid Life Cycle Analysis, based on Input-Output Analysis at international, national, and even regional scales.

The benefits of various pre-processing steps such as monopartite projection are assessed and higher-order formalism (e.g, hypergraphs) used for example in omics research are discussed.

Existing applications, for example community finding [2] are exposed, and limitations of conducting a network-oriented analysis are debated (such as data availability, uncertainty in databases, or lack of spatial information).

Special emphasis is drawn on criticisms raised over the interest of the network formalism with respect to scale-dependence. The way this topic is dealt with by linear algebraic tools [3] that are standard in those fields is analyzed in that light. The interest of mitigating this issue using recent multi-scale representations [4] and aggregation techniques in complex networks research is assessed.

References

[1] R.Heijungs, *Journal of Environmental Accounting and Management*, 3(2) (2015) 151.

[2] S. Kagawa, Okamoto, Suh, Kondo, Nansai, Social Networks 35 (2013).

[3] R.Wood, M.Lenzen, Journal of Industrial Ecology, 13(2) (2009).

[4] A.Hazan, *Proc. of French Regional Conference on Complex Systems, Dijon*, France, 2021.

Using Agent Based Modelling to Test Novel Remedies for Restoring Competition in Monetized Digital Markets

Todd Davies*[1]*. (1) Technical University Munich, 80636, Munich, Germany, todd.davies@tum.de

The rise of Chicago School doctrine in the field of Antitrust law oriented competition policy in major jurisdictions around the framework of neoclassical economics [1]. This viewpoint considered *consumer welfare* to be the main aim of antitrust laws, manifesting in regulatory activity that prioritized lower prices and higher output. Contemporary scholarship, motivated by current political economic debates, has highlighted the value of a broader set of goals. These include supply chain resilience, consumer privacy, and the distribution rather than accumulation of private power. Such concerns are hard to model using the framework of neoclassical economics.

This paper applies methods from complexity economics [2] to the question of how market power is accrued in zero-priced digital markets, and models the effect of various regulatory remedies. In such markets, auxiliary services such as advertising fund consumers' usage of digital products. Examples of such markets include social networks and online search. An agent-based model is used to simulate the accumulation of market power in monetized digital markets. Stylized facts about the function of such markets are recovered. It then investigates the effect of proposed regulatory remedies which aim to restore competition in these markets. Remedies which aim to alter the behavior of firms [3] and those which aim to change the structure of the market [4] are simulated, and their results analyzed.

Acknowledgements

I would like to thank my supervisors – Dr. Zlatina Georgieva and Dr. Prof. Amy Pond for their excellent mentorship. I would also like to thank the Young Researchers of the Complex Systems Society for their work to promote complex systems scholarship.

References

[1] Bougette, Patrice, Marc Deschamps, and Frédéric Marty. "When economics met antitrust: The Second Chicago School and the economization of antitrust law." Enterprise & Society 16.2 (2015): 313-353.

[2] Arthur, B., E. Beinhocker, and Allison Stanger. "Complexity Economics: An Introduction" *Proceedings of the Santa Fe Institute's 2019 Fall Symposium. SFI Press*. 2020.

[3] Gal, Michal, and Nicolas Petit. "Radical restorative remedies for digital markets." *Berkeley technology law journal* 37.1 (2021).

[4] Davies, Todd. "Reconceiving Big Tech: Decentralizing the Industry to Align Incentives and Maximize Total Welfare" *Concurrences Antitrust Writing Awards* 2021 (2021).

Characteristics of Pure Nash Equilibria in a Strategic Multiplex Network Formation Game

T.E. Willis[1]*, G. Punzo*[1]*. (1,2) Department of Automatic Control and Systems Engineering, The University of Sheffield. Sheffield, UK tewillis1@sheffield.ac.uk.*

The problem of strategic network formation (NF) unites elements of game theory and network theory. The setting applies to interactions of selfish players who may syndicate to share benefits of cooperation. In single layer networks, analytic results exist regarding the price of anarchy for pure Nash equilibria [1]. This work builds upon such results and expands them to a multiplex setting. We consider a multiplex network problem where players in one layer receive benefits from accessing nodes (e.g. physical location, services, etc…) in the other but this comes at a premium. To minimise the cost of a direct inter-layer connection, players establish links (coalitions) with nodes in their same layer who are willing to pay a premium to access the second layer. This could represent, for example, bus companies forging alliances to offer passages onto one another's routes and so connect to more nodes of a network of regional airports, which is contained in the second layer. Alternatively, the layers could represent congestion and information transfer with nodes communicating their findings and getting information in one layer about which parts of the second layer are congested. In this case, the players would seek to minimise the chance of encountering congestion when using the second layer, and to do so, they are willing to establish links with their peers. We looked at the nature of the equilibria under NF in the new model, with specific regard to the nature of the inter-layer interaction and the spectral radii of the adjacency matrices of the networks formed. To find strict bounds on the relationship between the spectral radii in separate layers and the outcome of the NF game, we introduce the incentive towards originality β, a variable which either encourages or discourages deviation from the network edges existing in other layers. This has the effect of changing the cost functions of Fabrikant's single layer model, with implications on the outcome of the NF game. Limitations on bandwidth within internet communication is an example of where utilising different edges may be advantageous, whereas in a transport infrastructure application companies that cooperate with respect to one type of vehicle are more likely to cooperate with other types of vehicle and so the use of similar edges is advantageous.

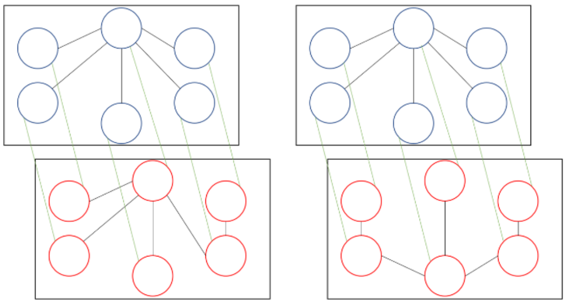


Figure 1: Sample Pure Nash Equilibria for β=-1 and β=1

[1] A. Fabrikant, A. Luthra, E. Maneva, C. Papadimitriou, S. Shenker. *On a Network Creation Game PODC '03: Proceedings of the twenty-second annual symposium on Principles of distributed computing.* 2003

Optimal Legal System in a Semi-Underground Economy

K. Mayer[1], *(1) Independent, Einsteinova 34, Bratislava, Slovakia, katarina.mayer1@gmail.com.*

Legal treatment of prostitution varies significantly among countries, from criminalizing all prostitution-related activities, through legalizing in attempt to regulate prostitution and its negative side effects to decriminalizing the sexual acts as a private matter between consenting adults [15]. Recognizing unequal position of the market agents in the sex market some countries chose to adopt mixed policies to better address the problem. The most frequently discussed example is so called "Swedish model" that decriminalizes the sale of sexual services but criminalizes all other related activities [4].

Besides a large number of studies attempting to evaluate and/or compare different legal responses to prostitution—in the area of law: [2], [9], [11], and [16]; women studies: [1], [4], and [5]; psychology: [8] and [7]; and sociology: [3], [14] and [15]—chosing which system is optimal remains difficult. First, different market agents tend to evaluate the success or failure of a legal system differently. Second, some legal systems while theoretically possible are not currently adopted anywhere, which makes them difficult to evaluate. Third, because the market for sex services is at least partially an underground economy, it is very difficult to obtain reliable data. As there is no synchronized framework for collecting the data, studies differ in the type of data collected and in how they define variables. Finally, the data collected under different legal systems is not directly comparable. Even under the same legal system, countries likely vary in factors that directly influence the performance of a given legal system such as the intensity of police enforcement and the level of corruption. For all these reasons, the only meaningful way to compare the performance of different legal systems is to build a model that controls for the differences in legal systems.

The current paper attempts to do just that—identify optimal legal system using an agent-based model. The model was motivated by Salop's model of spatial competition but it differs from the original model in two important aspects: agents are not restricted in their movement and both markets—legal and underground—are modeled as equal. Thus the model studies behavior of underground economy agents who compete with legal economy agents.

The model has three types of agents—prostitutes, pimps and punters. Each of these agents gets monitored at random by the government (depending on a legal system in place). The goal is to identify an optimal legal system under which negative impact of sex industry is minimized for the society.

Despite its simplicity —agents have minimal set of behaviors—the model can capture general dynamics we observe in real sex markets and therefore can be used as a proxy for realistic modeling of sex markets. Furthermore, it allows us compare different policies (even those nonexistent in real markets) and their social impact. Comparison of the legal systems shows that while no legal system is superior under all circumstances, under realistic parameter values, policy when sale of sexual services is legal and both purchase and pimping are illegal provides best results.

References

[1] Bindell, J., & Kelly, L. (2003). A Critical Examination of Responses to Prostitution in Four Countries: Victoria, Australia; Ireland; the Netherlands and Sweden. Abuse Studies Unit, London Metropolitan University.

[2] Bingham, M. (1998). Nevada Sex Trade: A Gamble for the Workers. *Yale Journal of Law and Feminism* 10, 69-99.

[3] Brents, B. G., & Hausbeck, K. (2005). Violence and Legalized Brothel Prostitution in Nevada: Examining Safety, Risk and Prostitution Policy. *Journal of Interpersonal Violence*, 20, 270-295.

[4] Egberg, G. (2004). The Swedish Law that Prohibits the Purchase of Sexual Services: Best Practices for Prevention of Prostitution and Trafficking in Human Beings. *Violence against Women*, 10, 1125-1218.

[5] Farley, M. (2004). Bad for the Body, Bad for the Heart: Prostitution Harms even if Legalized or Decriminalized. *Violence against Women*, 10, 1087-1125.

[6] Farley, M. (2007). Prostitution and Trafficking in Nevada: Making the Connections. San Francisco: Prostitution Research and Education.

[7] Farley, M., Isin, B., Kiremire, M., & Segzin, U. (1998). Prostitution in Five Countries. *Feminism and Psychology*, 8, 405-426.

[8] Farley, M. A., Cotton, A., Lynne, J., Zumbeck, S., Spiwak, F., Reyes, M. E., Alvarez, D., & Segzin, U. (2003). Prostitution, Trafficking and Traumatic Stress of Prostitution, New York: Haworth Maltreatment & Trauma Press.

[9] Farley, M. A., & Kelly, V. (2000). Prostitution: A Critical Review of the Medical and Social Sciences Literature. *Women and Criminal Justice*, 11, 29-64.

[10] Hindle, K., Barnett, L., & Cassavant, L. (2008). Prostitution: A Review of Legislation in Selected Countries. Parliamentary Information and Research Service, Library of Parliament, Canada.

[11] Johansson, M. C., & Kelemen, K. (2013). Still Neglecting the Demand that Fuels Human Trafficking: A Study Comparing the Criminal Laws and Practice in Five European States on Human Trafficking, Purchasing Sex from Trafficked Adults and from Minors. *European Journal of Crime*, 21, 247-289.

[12] Jordan, J. (2005). The Sex Industry in New Zealand: A Literature Review. Discussion Paper, Ministry of Justice, New Zealand.

[13] Kara, S. A. (2009). Sex Trafficking: Inside the Business of Modern Slavery. New York: Columbia University Press.

[14] Larsen, N. E. (1996). The Effect of Different Police Enforcement Policies on the Control of Prostitution, *Canadian Public Policy*, 22, 40-55.

[15] Shaver, F. M. (1985). Prostitution: A Critical Analysis of Three Policy Approaches. *Canadian Public Policy*, 11, 493-503.

[16] Waltman, M. (2011). Prohibiting Sex Purchasing and Ending Sex Trafficking: The Swedish Prostitution Law, *Michigan Journal of International Law*, 33, 130-157.

Cryptocurrency News Sentiment Networks

Ana Todorovska[1], Eva Spirovska[1], Gorast Angelovski[1], Hristijan Peshov[1], Ivan Rusevski[1], Jovana Marojevikj[1], Irena Vodenska[2,3], Lubomir Chitkushev[2,4] and Dimitar Trajanov[1]

[1]Faculty of Computer Science and Engineering, Ss. Cyril and Methodius University in Skopje, Republic of North Macedonia
[2]Laboratory for Interdisciplinary Finance and Economics (LIFE) Research, Metropolitan College, Boston University, USA
[3]Administrative Sciences Department, Financial Management, Metropolitan College, Boston University, USA
[4]Computer Science Department, Metropolitan College, Boston University, USA
E-mail: ana.todorovska@finki.ukim.mk

Abstract

In a world where no country, market, or economy is an island, interconnectivity is becoming a fundamental feature of almost all social and economic systems. In the case of digital assets like cryptocurrencies, the impact of interconnectivity on their performance and price trajectory is amplified. Studying these phenomena is essential for understanding the processes driving the crypto-markets. In this paper, we propose seven different approaches to create a network of eighteen most important cryptocurrencies. The first three approaches discover correlations between cryptocurrencies based on their daily prices, daily returns, and sentiment extracted from Reddit data. The following two approaches offer insights from the frequency of joint appearance of cryptocurrencies in Google news and Reddit data. The remaining two approaches determine each cryptocurrency's impact over the others when forecasting prices and returns. Furthermore, we explore the networks' interdependencies to explore the similarities of the cryptocurrency networks generated by different approaches. The proposed methodology allows us to understand the dynamics in the cryptocurrency markets and the different processes that influence their performance.

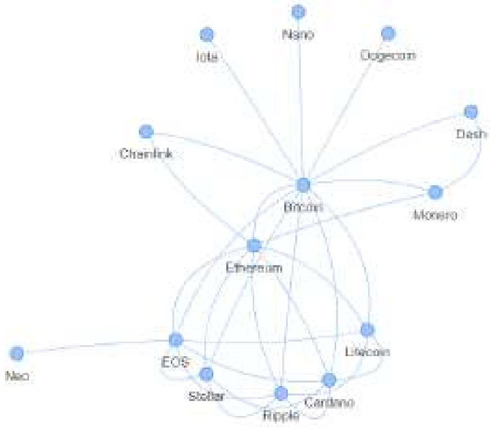


Figure 1: Network based on cryptocurrency co-occurrences in Reddit titles

4 Infrastructure, Planning and Environment

Contents

A Lightweight Approach for Origin-Destination Matrix Anonymization

B. Matet[1,2], *E. Côme*[1], *A. Furno*[2]. *(1) Univ. Gustave Eiffel, COSYS, GRETTIA F-77447 Marne-la-Vallée, France, benoit.matet@univ-eiffel.fr. (2 Univ. Gustave Eiffel, Univ. Lyon, ENTPE, LICIT,F-69518, Lyon, France*

Personal trajectory data are a category of personal data consisting in sequences of spatiotemporal points that represent the trajectory of users over a given timespan. It is clear that a dataset of trajectories must be anonymized in order to protect user's privacy. Data protection approaches rely mostly on k-anonymization [1], which is achieved when all users in the dataset are indistinguishable from at least k−1 other users. k-anonymization is usually attained through generalization and suppression, i.e., replacing exact values with intervals and deleting outliers. However, processing of whole trajectories with k-anonymization proves to be difficult to achieve for k > 5 and does not offer a truly foolproof anonymization scheme. Moreover, we argue that according to European regulation, any link inferred between two spatiotemporal points constitutes a data breach. As such, k-anonymized trajectories do not offer enough protection to be safely released in open data. We propose simplifying the problem to its extreme by considering trajectories defined only by their first and last spatiotemporal points, i.e., an Origin/Destination (OD)-matrix. Although they leave out most of the trajectory information, OD-matrices are a key element in the transport analysis framework as they can be used to understand the dynamic of transport demand, make long-term prediction, and allow for transport simulations. We achieve k-anonymization of an OD-matrix with $6 \leq k \leq 16$ by setting a quadtree structure for the spatial generalization of origins and destinations separately. The aggregation of origins consists in finding a quadtree whose set of leaves represent a spatial partition of the area of study. We formalize the problem as finding the tree that minimizes the Mean Squared Error compared to a given target outgoing volume. Each origin leaf is then associated to a second quadtree-based spatial partition for destination areas. The problem of finding destinations is formalized as finding the tree that minimizes a generalization error [2]. Resulting volumes v_{od} from an origin o to a destination d are considered anonymized if $v_{od}>k$ and are suppressed if $v_{od}\leq k$. As the objective functions for both problems are modular, we can easily find the optimal solutions, as well as the best solution under a given suppression threshold, by reducing it to a Tree Knapsack Problem. These simplifications let us find the least destructive anonymization as measured by the generalization error for high values of k, ensuring complete anonymization of the data for a reasonable computational cost. We apply our approach at scale on massive mobile network data from French telecom operator Orange. The results show that we are able to reconstruct the hourly profiles for in- and outgoing mobility flows of areas in Lyon and retrieve the land use of the city via clustering of the profiles. We are also able to retrieve typical profiles of OD-matrices with respect to the hour of the day via Latent Dirichlet Allocation (LDA).

Acknowledgements

This work is supported by the French ANR research project PROMENADE (grant number ANR-18-CE22-0008).

References

[1] M. Fiore, P. Katsikouli, E. Zavou, M. Cunche, F. Fessant, D. Le Hello, U. Matchi Aivodji, B. Olivier, T. Quertier, and R. Stanica. Privacy in trajectory micro-data publishing : a survey, 2020.

[2] Y. Liang and R. Samavi. Optimization-based k-anonymity algorithms. Computers Security, 93:101753, 2020.

The Evolution of City Infrastructure Scaling and Spatial Correlations in the United States: 1900-2015

K. Burghardt[1]*, J. Uhl*[2,3]*, K. Lerman*[1]*, S. Leyk*[4]*. (1) USC Information Sciences Institute, Marina del Rey, USA, keithab@isi.edu. (2) Institute of Behavioral Science (IBS), University of Colorado Boulder, Boulder, USA, (3) Cooperative Institute for Research in Environmental Sciences (CIRES), University of Colorado Boulder, Boulder, USA, (4) Department of Geography, University of Colorado Boulder, Boulder, USA*

Over 50% of the world population, and over 80% of the US population live in cities or other urban areas. These areas cause a disproportionate amount of pollution and congestion, which affects our health and quality of life. We aim to better understand a key component mediating pollution and congestion within cities: the street network and its evolution over time as cities have expanded. We conduct one of the largest and longest longitudinal studies in this field by reconstructing and analyzing the evolution of over 850 US urban street networks since 1900, based on the integration of novel geospatial data sources. We extract different measures from these spatio-temporal networks, including road network lengths and approximate urban areas, examine how these measures are related to city population and assess their spatial autocorrelation over time. As can be seen in Fig. 1a, the length of road networks scale nearly linearly with population in 1900 but changes to strongly sublinear by 2015, suggesting a varying economy of scale. Similar results are seen for urban areas within a city versus population (Fig. 1b). We also find that the correlation of city-level statistics, such as road density within cities, exhibits strong spatial patterns at a regional scale. This correlation decreases only logarithmically with distances between cities and have only slightly decreased over 100 years (Fig. 1c,f). Finally, Figs. 1d & e show metropolitan areas shaded by their scaling exponents derived from individual city road length or built area versus population plots. These findings are in strong contrast to theoretical expectations that there is a universal scaling law across or within cities. Moreover, the spatial correlations suggest that cities do not grow independent of each other, thus scaling relations have a strong regional dependence. Similarities over time and space, however, also suggest common mechanisms of growth.

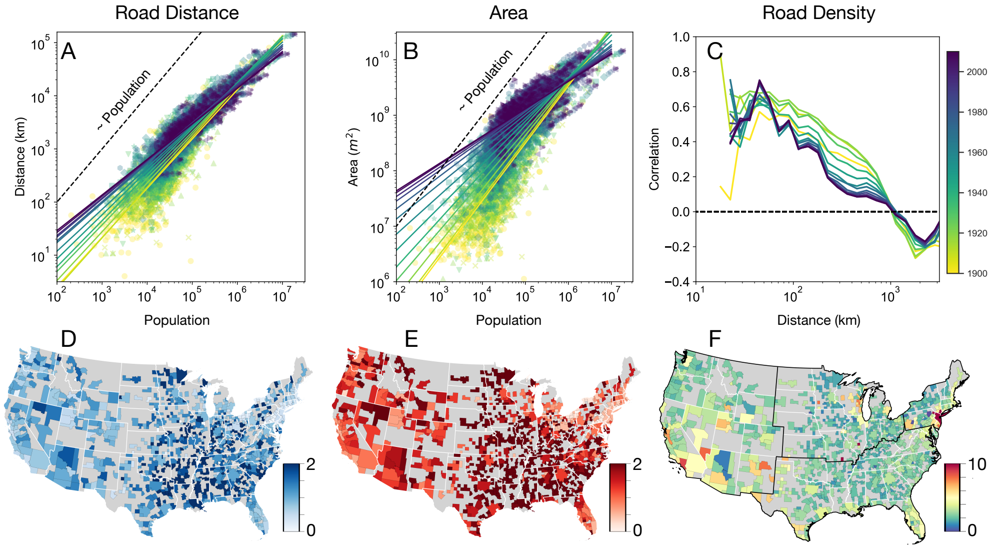


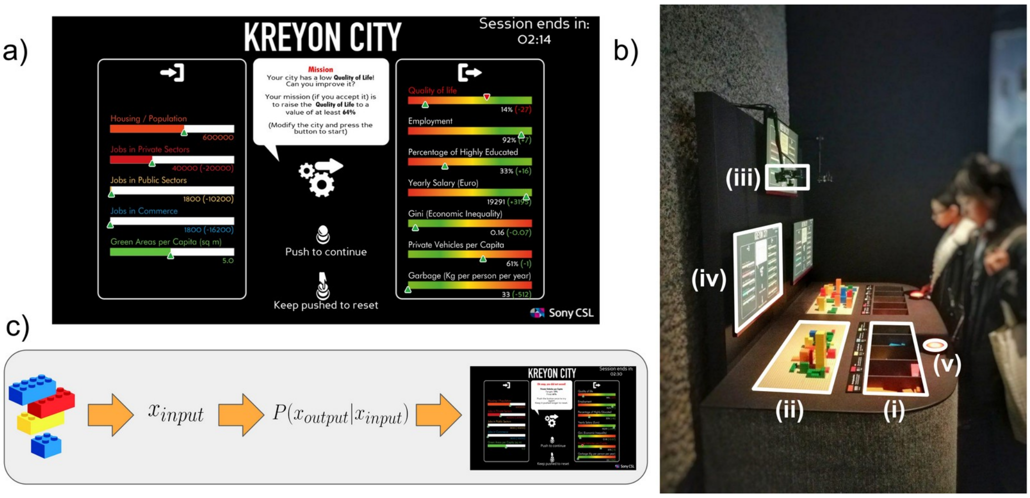
Figure 1: Road network scaling. Scaling laws between (a) road distance or (b) urbanized area and population across cities from 1900-2015, (c) Spatial correlations of measures of road density between cities over time, (d) map of scaling law exponents between road lengths and city population over time, (e) map of scaling law exponents between urbanized area and city population over time, and (f) road density (km of road per square km) in 2015.

Finding Successful Strategies in a Complex Urban Sustainability Game

B. Monechi[1], *E. Ubaldi*[1], *P. Gravino*[1] , *I. Chabay*[2,] *V. Loreto*[1,3,4]

(1) Sony Computer Science Laboratories, 6, Rue Amyot, 75005, Paris, France, bernardo.monechi@sony.com. (2) Institute for Advanced Sustainability Studies, Berliner Strasse 130, 14467 Potsdam, Germany (3) Sapienza University of Rome, Physics Department, Piazzale Aldo Moro 2, 00185, Rome, Italy (4) Complexity Science Hub Vienna, Josefst\"adter Strasse 39, A-1080 Vienna, Austria

The adverse effects of unsustainable behaviors on human society are leading to an increasingly urgent and critical need to change policies and practices worldwide. This requires that citizens become informed and engaged in participatory governance and measures leading to sustainable futures. Citizens' understanding of the inherent complexity of sustainable systems is a necessary (though generally not sufficient) ingredient for them to understand controversial public policies and maintain the core principles of democratic societies. In this work, we present a novel, open-ended experiment where individuals had the opportunity to solve model urban sustainability problems in a purposeful game. Participants were challenged to interact with familiar LEGO blocks representing elements in a complex generative urban economic indicators model. Players seeks to find a specific urban configuration satisfying particular sustainability requirements. We show that, despite the intrinsic complexity and non-linearity of the problems, participants' ability to make counter-intuitive actions helps them find suitable solutions. Moreover, we show that through successive iterations of the experiment, participants' can overcome the difficulties linked to non-linearity and increase the probability of finding the correct solution to the problem.



Mission
Your city has a low **Quality of Life**! Can you improve it?
Your mission (if you accept it) is to raise the **Quality of Life** to a value of at least **64%**
(Modify the city and press the button to start)

Figure 1: (a) Example of the interface. The generativ model's input is based on the number of bricks for each color present in the building area and are shown on the left bars. The city's status is represented by bars on the right (e.g. employment rate, produced garbage). Players are asked to find a specific value for the city's stauts by modifing the number of bricks on the board. (b) The installation with two gaming stations with brick conainers (i), a building area (ii) with an RGB/depth-sensor (iii) above it, a feedback monitor (iv) with a red button to control it (v). (c) Scheme of user interaction during a task. Players modify the city, then pressing a button they trigger the computation of the city status with a generative model, then the output is displayed on the monitor.

Understanding and Forecasting the Illicit Drug Distribution in England: a Data-Centric Approach to the County Lines

Leonardo Castro-Gonzalez[1, 2] lcastrogonzalez@turing.ac.uk *(1): University of Birmingham, Edgbaston, United Kingdom (2): The Alan Turing Institute, British Library, 96 Euston Road, London, United Kingdom*

For over a decade, the illicit drug market in Great Britain has developed a new distribution method, acknowledged as 'the County Lines Model' by the UK Gov. This model has caused important problems regarding modern slavery and coerced the polices' capacity to dismantle the network [1], making it a top priority for the government [2]. Our goal is to forecast the County Lines dynamics between English police forces, so an efficient coordination is possible. To the best of our knowledge, this is the first attempt to study the CL Model quantitatively.

We use the Boltzmann-Lotka-Volterra (BLV) model [3] to forecast the spatial dynamics, i.e. the flow of county lines from one place in England to another. The BLV model considers the benefits and costs of establishing a *line* between locations. Using data from the Metropolitan Police of London, we infer the most influential variables for the County Lines dynamics. We validate our results using a differences equation given by the BLV model, to then forecast the number of lines going from one police territory to another.

Among the different variables tested, the most influential are the knife crime events and the police workforce present in the territory, resulting as important costs for operators. This is an interesting result given that knife crime is another top security priority for the UK gov. When forecasting the number of lines throughout England, we observe how there is decrease of activity from 2019 to 2020. This can be seen in the heatmaps presented in Figure 1.

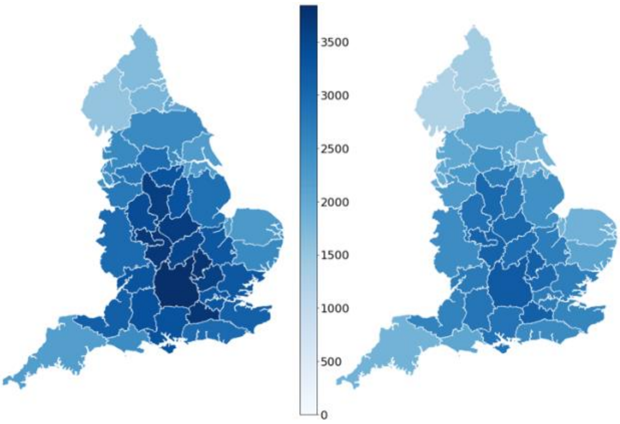


Figure 1. Heatmap of forecasted County Lines in England in 2019 (left) and 2020 (right). Divisions account for the different local police forces in England.

References

[1] R. Coomber, L. Moyle. The Changing Shape of Street-level Heroin and Crack Supply in England: Commuting, Holidaying, and Cuckooing Drug Dealers across 'County Lines', Brit. J. of Criminol., **58**: 1323-1342, 2018

[2] C. Black. Review of Drugs: evidence relating to drug use, supply and effects, including current trends and future risks. U.K. Gov. Home Office, Feb. 2020

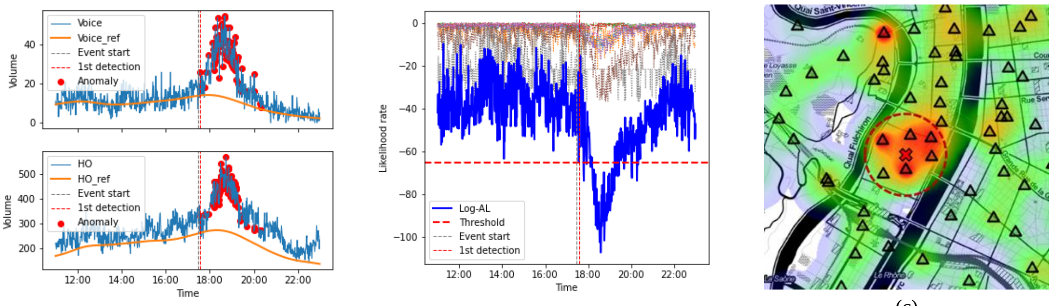
[3] A. Wilson. Boltzmann, Lotka and Volterra and spatial structure evolution: an integrated methodology for some dynamical systems. J. Roy. Soc. Interface, **5**(25): 865-871, 2008

Unsupervised Real-time Detection of Urban Anomalies Using Massive Mobile Phone Network Data

E. Akopyan[1], *A. Furno*[1], *N-E. El Faouzi*[1], *E. Gaume*[2]. *(1) Univ. Gustave Eiffel, Univ. Lyon, ENTPE, LICIT, F-69518 Lyon, France, evelyne.akopyan@univ-eiffel.fr. (2) Univ. Gustave Eiffel, GERS, LEE, F-44344 Bouguenais, France*

Urban areas of large cities are complex systems involving the interaction of people, ICT infrastructures, transport means, that are frequently impacted by perturbations and hard-to-predict anomalies of different nature. Such phenomena can produce dramatic cascading effects with tremendous socio-economic consequences for the masses of population present in these areas. Monitoring urban areas with real-time anomaly detection systems relying on massive data represents a recent research field with challenging requirements.

We propose a lightweight and robust framework for real-time anomaly detection in multivariate time-series extracted from large-scale Mobile-phone Network Data (MND). Our solution relies on unsupervised machine learning applied to MND collected at individual antennas of the largest mobile phone network provider of France, Orange. The framework requires a minimal number of parameters to be set and is based on statistical models that ensure lightweight computation and almost instantaneous anomaly detection, which makes our solution adequate for real-time urban monitoring, especially on densely populated areas. Our solution is based on a two-step approach: (i) the offline stage aims at assessing the typical behavior of the antennas of the region of interest; (ii) the online stage performs real-time comparison of incoming data with respect to the detected typical behavior. Results related to a real case-study of terrorist attack in Lyon showcase that our framework can successfully detect an emergency situation almost instantaneously and locate the anomalous area with high precision. Nodes that are farther away from the attack location do not detect any anomaly at all along that day, thus making the anomaly easy to locate with accuracy at the scale of the city as shown in Fig. 1.



(a) (b) (c)

Figure 1: Real-time detection: comparison with the typical behavior of the MND traffic for different mobile network services over time (a), anomaly likelihood rate temporal evolution (b), and spatial extent of the detected anomalous area (red circle) (c)

Additional results related to different case studies will be presented as well as a detailed analysis of the performance of our approach, with a particular focus on its accuracy with respect to known anomalous and normal situations.

Acknowledgements

This work is supported by the French ANR research projects DISCRET (grant number ANR-19-FLJO-0002-01) and PROMENADE (grant number ANR-18-CE22-0008).

Derivation of the potential field of collective human flow as an analogy of electric circuits

Yohei Shida[1], *Jun'ichi Ozaki*[2], *Hideki Takayasu*[2,3], *Misako Takayasu*[1,2]. *(1) Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan, shida.y.aa@m.titech.ac.jp. (2) Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan (3) Sony Computer Science Laboratories, 3-14-13 Higashi-Gotanda, Shinagawa-ku, Tokyo, Japan*

For the last decade, scientific studies of human mobility using human location data can be roughly categorized into two: One is focusing on statistical properties of individual human trajectory patterns (microscopic scales)[1], and the other is mass human movement between the origin and the destination (macroscopic scales)[2]. Collective human flow, which cumulates the movement of a large number of people at each point, makes it possible to generate a vector field in a metropolitan area and opens a new door for human mobility patterns (mesoscopic scales)[3, 4]. It is argued that rotational flow does not contribute in the collective human movement in urban area, so a scaler potential function for the velocity can be estimated from the data [4]. Here, we introduce an analogy of an electric circuit to regard human flow as electrical current and the velocity potential as voltage.

Our GPS data, provided for scientific study by a Japanese private company Agoop, is consisted of user's ID, individual locations and velocities for about 250,000 mobile phone users in Japan with about 30 minutes data transmission intervals.

We divide the area into 500 m × 500 m grid-like squares and calculate the population density and mean velocity in each square every 30 minutes. Regarding the number of people as the number of charged particles, the product of the population density and the mean velocity corresponds to the electric current $\vec{I}$ for each square. We consider that each square has a resistance R, then the vector field $\vec{I}R$ (product of current $\vec{I}$ and resistance R) should be rotation free. We estimate the value of R to minimize the sum of squares of the rotation. The estimated value of resistance tends to take a low value for those squares including railways or highways.

The morning and evening potentials around the Tokyo metropolitan area are valley-shaped (mountain-shaped) centered on the central Tokyo. These are consistent with the results that human flow patterns in the commuter and returning rush have specified directions[3]. On the other hand, the potential in the afternoon is nearly flat, which is also consistent with the result that the daytime flow patterns is close to random[3].

References

[1] M. C. González, C. A. Hidalgo, and A. L. Barabási, Nature **453**, 779 (2008).

[2] F. Simini, M. C. González, A. Maritan, and A. L. Barabási, Nature **484**, 96 (2012).

[3] Y. Shida, H. Takayasu, S. Havlin, and M. Takayasu, Sci. Rep. **10**, 1 (2020).

[4] M. Mazzoli, A. Molas, A. Bassolas, M. Lenormand, P. Colet, and J. J. Ramasco, Nat. Commun. **10**, 1 (2019).

Deploying for Resilient Cities: Utilizing Social Media Network Data to Aid in Humanitarian Assistance and Natural Disaster Relief

Thomas Y. Chen[1], *(1) Academy for Mathematics, Science, and Engineering, 520 W Main Street, Rockaway, NJ 07058, USA, thomaschen7@acm.org*

Satellite images and social media images have for some time been generally seen by the research community as almost equivalently valuable data sources for training deep learning models for disaster relief and management. However, given that social media is a more versatile source of dataset curation, we posit that it is the most applicable means for computational damage assessment. The multi-temporal aspect of datasets sourced from platforms like Twitter, Facebook, Flickr, Instagram, etc., rather than having simple bitemporal properties, allows for researchers to use a wide variety of temporal and geographical locations in their models. In essence, social media networks are multi-temporal catalogues/repositories for pre-, present, and post- disaster scene and object imagery. This reduces overfitting and allows for better cross-region and cross-temporal transfer learning. Additionally, obtaining satellite and GIS data is often very expensive. Extracting social media imagery is often a simple matter of web mining, which only requires computational power, which is far more available to researchers and students. In tandem to this advantage, the rapid access to social media networks allows for the timely analysis and subsequent response. Furthermore, the quality of smart phones in the 21st century complements the value of the data mined from the web. In satellite images, many times the results are nonoptimal precisely because the data is not as high-quality as desired, given that each building, structure, or person within each image and its corresponding annotated bounding box constitute a much smaller percentage of the total image. Therefore, in many instances, the more close-up, high quality nature of images of buildings and other infrastructure found on social media provide more optimal opportunities for the assessment of damage and change detection. Because social media is often the primary method by which users communicate with friends, family, and the world, these platforms can be relied upon to contain a diverse set of data from the disaster site immediately upon the natural disaster's commencement. By collecting and analyzing this data in real time, systems developed in academic research can be seriously put to use in crisis situations. Although further work has to be conducted in the realm of interpretability of the models (opening black boxes, keeping end users and operators in mind and preventing unforseen biases), such architectures have already been utilized for real-world deployment. Although the widespread use of social media, specifically for posting high-quality images, is a relatively recent development, the popularity of deep learning approaches is also very recent. Therefore, their combination is quite natural and many cutting-edge methodologies have been published. Researchers at the Qatar Computing Research Institute (QCRI) have particularly focused in this area, presenting Image4Act, an end-to-end pipeline for social network image content machine learning analysis to help humanitarian organizations direct targeted relief and recovery efforts. Image4Act has already been used during a recent cyclone, demonstrating is real-world applicability. While further work remains to be completed, such as conducting in-detail experiments and ablation studies on the recent Incidents Dataset and finding additional optimal sources of data collection for damage and natural disaster assessment, it can be concluded that social media present themselves as burgeoning opportunities in this area, as well as other related fields and subfields. The interdisciplinary nature of this research, combining the expertise of artificial intelligence and computer scientists, industry professionals, and temporal web analysts with the needs of end operators in disaster-prone and underserved regions, promotes future scientific and academic cooperation in the pursuit of improved humanitarian aid mechanisms.

Urban mobility data uncovers socioeconomic segregation in the US

R. Hilman[1], *G. Iñiguez*[1,2,3,], *M. Karsai*[1]. *(1) Department of Network and Data Science, Central European University, 1100 Vienna, Austria, Hilman_Rafiazka@phd.ceu.edu (2) Department of Computer Science, Aalto University School of Science, 00076 Aalto, Finland (3) Centro de Ciencias de la Complejidad, Universidad Nacional Autóma de México, 04510 Ciudad de México, Mexico*

Urban areas serve as a melting pot of people across various socioeconomic backgrounds, who may not only live in segregation from each other but follow characteristically different mobility patterns in the city. Individual mobility is dictated by one's meaningful places, while the choice of venues to visit is not random but strongly depends on the socioeconomic status (SES) of a person. The interaction between people and places they visit, given their attributes and homophily leaning, may bring particular insight to socioeconomic stratification patterns and urban segregation at large [1]. In this data-driven study we investigate biased mobility mixing patterns in the 20 largest urban areas in the United States using a large mobility dataset collected from Foursquare [2] between 2012 and 2013. Using this mobility data, constituting $\sim$ 40K users with $\sim$ 2.5M check-ins, we infer the home location of $\sim$ 30K users. After identifying the census tract corresponding to home location, we assign a SES indicator to people using the 2010 American Community Survey (ACS) 5-year estimates from 2010-2014. We also assign a SES indicator to each place people visit, based on the location of these venues. As a SES indicator we choose per capita income, selected from out of 78 socioeconomic variables after evaluating its performance in capturing inequality based on the implementation of various feature extraction (PCA and Gini coefficient) and feature selection (information gain and decision tree) techniques. Furthermore, we define socioeconomic status as a discretization of income distribution into 10 equally populated percentiles.

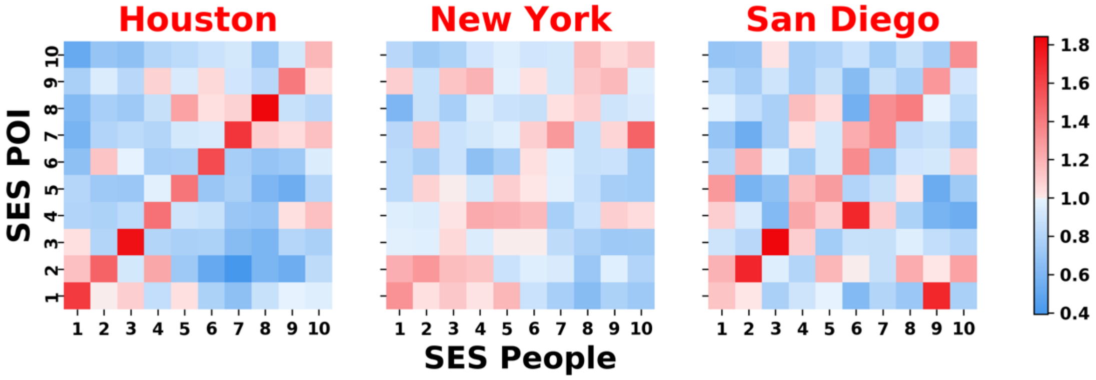


Figure 1: Normalized Interaction Matrix measures between user-venue interaction patterns (a) Houston, (b) New York, (c) and San Diego. While x-axes indicate the SES of people, y-axis indicates the SES of venue visited. Homophilic patterns of mixing are evident from the diagonal components. Blue grade bins are found relatively more frequent in the upper diagonal elements. It captures the preliminary observation of upward biased mobility which is higher than expected.

We define a bipartite network connecting individuals to places they visit. This network can be represented as an aggregated connection matrix Mi, j stratified by the SES of people and places. This matrix gives the number of times people of a given SES class i visited places of SES status j. It not only codes the distribution of SES of places people prefer to visit, but also the mixing patterns of people of different SES status. We define a reference network model where we shuffle the links of people among places visited by others from their own SES class, and measure the average individual visits between the original and random expected number of visits of places from different classes. This Normalized Interaction Matrix, summarised in the matrices shown in Fig. 1 for Houston, New York, and San Diego, shows strong signs of stratification indicating that people visit most likely places of their own socioeconomic class. At the same time, interestingly they indicate patterns of upward bias, signalling preferences of people to visit places with higher SES class than their own. We investigate this upward bias further and find that this effect is present in most cities even after we account for boundary effects of the lowest and highest SES classes. Our results indicate even stronger segregation in individual mobility what one would expect from system level distributions, providing further understanding on uneven mobility mixing patterns in cities.

Acknowledgements (optional)

G.I. acknowledges partial support from the Air Force Office of Scientific Research under award number FA8655-20-1- 7020, and by the EU H2020 ICT48 project Humane AI Net under contract 952026.

References (optional)

[1] Atlas of Inequality https://inequality.media.mit.edu (date of access 24 January 2021).

[2] D. Yang, D. Zhang, B. Qu. Participatory Cultural Mapping Based on Collective Behavior Data in Location Based Social Networks. ACM Trans. on Intelligent Systems and Technology (TIST), 2015.

[3] M. C. Gonzalez, C. A. Hidalgo, and A-L. Barabasi. Understanding individual human mobility patterns. nature, 453(7196):779–782, 2008.

[4] C. Song, T. Koren, P. Wang, and A-L. Barabasi. Modelling the scaling properties of human mobility. Nature physics, 6(10):818–823, 2010.

[5] J. Tang, F. Liu, Y. Wang, and H. Wang. Uncovering urban human mobility from large scale taxi gps data. Physica A: Statistical Mechanics and its Applications, 438:140–153, 2015.

[6] R. Gallotti, A. Bazzani, S. Rambaldi, and M. Barthelemy. A stochastic model of randomly accelerated walkers for human mobility. Nature communications, 7(1):1–7, 2016.

[7] L. Alessandretti, P. Sapiezynski, V. Sekara, S. Lehmann, and A. Baronchelli. Evidence for a conserved quantity in human mobility. Nature human behaviour, 2(7):485–491, 2018.

[8] B. Hawelka, I. Sitko, E. Beinat, S. Sobolevsky, P. Kazakopoulos, and C. Ratti. Geo-located twitter as proxy for global mobility patterns. Cartography and Geographic Information Science, 41(3):260–271, 2014.

[9] L. Wu, Y. Zhi, Z. Sui, and Y. Liu. Intra-urban human mobility and activity transition: Evidence from social media check-in data. PloS one, 9(5):e97010, 2014.

[10] R. Jurdak, K. Zhao, J. Liu, M. AbouJaoude, M. Cameron, and D. Newth. Understanding human mobility from twitter. PloS one, 10(7):e0131469, 2015.

[11] D. Brockmann, L. Hufnagel, and T. Geisel. The scaling laws of human travel. Nature, 439(7075):462–465, 2006.

[12] A. Baronchelli and F. Radicchi. Levy flights in human behavior and cognition. Chaos, Solitons & Fractals, 56:101–105, 2013.

[13] X-W. Wang, X-P. Han, and B-H. Wang. Correlations and scaling laws in human mobility. PloS one, 9(1):e84954, 2014.

[14] L. Alessandretti, U. Aslak, and S. Lehmann. The scales of human mobility. Nature, 587(7834):402–407, 2020.

[15] Mariano G Beiró André Panisson, Michele Tizzoni, and Ciro Cattuto. Predicting human mobility through the assimilation of social media traces into mobility models. EPJ Data Science, 5:1–15, 2016.

[16] C. Comito. Human mobility prediction through twitter. Procedia computer science, 134:129–136, 2018.

[17] F. Luo, G. Cao, K. Mulligan, and X. Li. Explore spatiotemporal and demographic characteristics of human mobility via twitter: A case study of chicago. Applied Geography, 70:11–25, 2016.

[18] Y. Leo, E. Fleury, J. Ignacio Alvarez-Hamelin, C. Sarraute, and Márton Karsai. Socioeconomic correlations and stratification in social-communication networks. Journal of The Royal Society Interface, 13(125):20160598, 2016.

[19] X. Dong, A. J. Morales, E. Jahani, E. Moro, B. Lepri, B. Bozkaya, C. Sarraute, Y. Bar-Yam, and A. Pentland. Segregated interactions in urban and online space. EPJ Data Science, 9(1):20, 2020.

[20] S. A. Marston. The social construction of scale. Progress in human geography, 24(2):219–242, 2000.

[21] A.Paasi. Place and region: looking through the prism of scale. Progress in human geography, 28(4):536–546, 2004.

[22] W. R. Boterman and S. Musterd. Cocooning urban life: Exposure to diversity in neighbourhoods, workplaces and transport. Cities, 59:139–147, 2016.

[23] N. Bora, Y-H. Chang, and R. Maheswaran. Mobility patterns and user dynamics in racially segregated geographies of us cities. In International Conference on Social Computing, Behavioral-Cultural Modeling, and Prediction, pages 11–18. Springer, 2014.

[24] N. M. Yip, R. Forrest, and S. Xian. Exploring segregation and mobilities: application of an activity tracking app on mobile phone. Cities, 59:156–163, 2016.

[25] Q. Wang, N. E. Phillips, M. L. Small, and R. J. Sampson. Urban mobility and neighborhood isolation in america's 50 largest cities. Proceedings of the National Academy of Sciences, 115(30):7735–7740, 2018.

[26] A. J. Morales, X. Dong, Y. Bar-Yam, and A. S. Pentland. Segregation and polarization in urban areas. Royal Society Open Science, 6(10):190573, 2019.

[27] S. Schönfelder and K. W. Axhausen. Activity spaces: measures of social exclusion? Transport policy, 10(4):273–286, 2003.

[28] D. WS Wong and S-L. Shaw. Measuring segregation: An activity space approach. Journal of geographical systems, 13(2):127–145, 2011.

[29] S. Farber, T. Neutens, H. J. Miller, and X. Li. The social interaction potential of metropolitan regions: A time-geographic measurement approach using joint accessibility. Annals of the Association of American Geographers, 103(3):483–504, 2013.

[30] S. Farber, M. O'Kelly, H. J. Miller, and T. Neutens. Measuring segregation using patterns of daily travel behavior: A social interaction based model of exposure. Journal of Transport Geography, 49:26–38, 2015.

Capturing Neighborhood Feel Using Local Graph Characteristics in Large Road Networks

A. Bramson[123]*. (1) GA Technologies Inc., Roppongi 3-2-1, Minato-ku, Tokyo, Japan, a_bramson@ga-tech.co.jp. (2) RIKEN Center for Biosystems Dynamics Research, 6-7-3 Minatojima-Minamimachi, Chuo-ku, Kobe, Japan (3) Ghent University, Tweekerkenstraat 2, Ghent, Belgium*

The network analysis of transportation networks combines graph theoretical and geographical measures to reveal features such as traffic flows, efficient routing, robustness, structural patterns, among others. Here we use a variety of road network characteristics to classify neighborhoods and capture their ambiance; i.e., discover neighborhood types and identify areas that provide a similar neighborhood feel. Because the network for our test region (the Tokyo Metropolitan Area) has millions of nodes and edges, it is too large to use traditional transportation network measures such as closeness, betweenness, or straightness centralities. Instead, we propose a set of local measures applied to a subnetwork grid that can capture salient features of the network in linear time: total length of road segments, mean segment length, number of intersections, and the straightness of road segments. We also examine the relationship of these measures to the distribution of road types (residential, tertiary, motorway, etc.) and to population, land use, land area covered by structures, and building heights. When combined with neighborhood store profiles and measures of accessibility/walkability these measures can form a crucial and missing component of livability and walkability studies with applications to real estate matching and property valuation.

Figure 1: Measures of the a) number of intersections, b) total road length, c) mean road length on subnetwork grids reveal consistent patterns in the road structure that create varying neighborhood ambiances (selected area of central Tokyo shown).

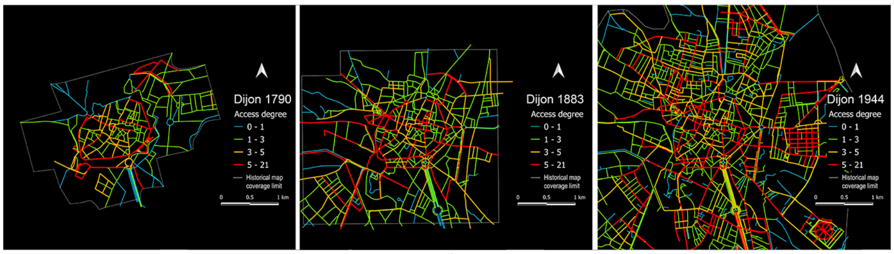
Road Network Analysis Based on Geohistorical Data

H. El Gouj[1]*, C. Lagesse*[1]*. (1) ThéMA UMR 6049, Burgundy Franche-Comté University, 32 rue Mégevand, Besançon, France, hanae.el_gouj@univ-fcomte.fr.*

Road network construction results from a subtle balance between geographical coverage and rapid access to the strategic points of space. The understanding of the road network structure is fundamental to evaluate and improve territorial accessibility.

To better understand the progressive structuring of territorial patterns, this work proposes an analysis of road network evolution. Studying actual road network morphogenesis requires geohistorical data [3], which was provided by historical maps within this framework. Those maps allow us to digitize ancient road networks. From those networks, a hypergraph is built, where road segments are combined into "ways" [2], based on a methodology to reconstruct the road's continuity. Then we compute indicators on those ways, based on topological and geometrical aspects [4]. A morphological analysis of road patterns over time is then developed on three cities of the Burgundy Franche-Comte region in France (Dijon, Besançon and Pontarlier), over three historical periods (the 18th, 19th and 20th centuries). Thus, it is possible to compare their topological features and a different set of characteristics, especially about their centralities; to help revealing information about their evolution through time.

In a prospective vision, this work aims to identify mechanisms leading evolution of road networks [1,5]. Detecting indicator stability or variation over time can help identifying similar behavior, despite geographic and cultural distances. Studying road network morphogenesis can be of a major support for better understanding their impact on accessibility and mobility.



(a) (b) (c)

Figure (1): Access degree calculated on Dijon's historical road networks.

Acknowledgements

This work is based on the COVADEO project (2018-2021 direction: C. Lagesse), funded by the Burgundy Franche-Comté region, within a funding program for new research team support.

References

[1] Marc Barthelemy. 2015. *Time Evolution of Road Networks. In Traffic and Granular Flow '13, Springer International Publishing, Cham, 317–337.*

[2] C. Lagesse, P. Bordin, and S. Douady. 2015. *A spatial multi-scale object to analyze road networks. Netw. Sci. 3, 1 (March 2015), 156–181.*

[3] Julien Perret. 2016. *Des données spatio-temporelles aux dynamiques urbaines. Habilitation à diriger des recherches. Université de Paris-Est.*

[4] Sergio Porta, Paolo Crucitti, and Vito Latora. 2006. *The network analysis of urban streets: a primal approach. Environ. Plan. B Plan. Des. 33, 5 (2006), 705–725.*

[5] Juste Raimbault. 2018. *Modeling the co-evolution of cities and networks. ArXiv180409430 Phys. (September 2018).*

“Towards a science of traffic networks: A unifying approach"

Lukas Ambuehl (ETH Zuerich), James Parker (UC Berkeley), Monica Menendez (NYU Abu Dhabi), Marta C. Gonzalez (UC Berkeley)

This work focuses on traffic congestion and unifies different approaches, perspectives, and fields into a single science of traffic. A century passed the age of the automobile, traffic networks have been described as engines of global growth and prosperity. However, their uninformed expansion is a leading cause of pollution and other negative externalities. The collapse of the traffic network, i.e. network-wide congestion, causes a loss in social and economic opportunities and increased carbon emissions; however, it is common across cities worldwide. In our endeavor towards greener and more livable cities, a clear understanding of traffic network congestion is necessary, to simplify science informed policy. A unified science of traffic networks facilitates the planning of cities from a social, environmental, and economic perspective.

Current models of urban traffic congestion implement either a network level or link level perspective. Both perspectives have proven useful in understanding urban traffic congestion - the network level perspective capturing the aggregated global system dynamics and the link level perspective accounting for the dynamics of individual roads. Despite the abundance of literature on both perspectives, they have not yet been unified into a single science of traffic consistently describing the dynamics of the congestion at both scales simultaneously. Traffic research combines diverse fields such as engineering, city planning, and physics. This has resulted in different fields preferring different analysis metrics which further hampers communication and research cohesion. In this work, we address these issues by connecting the two perspectives (link and network) for five major cities worldwide: Boston, Los Angeles, and San Francisco Bay Area in the United States, Rio de Janeiro in Brazil, and Lisbon in Portugal. We begin by modeling the traffic networks through state-of-the-art simulations spanning an extended morning traffic peak.

We then assess and connect different approaches and perspectives used by various fields to study traffic. In particular, we uncover a correspondence between *network level* traffic performance measures including the macroscopic fundamental diagram (MFD) as discussed by Helbing in *Eur. Phys. J. B* **70,** 229-241 (2009) and the flow over capacity measure Γ as proposed by Çolak et al. in *Nature Comm.* **7**:10793 (2016). Due to its ability to assess the network state with minimal data, the MFD has received a lot of attention by traffic engineers since it allows to parsimoniously model, manage, and optimize traffic in urban networks. On the other hand, \Gamma, as proposed by network scientists, measures traffic network efficiency in terms of vehicles-travelled kilometers, thus appealing to urban planners. Therefore, by combining these measures, we unify efforts by different fields to understand and model traffic at the network level into a single actionable framework.

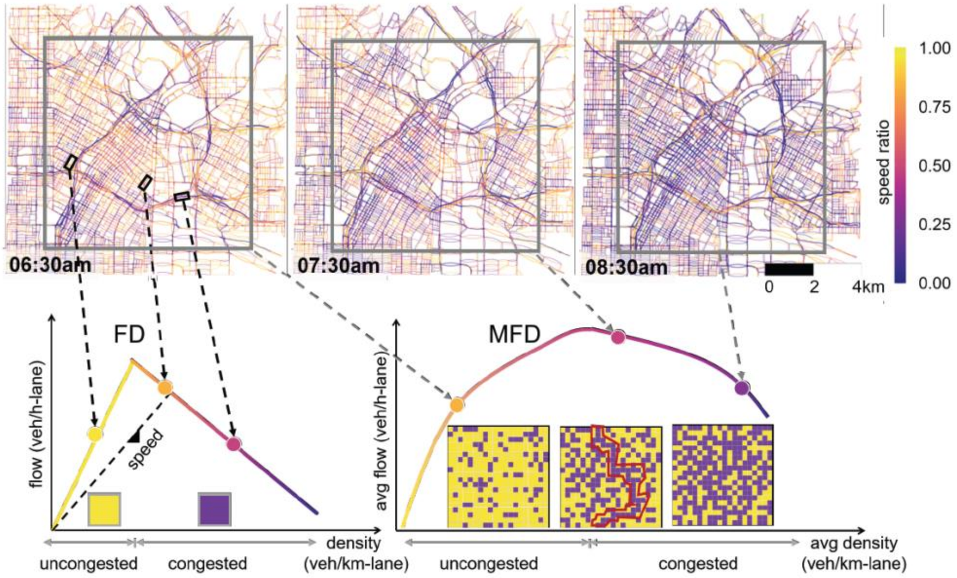


Figure 1: Examples of link and network level perspectives compared.

Building upon this unification process, we use percolation theory, inspired by Li et al. in *PNAS* **112**, 669-672 (2015), to characterize the growth of link level congestion from small isolated pockets of congestion to the sudden emergence of a single giant connected cluster of congestion. We then correlate this process with the evolution of the network level measures. Through this, for the first time, we directly connect the link level and network level perspectives as well as uncover the deeper meaning of features observed in the network level measures. Furthermore, because percolation is a tool for assessing network resilience, our results have the potential to enable planners to take advantage of the minimal data requirements of the network level measures to assess road network resilience. While traffic is typically studied with urban planning in mind, the actual impact of traffic congestion is on individual drivers. Thus, we relate our network level measures to their impact on individual drivers through the average percent congestion (equivalent to the TomTom Traffic Index). This process reveals a universal relationship which holds between cities around the world and should allow the same sparse network level data and resiliency models to be translated to the average driver.

Each of these results is interesting in its own right; when combined their implications are even more significant. The unification of perspectives, approaches and fields of study into a single *Science of Traffic* updates our understanding of urban traffic substantially. It allows traffic scientists to draw on knowledge from a coherent set of models and approaches and to communicate their results more efficiently to other researchers in the field.

Modelling the complex system of the response to a natural disaster by a multilayer network approach

C.Cherifi[1], *A.Charles*[1], *G.Bouleux*[2] *(1) DISP Université Lyon 2, 160 Boulevard de l'université, Bron, France (2) DISP, INSA Lyon, Avenue Jean Capelle, Villeurbanne, France,* **Chantal.BonnerCherifi@univ-lyon2**

The complexity of the situation during a humanitarian crisis makes its overall vision difficult. The data available in real time is often incomplete and dynamic. Indeed, collecting and updating data takes time and is not always considered as a priority. Despite their high impact, few papers studied humanitarian crisis with a complex network approach [1-3]. While all provide elements to better analyze the situation, they focus on specific, limited scopes. The complexity and the overall understanding of the situation is therefore limited. This is the limitation we focus on: providing an exhaustive view of the situation, combining various elements. We propose to model the response to a humanitarian crisis with a multi-layer network approach, as illustrated in Figure 1. The layers represent humanitarian actors and infrastructures. The intralayer links capture the relation between nodes of the same type, like two cities being connected by a road. The interlayer links represent the relation between nodes of different type, like donors providing funding or in-kind donation to field actors. The model can provide useful insights during operations, to better understand the situation, and simulate various response scenarios to choose the best way to utilize the resources available. It can also be used in the preparedness phase, to propose ways to improve the resilience of the actual network. Figure 2 illustrates an application for the earthquake in Haiti in 2010, with 240 nodes and 845 edges in the 4 layers considered.

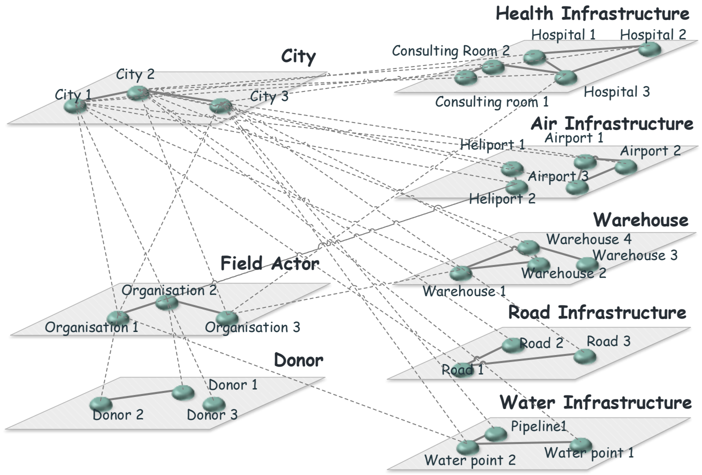


Figure 1: Multilayer network model: 8 layers of nodes, intra-layer edges and inter-layer edges

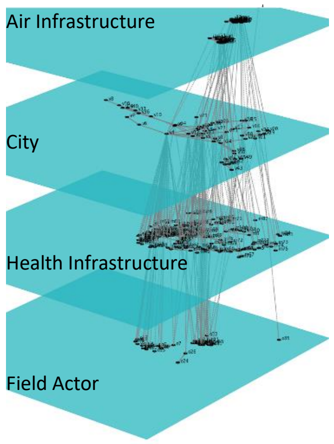


Figure 2: Multilayer network 4 layers: application to Haiti earthquake 2010

References

[1] X. Guo, N. Kapucu, Network performance assessment for collaborative disaster response, *Disaster Prevention and Management*, *24*(2) (2015) 201–220.

[2] J. A. Lassa, Post disaster governance, complexity and network theory, In *PLoS Currents*, Vol. 7 (Issue DISASTERS) (2015).

[3] M. Warnier, V. Alkema, T. Comes, B. Van de Walle, Humanitarian access, interrupted: dynamic near real-time network analytics and mapping for reaching communities in disaster-affected countries, *OR Spectrum*, *42*(3) (2020) 815–834.

Large scale transportation networks: complex behaviour induced by communication and multimodality

M.M. Khoshyaran [1], J.P. Lebacque[2],. (1) Economics Traffic Clinic (ETC) Paris. France. Email : etclinic@wanadoo.fr . (2) UGE (University Gustave Eiffel) COSYS GRETTIA-– France. Email: jean-patrick.lebacque@univ-eiffel.fr

The context of this contribution is given by regional sized transportation systems. Vehicular traffic still accounts for a majority of trips and thus provides a huge share of mass transportation. Owing to the consequences of Covid-19 the share of public transportation is not expected to increase in the close future. The current trend points towards a diversification of modes and services on the road infrastructure: this constitutes vehicular multimodality. Thus vehicular traffic will be increasingly be constituted by various subsystems competing for the same infrastructure, and benefitting from competing services in terms of information (vehicle to vehicle, web based or crowd sourcing) and connectivity. From the individual traveler's point of view, the most important decisions are departure time and path choice (i.e. traffic assignment). These decisions are taken individually and modified in real time by the available information, but they affect collectively the dynamics and efficiency of the whole transportation system. The collective behavior of travelers can be described as a mean field game [1]. Even such a simple process as departure time choice exhibits instability [4] or periodic behavior, as can be shown based on the relatively simple model [1]. Indeed links or time slots which attract too many travelers will become unattractive, which accounts for oscillations and instabilities. These instabilities are enhanced when the traveler decisions are based on real-time instantaneous information pertaining to travel times and traffic state evolution in the network [3]. Indeed vehicles travel at finite speed whereas web- or crowd-sourcing information is long range and instantaneous [2]. If there are competing information providers, chaotic dynamics may occur. With the advent of communicating vehicles, and in a more distant future, of automated vehicles, these issues should gain in importance, since the communicating/automated vehicles will constitute another sub-system with its own information sources and treatment.

The object of the contribution is to review some aspects of complexity in large scale transportation networks, and to present some new results concerning the dynamics of such systems.

References

[1] Ameli, M., Faradonbeh, M. S. S., Lebacque, J. P., Abouee-Mehrizi, H., & Leclercq, L. (2021). Departure Time Choice Models in Congested Transportation Systems Based on Mean Field Games. arXiv preprint arXiv:2101.02238. (submitted)

[2] Codeca, L., Frank, R., Engel, T., 2014. Traffic routing in urban environments: The impact of partial information, in: Proceedings of 17th IEEE International Conference on Intelligent Transportation Systems, ITSC 2014. pp. 2511 – 2516

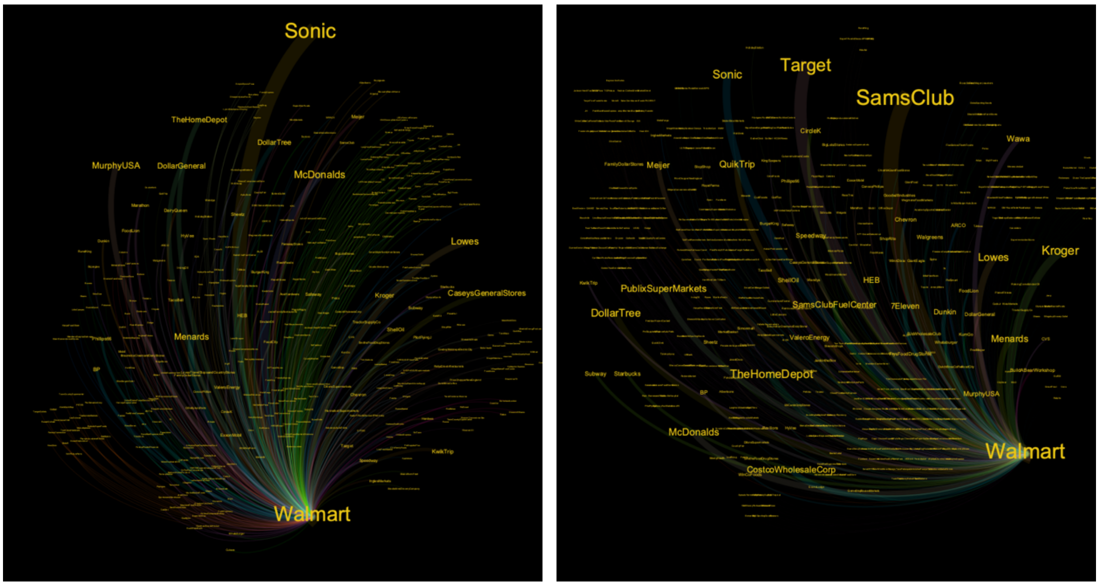
[3] Khoshyaran, M. M., & Lebacque, J. P. (2020). Reactive dynamic traffic assignment: impact of information. Transportation Research Procedia, 47, 59-66.

[4] Iryo, T. (2019). Instability of departure time choice problem: A case with replicator dynamics. Transportation Research Part B: Methodological, 126, 353-364.

Similarity in Place-based Networks Through a Rural-Urban Perspective

***Gautam Thakur**, Oak Ridge National Laboratory, Oak Ridge, Tennessee, thakurg@ornl.gov*

Recent studies in platial geography[1] have brought renewed interest in understanding the collective sense of a Place. Several definitions for a Place have been proposed in the past, a meaningful segment of geographical space, a large settlement, well-known public spaces, among others. A much agreed-upon definition is that Places are locations with meaning and they are not fixed in space or time[2]. Thus, urban morphologies and the collective sense of human settlements can be transpired by modeling the agglomeration dynamics of places. Under a bigger umbrella of this modeling, this work studies the similarity of places through human mobility movements among them. We model rural and urban areas in the United States demonstrate starkly different but consistent similarity patterns at varying spatio-temporal scale.



a) Rural region *b) Urban Region*

Figure 1: Urban and rural decomposition as examined by human mobility patterns.

The figure shows the collective similarity of places through the human mobility movement data. a) shows cluster of places that are similar in a rural area, b) shows the similarity among places in an urban area. The edges representing the frequency of the visitation patterns. There is a difference in the composition of mobility patterns when visiting places. The subset of places visited in rural are smaller than in urban areas; meaning in rural regions population is frequented among a small number of places more often; while in urban regions, the choices of abundant and hence less predictable. We study patterns' characteristics both spatially and temporality over varied duration and seasonality. This allowed modeling the stability and predictability of human mobility patterns over a longer period. Also, we examine the universality and stability of these similarities through the functional relationship between them that scales with each other over a significant interval (existence of scaling laws).

References

[1] "Platial - Wikipedia." https://en.wikipedia.org/wiki/Platial (accessed Mar. 18, 2021).

[2] T. Cresswell, "Place: encountering geography as philosophy," Geography, vol. 93, no. 3, pp. 132–139, Mar. 2008,

Impact of Urban Structure on Infectious Disease Spreading

J. Aguilar[1], *A. Bassolas*[2], *G. Ghoshal*[3,4],*S. Hazarie*[3,4],*S. Hazarie*[3], *A. Kirkley*[5], *M. Mazzoli*[1] ,*S. Meloni*[1],*S. Mimar*[3],*V. Nicosia*[2],*J. J. Ramasco*[1],*and A. Sadilek*[6] .

(1) IInstituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), 07122 Palma de Mallorca, Spain javieraguilar@ifisc.uib-csic.es.(2) School of Mathematical Sciences, Queen Mary University of London, E1 4NS, London, UK (3) Department of Physics & Astronomy, University of Rochester, Rochester, NY, 14627, USA. (4) Department of Computer Science, University of Rochester, Rochester, NY, 14627, USA. (5) Department of Physics, University of Michigan, Ann Arbor, MI, 48109, USA. (6) Google, 1600 Amphitheatre Parkway, Mountain View, CA, 94043, USA.

The ongoing SARS-CoV-2 pandemic evidences that promptly detecting which areas and regions are most susceptible to reinfections is critical to implement adequate containment policies and prevent new waves by potentially deadlier strains. Mobility has always been key to viral spreading, but the relatively slow pace of vaccination in the majority of countries, together with the appearance new aggressive variants, have forced mitigation measures to rely primarily on non-pharmaceutical interventions, such as lockdowns. This puts urban centers in the focus of epidemic surveillance and intervention. Here we show that the organization of urban flows has a tremendous impact on disease spreading and on the amenability of different mitigation strategies. By studying anonymous and aggregated intra-urban flows in a variety of cities in the United States and other countries, and a combination of empirical analysis and analytical methods, we demonstrate that the response of cities to epidemic spreading can be roughly classified in two major types, according to the overall organization of those flows. Centralized cities, where flows are concentrated primarily between mobility hotspots, are particularly vulnerable to the rapid spread of epidemics. Nevertheless, mobility restrictions in such types of cities are very effective in mitigating the spread of a virus. Conversely, in sprawled cities which present many centers of activity, the spread of an epidemic is much slower, but the response to mobility restrictions is much weaker and less effective.

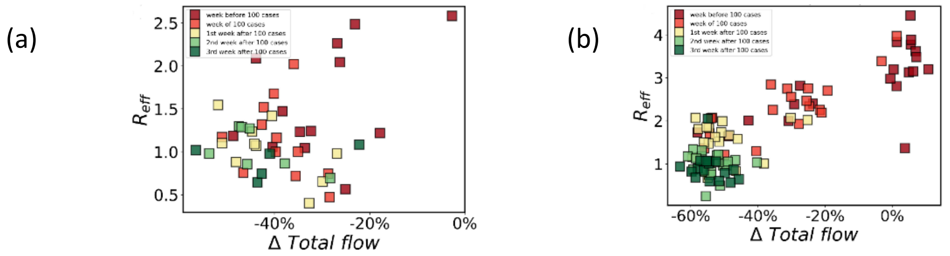


Figure 1: The temporal evolution of the effective reproduction number before and after lockdown versus the mobility change one week before Reff is measured. While sprawled cities like Atlanta (b) have regions responding independently, in centralized New York City (a), we see a clear synchronized and monotonically decreasing reduction in Reff as a function of mobility reduction.

References

[1] Aguilar, Javier, et al. "Impact of urban structure on COVID-19 spread." *arXiv preprint arXiv:2007.15367* (2020).

Morphogenesis, Evolution and Co-evolution of Cities

J. Raimbault[1,2] *(1) CASA, University College London, UK, j.raimbault@ucl.ac.uk. (2) UPS CNRS 3611 ISCPIF, France.*

Urban systems have their own complexity and specific characteristics, but still have been described and modelled using analogies and models imported from other disciplines. In particular, biological metaphors have been widely used in urban planning and design. We argue that fruitful transfer of concepts between biology and urban science can be achieved within a complexity interdisciplinary framework. We illustrate this idea by synthesising a recent stream of research focusing on urban morphogenesis, urban evolution and urban co-evolution. First, urban morphogenesis can be understood as the growth of urban form, at the scale of a city or an urban area. [1] compares and calibrates morphogenesis models at the district scale. [2] shows for urban areas in Europe that reaction-diffusion models can accurately reproduce existing urban forms. Then, the concept of urban evolution can be related to cultural evolution in urban systems, but also to cities themselves as evolving entities. [3] proposed an *evolutionary theory of cities* to understand the dynamics of urban systems, in the sense of adaptive complex systems. In that context, a model of urban evolution is introduced by [4] in which the main ingredients are innovations propagating between cities. Finally, [5] introduces a multi-level definition for a concept of urban co-evolution, with at the intermediate level the possibility of *statistically co-evolving population of urban entities within territorial niches*. A method to characterise it on spatio-temporal data was introduced by [6]. At the mesoscopic scale of urban areas, the morphogenesis model of [7] effectively captures such a co-evolution between indicators of urban morphology and the topology of road networks. At the macroscopic scale of the system of cities, [8] study a co-evolution model between cities and inter-city networks, and show that diverse regimes of co-evolution can be produced. We suggest that other several interdisciplinary bridges could be relevant to urban science. For example, biomimicry has been put forward as an effective tool for urban design and architecture, while the concept of autopoiesis has not been applied yet to urban systems. Urban computing and collective intelligence, as research fields such as smart cities and digital twins are emerging, may also be linked to studies of collective computation in biological systems. Altogether, understanding urban systems as complex systems implies an interdisciplinary integration of concepts, methods and models.

References

[1] Raimbault, J., & Perret, J. (2019). Generating urban morphologies at large scales. In *Artificial Life Conference Proceedings* (pp. 179-186).

[2] Raimbault, J. (2018). Calibration of a density-based model of urban morphogenesis. *PloS one*, *13*(9), e0203516.

[3] Pumain, D. (1997). Pour une théorie évolutive des villes. *L'Espace géographique*.

[4] Raimbault, J. (2020,). A model of urban evolution based on innovation diffusion. In *Artificial Life Conference Proceedings* (pp. 500-508).

[5] Raimbault, J. (2018). *Caractérisation et modélisation de la co-évolution des réseaux de transport et des territoires* (Doctoral dissertation, Université Paris 7 Denis Diderot).

[6] Raimbault, J. (2017). Identification de causalités dans des données spatio-temporelles. In *Spatial Analysis and GEOmatics 2017*.

[7] Raimbault, J. (2019). An urban morphogenesis model capturing interactions between networks and territories. In *The mathematics of urban morphology* (pp. 383-409).

[8] Raimbault, J. (2021). Modeling the co-evolution of cities and networks. In Handbook of Cities and Networks, in press.

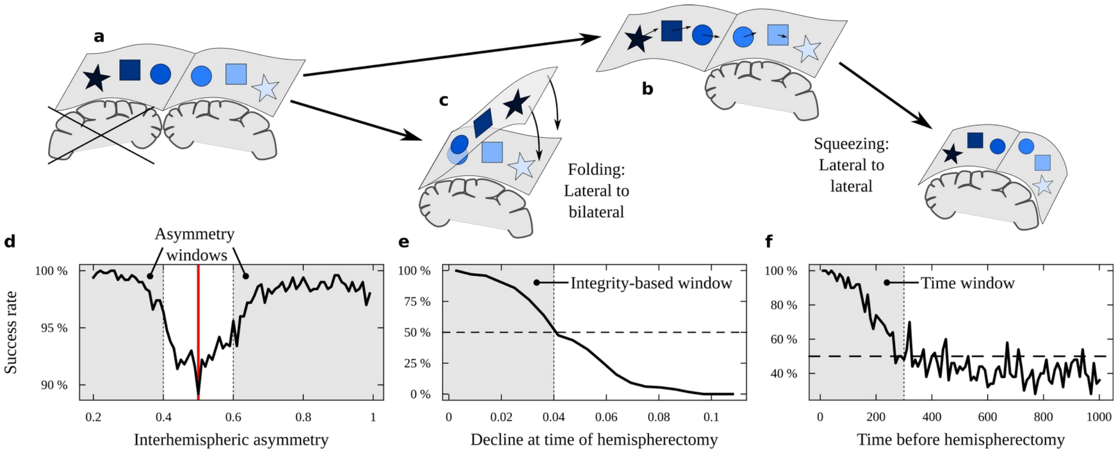
5 Biological and (Bio)Medical Complexity

Contents

Modeling Brain Reorganization after Hemispherectomy

L.F. Seoane[1,2,3]*, R. Solé*[4,5,6]*. (1) Departamento de Biología de Sistemas, Centro Nacional de Biotecnología (CSIC), C/ Darwin 3, Madrid, Spain. brigan@gmail.com. (2) Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain. (3) Instituto de Física Interdisciplinar y Sistemas Complejos, IFISC (CSIC-UIB), Campus UIB, 07122, Palma de Mallorca, Spain. (4) ICREA-Complex Systems Lab, Universitat Pompeu Fabra (GRIB), Dr Aiguader 80, 08003 Barcelona, Spain. (5) Institut de Biologia Evolutiva, CSIC-UPF, Pg Maritim de la Barceloneta 37, 08003 Barcelona, Spain. (6) Santa Fe Institute, 1399 Hyde Park Road, Santa Fe NM 87501, USA.*

Hemispherectomy is a last-resource treatment for some neurological disorders. This radical intervention allows some patients to live normally, with better odds the earlier in life it happens. Somehow, the remaining hemisphere takes on the outstanding computational burden. Brain plasticity at smaller scales shows how functionality is adopted by adjacent tissue. In models of brain rewiring after stroke, circuits accepting new workload are close and similar to the damaged ones. Hemispherectomy demands more drastic changes, mixing far and functionally diverse regions. We lack mathematical models of this. We introduce a simple model of brain reorganization after hemispherectomy [1] based on Self-Organized Maps (SOMs). We show how emerging representations in SOMs constrain brain reorganization after simulated hemispherectomy, resulting in some forbidden and some other favored rearrangement pathways, each with distinct symmetries and properties (Figure 1**a**-**c**). We discuss what the enabled paths imply for the recovery of topographic maps and language functionality after hemispherectomy. We find how too much symmetry can be detrimental for the proper formation of representation systems (Figure 1**d**). We also obtain results regarding the existence of window periods – a critical age after which hemispherectomy causes irreversible function loss (Figure 1**e**-**f**). These findings illuminate various (hitherto unexplained) clinical facts about window periods for language recovery. Our model offers a powerful thinking tool and suggests simple guiding principles for large-scale brain plasticity – notably, that the geometry of emerging representations turn into topological constraints for large-scale brain rearrangement. This offers insights about why such an aggressive intervention results in highly functional brains nevertheless, and suggests specific treatments for simulated, pathological disorders observed in our SOM models.



References:

[1] L.F. Seoane, R. Solé. Modeling brain reorganization after hemispherectomy. https://www.biorxiv.org/content/10.1101/2020.12.25.424412v1 (2020).

Diet influences both cerebral and gut microbiota networks connectivity in early ages, a case of study.

Elvia Ramírez-Carrillo[1*], Isaac G-Santoyo[1*], Oliver López-Corona[2,3], Olga A. Rojas-Ramos[1,4] and Javier Nieto[5].

[1]NeuroEcology Lab, Faculty of Psychology, Facultad de Psicología, UNAM, México, 04510
[2]Cátedras CONACyT, Comisión Nacional para el Conocimiento y Uso de la Biodiversidad (CONABIO), CDMX, México
[3]Centro de Ciencias de la Complejidad (C3), Universidad Nacional Autónoma de México, CDMX, México
[4]Coordinación de Psciobiología y Neurociencias, Facultad de Psicología, UNAM, México, 04510
[5]Laboratorio de Aprendizaje y Adaptación, Facultad de Psicología, UNAM, México, 04510

The gut-brain axis is a complex communication network that integrates physiological mechanisms connecting endocrine, immunological, dietary, and neuronal signals between the gut system and the brain, and is essential for maintaining physiological homeostasis; intervening in several neurological diseases, in mood, in motivation, and other higher cognitive functions. It is well established now that diet plays a crucial role in both the ecosystem dynamics of gut microbiota (GM) and brain functioning. Hence, even temporary changes in diet habits may cause essential effects on the gut-brain axis functioning, especially at an early age when both systems go through maturation processes. Here we use a novel combination of Machine Learning and Network Theory techniques to study the effect of protein and lipids intake in the gut microbiota and cerebral network connectivity from 5-10-year-old children from an isolated indigenous population. This study case is of the outermost interest because this population is one of the most contrasting groups in terms of lifestyle concerning typical westernized urbanized areas and, therefore, western diet. Results suggest that both network connectivity, evaluated by Minimum Spanning Tree as the critical backbone of information flow, diminish under low protein and lipid intake. So even under nonwestern diet regimens, deficient diets in the early years may produce long-term detrimental effects in the gut-brain axis due to loss of connectivity in crucial development stages. Also relevant, network theory (i.e. MST) allows to use a unifying metric of connectivity regardless of material substrc of the system (i.e. neurons vs bacteria) and assess how its being affected by the same ecological pressures, for example diet (in this case lipid and animal protein consumption).

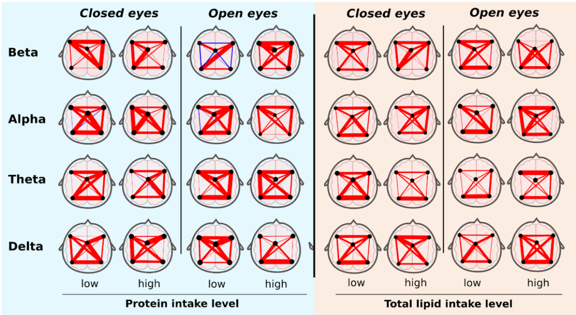


Figure 2. We show a resting-state EEG Mutual Information networks (MI), considering five regions of interest: front left, front right; posterior left, posterior right; and midline. MI was obtained for each broad brain band and resting-state condition; Open and Closed eyes. The color red means positive relationship, blue negative, just as in a standard correlation. The thickness of the line represents the intensity of the relationship.

A new framework for measuring diversity in temporal ecological networks: application to the demersal fish community in the Mediterranean sea

Z. GHALMANE[1], *B. Mérigot*[2], *R. Interdonato*[3], *D. Ienco*[4], *P. Carbonara*[5], *A. Cau*[6], *A. Esteban*[7], *M.T. Farriols*[8], *M. C. Follesa*[6], *C. Garcia-Ruiz*[9], *G. Garofalo*[10], *M. Gristina*[11], *T. Imzilen*[1], *I. Isajlović*[12], *S. Kavadas*[13], *I. Maina*[13], *P. Maiorano*[14], *C. Manfredi*[15], *J. Mifsud*[16], *P. Peristeraki*[17], *M. Sbrana*[18], *M.T. Spedicato*[5], *I. Thasitis*[19], *N. Vrgoc*[12], *P. Poncelet*[1]

The development of advanced methods able to measure the spatial and temporal variation in the diversity among communities of different species is a crucial issue in face of global change. To address this issue, many frameworks were proposed based on species proportion (e.g., abundances, biomasses) as well as on their interactions (e.g., co-occurrence) [1]. However, the existing measures of diversity need to be improved by taking into account the differences among the species (e.g., function, phylogeny, taxonomy), the topological properties of the interaction network, and the variability along the time dimension. Indeed, assemblages of species can be viewed as networks, where nodes stand for species and links stand for their interactions (e.g., co-occurrence, function). Some assemblages can have the same number of species while showing totally different topologies. A centrality measure that indicates the influence and the structural role of each node (species) can be used in order to take into account network topology. In this work, the diversity of species is computed based on their so-called 'strength-connectivity' proportions. This measure is based on a combination of species strength and connectivity. Firstly, the strength of a given species is computed as the sum of the weights of their links with other species (weight being their co-occurrences, and/or their distances such as functional ones). Then, the degree (or connectivity) of species is computed as the number of their connections. According to the type of studied network, it is also possible to replace the degree by a global centrality measure. Furthermore, assemblages of species and their diversity may change over time. In this context, a new measure of diversity, *TD*, is also introduced to quantify the temporal dynamic of species networks, that can be interpreted as the effective number of species having equal strength-connectivity proportions in all the timestamps of a given time series.

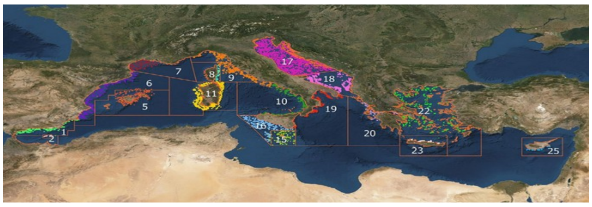


Figure 1: Studied area, northern Mediterranean sea [2].

This framework is applied to the Mediterranean exploited fish communities sampled by the MEDITS program between 1994 and 2019 [2]. Environmental and fishing conditions are also considered to assess processes behind the observed spatio-temporal patterns. *TD* show that the Geographical Sub-Areas (GSAs) having close environmental and fishing conditions have very similar temporal diversity (and vice-versa). The proposed framework encompasses both species differences (e.g., function, phylogeny, taxonomy), the topological properties of species network and its temporal variability. It allows to deepen diversity assessment, necessary to identify priority zones of interest for management and conservation.

References

[1] Ohlmann, Marc, et al. "Diversity indices for ecological networks: a unifying framework using Hill numbers." Ecology letters 22.4 (2019): 737-747.

[2] Spedicato, Maria-Teresa, et al. "The MEDITS trawl survey specifications in an ecosystem approach to fishery management". Scientia marina 83S1 (2019): 9-20.

(1) LIRMM, University of Montpellier, Montpellier, France. (2) MARBEC, University of Montpellier, CNRS, Ifremer, IRD, Sète, France. (3) CIRAD, TETIS, Montpellier, France. (4) INRAE, TETIS, Montpellier, France. (5) COISPA Tecnologia e Ricerca, Bari, Italy. (6) Department of Life and Environmental Science, Universit*y* of Cagliari, Cagliari, Italy. *(7) Instituto Español de Oceanografía (IEO), Centro Oceanográfico de Murcia, Murcia, Spain. (8) Intituto Español de Oceanografía, Centre Oceanogràfic de les Balears, Palma, Spain.* (9) *Instituto Español de Oceanografía (IEO), Centro Oceanográfico de Málaga, Fuengirola, Málaga, Spain. (10) Institute for Biological Resources and Marine Biotechnology, National Research Council (CNR-IRBIM), Messina, Italy. (11) Institute of Anthropic Impacts and Sustainability in marine environment, National Research Council (CNR), Castellammare del Golfo, Italy. (12) Institute of Oceanography and Fisheries, Split, Croatia. (13) Institute of Marine Biological Resources and Inland Waters, Hellenic Centre for Marine Research (IMBRIW/HCMR), Anavyssos, Greece. (14) Department of Biology, University of Bari Aldo Moro, Bari, Italy. (15) Laboratorio di Biologia Marina e Pesca, Dip.to BiGeA, Università di Bologna, Fano, Italy. (16) Department of Fisheries and Aquaculture, Ministry for Agriculture, Fisheries, Food and Animal Rights, Marsa, Malta. (17) Institute of Marine Biological Resources and Inland Waters, Hellenic Center for Marine Research, 71003 Heraklion, Crete, Greece. (18) Consorzio per il Centro Interuniversitario di Biologia Marina ed Ecologia Applicata, Livorno, Italy. (19) Department of Fisheries and Marine Research, Ministry of Agriculture, Natural Resources and Environment, Nicosia, Cyprus.*

Multi-levels Hepatitis E Virus Spread In Various Epidemiological Contexts: Combining Complex Network Analysis & Disease Transmission Model

P. Hammami[1]*, S. Widgren*[2]*, V. Grosbois*[3]*, A. Apolloni*[3]*, N. Rose*[1]*, M. Andraud*[1]*. (1) ANSES, French Agency for Food, Environmental and Occupational Health & Safety Epidemiology, Health and Welfare Research Unit, Ploufragan-Plouzané-Niort Laboratory, F22440 Ploufragan, France, pachka.hammami@anses.fr. (2) Department of Disease Control and Epidemiology, National Veterinary Institute (SVA), SE-751 89 Uppsala, Sweden. (3) CIRAD, Agricultural Research for Development, Univ. Montpellier, INRAE, French National Institute for Agricultural Research, Animal, Health, Territories, Risks, Ecosystems Research Unit, Montpellier, France*

Hepatitis E virus (HEV) infects pigs and humans. Its spread among pigs occurs within-farm through faecal–oral route including direct and environmental transmission [1]. Live animal movements are known as the major driver of between farms spread [2]. Strongly structured in space and time, the swine production chain is made of various interconnected farms, forming a complex network. We aimed to simulate realistic between-farm movements to feed a multilevel epidemiological model. Swine movements from 2017 to 2019 provided by the National Swine Identification database (BDporc) were analysed using network analysis (NA) methods. Exponential random graph models (ERGM) were then used to identify the key drivers of trade partner choices and the associated probabilities of contact between farms. In the 36-month period, 2.512.174 loading and unloading events were recorded involving 16.377 premises. Because of the large size of the dataset, analysis was limited to specific subsets of the data, using semestrial data and type of transported animals (breeding sows, piglets and growing pigs). For each data subset, an ERGM was selected using a stepwise forward approach based on AICs comparison and analyzing goodness-of-fit. In all cases, the company, type of farm, free-ranging characteristics, size and batch rearing systems were identified as explanatory variables of the network structure. Network statistics and levels included were subnetwork-specific. ERGM outcomes allow simulating pig trade networks with characteristics similar to the observed one (Figure 1), to prospect on the global impact of the network structure on pathogen transmission. Simulated movements fed a multi-level model developed with SimInf package [3]. First run on French data, the model will be extended to England and Germany. The simulated prevalence of HEV positive pigs sent to slaughterhouses, will feed a Quantitative microbiological risk assessment model (QMRA) to evaluate the risk of human exposure to HEV through swine products [4].

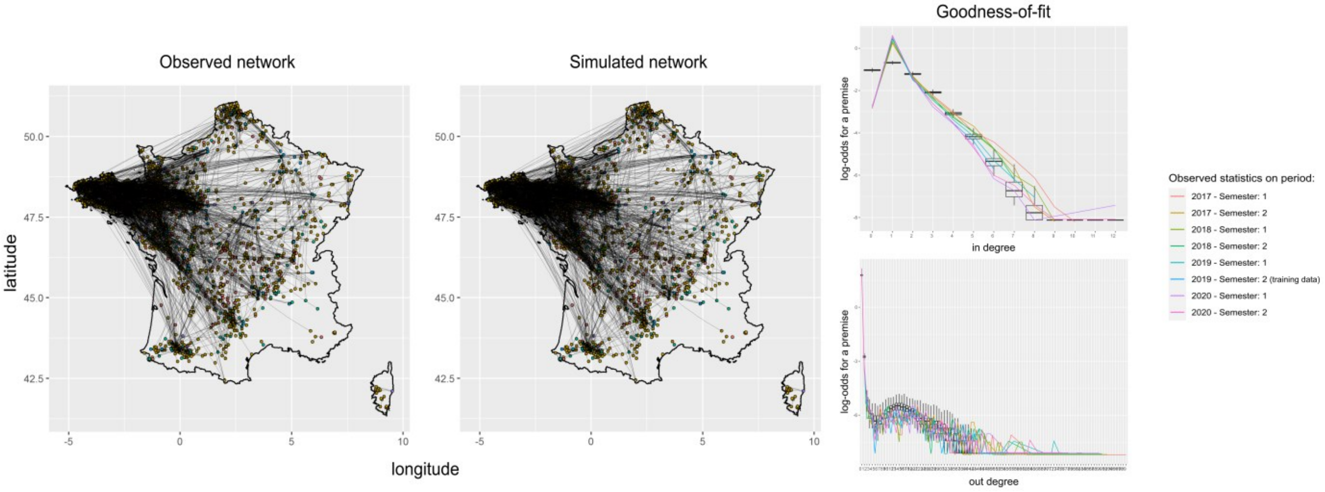


Figure 1: Breeding sows subnetwork. Left: Observed and simulated networks based on 2nd semester of 2019. Right: Goodness-of-fit and cross-validation on indegree and outdegree.

Acknowledgements

We would like to thank the comprehensive French national pig identification database (BDporc) for providing the data. This work was supported by funding from the European Union's Horizon 2020 Research and Innovation programme under grant agreement No 773830: One Health European Joint Programme.

References

[1] M. Andraud, M. Dumarest, R. Cariolet, B. Aylaj, E. Barnaud, F. Eono, N. Pavio & N. Rose. *Vet. Res.,* 44 (2013) 102.

[2] M. Salines, M. Andraud, N. Rose & S. Widgren. *PLoS ONE*, 15 (2020).

[3] S. Widgren, P. Bauer, R. Eriksson & S. Engblom. *Journal of Statistical Software*, 91 (2019) 42.

[4] E.L. Snary, A.N. Swart, R.R. Simons, A.R. Domingues, H. Vigre, E.G. Evers, T. Hald & A.A. Hill. *Risk analysis : an official publication of the Society for Risk Analysis*, 36 (2016) 437-449.

Disentangling the role of external and intrinsic dynamics on the critical signatures of neural activity

B. Mariani[1,2]*, G. Nicoletti*[2]*, M. Bisio*[3]*, M. Maschietto*[3]*, S. Vassanelli*[2,3]*, S. Suweis*[1,2] *(1) Department of Physics, University of Padova, Via Marzolo 8, Padova, Italy, benedetta.mariani@studenti.unipd.it. (2) Padova Neuroscience Center, University of Padova, Via Orus 2, Padova, Italy. (3) Department of Biomedical Sciences, University of Padova, Via Ugo Bassi 58/B, Padova, Italy*

The critical brain hypothesis has emerged as an attractive framework to understand variability in brain activity. Historically, the main signature of criticality has been the presence of scale-free neuronal avalanches and the associated "crackling-noise" relation among exponents. Yet, they are compatible with different models and phase transitions. Here we propose a modeling framework to reconcile apparent contrasting results. We develop a model of Ornstein Uhlenbeck processes, driven by a time varying stochastic modulation, that can be either non-interacting - extrinsic model - or connected through an interaction network inferred from our data - interacting model. We show that the presence of scale-free neuronal avalanches and the associated crackling-noise relation can be explained by an external, stochastic-induced mutual information in the system, whereas the intrinsic spatial correlation structure, deeply connected to the interaction network, shows the typical features of systems close to their critical point. We test our results on data obtained from the rat's cortex through state-of-the-art multi-array probes.

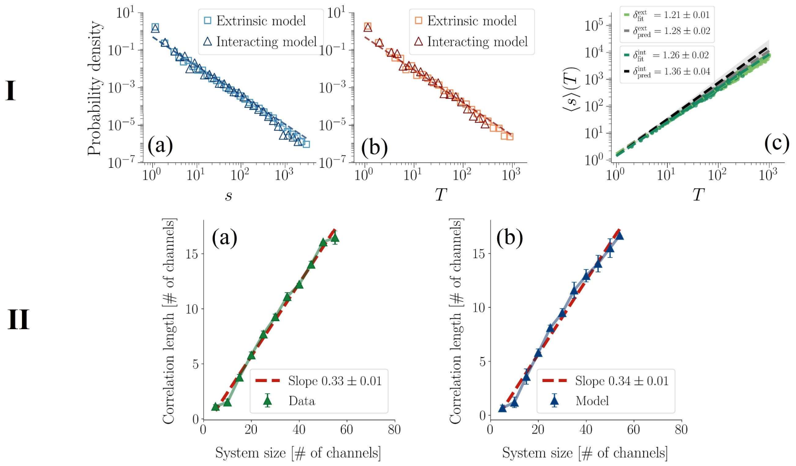


Figure 1: **Panel I**: As figures (a-b-c) show, avalanches results are very similar in our extrinsic model and in our interacting model. We show that avalanches are the sign of a high mutual information among the units, but by looking only at avalanches we could not infer nothing about the nature of the interactions. **Panel II**: The scaling of the correlation length of both our data (a) and our interacting model (b) shows features typical of critical points, while the correlations are vanishing in the extrinsic model.

References

[1] B. Mariani, G. Nicoletti, M. Bisio, M. Maschietto, S. Vassanelli and S. Suweis, "On the critical signatures of neural activity", 2021, arXIv: 2105.05070

Evolutionary escape from local fitness peaks through inversion mutations

L. Trujillo[1]*, P. Banse*[1]*, G. Beslon*[1]*. (1) Inria Beagle Team - Artificial Evolution and Computational Biology, Université de Lyon, Inria, ECL, INSA Lyon, Université Claude Bernard Lyon 1, Univ. Lumière Lyon 2, CNRS, LIRIS UMR 5205, F-69622, France, leonardo.trujillo@inria.fr.*

Adaptive molecular evolution is often conceptualised as adaptive walks on rugged fitness landscapes that are driven by changes in the genome. However, once a local fitness peak is reached there are no more selectively accessible pathways by the mechanism of point mutations. Then, the evolutionary dynamics get stuck in a suboptimal peak, hindering any possibility to reach other higher peaks. We present a simple computational model to simulate the escaping process from a local fitness peak combining point and inversion mutations as sources of genetic variation. The inversions are intended to simulate large-chromosomal rearrangements, which are one of the most frequent mutations after point mutations (for example in the Long-Term Evolution Experiment [1]). We investigate the impact of epistasis and inversion mutations on adaptive dynamics by studying the molecular evolution of asexual organisms using the NK fitness landscape model [2] for N interacting loci, where each locus interacts with K other loci. Our approach is framed within the strong selection weak mutation limit, where evolution can be simulated as random adaptive walks driven by successive mutational events constrained to incremental fitness selection. We show that inversion mutations *compactify* the topology of accessible paths in the genotype space which, together with point mutations, results in a *mutational network* with shortcuts path allowing the evolutionary process to jump-off from local fitness peaks. Our model suggests a simple mechanistic rationale to analyse the molecular evolutionary escape from local fitness peak by chromosomal rearrangements.

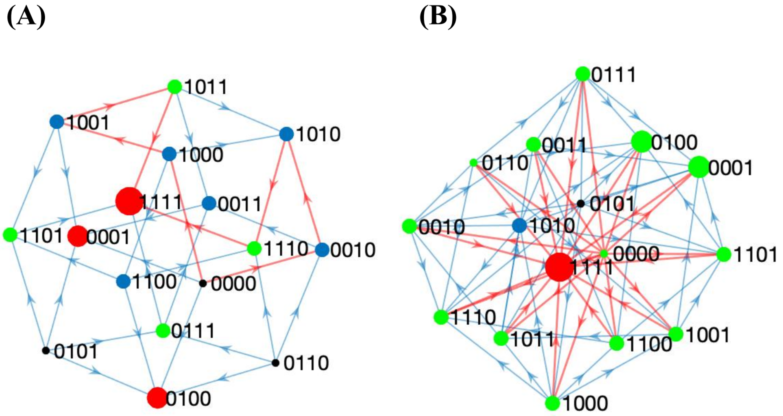


Figure 1: (A) The Levin-Kauffman fitness landscape (point mutations) [2, Fig. 1]; (B) The "reshaped" landscape (inversion). Vertices' marker sizes denote fitness ranks. Red vertices are the highest fitness ranks. Green vertices indicate the predecessors to the global fitness maximum, while the black ones are local minima. Shortest paths form the global minimum to the global maximum are highlighted in red.

References

[1] C. Raeside, et al., *MBio*, 5 (2014) 1.

[2] S. Levin, S. Kauffman, *J. Theoretical Biology*, 128, 1987, 11.

The Complexity Ratchet

***G. Beslon**[1], V. Liard[1], J. Rouzaud-Cornabas[1]. (1) Inria Beagle Team, Université de Lyon, Inria, INSA Lyon, CNRS, LIRIS UMR 5205, F-69622, France, guillaume.beslon@inria.fr.*

The evolutionary origin of biological complexity is a recurrent matter of debate. In short, two general categories of explanations are proposed in the literature. Neutralist ones state that the increase of complexity is due to random variations given that the initial complexity was close to zero. On the opposite, selectionnist ones state that it is due to selection, complex organisms being fitter than simple ones. One of the central difficulties is that this debate is mainly based on thought experiments, experimental assays lacking to test theses different hypotheses.

With this idea in mind, we used the Aevol in silico experimental evolution platform (www.aevol.fr) to design an in silico experiment able to test the selective/neutralist hypotheses [1]: We evolved populations of organisms in two different environments, one designed to enable survival of the simplest possible organisms (i.e. organisms which genomes contain only one gene), the other designed to favor complex organisms (i.e. organisms which genomes contain a virtually infinite number of genes). By repeatedly evolving organisms in these two conditions, we were able to quantify the evolution of complexity – complexity being here measured as the amount of information in the genomes and in the proteomes.

Surprisingly, in the simple environment we observed that complex organisms evolve in more than three fourths of independent replicates. Moreover, when we compared the fitnesses of the simple individuals to the fitnesses of the complex individuals, we observed that simple organisms ended up with very high fitnesses while the fitnesses of complex organisms were typically 10 to 100 times lower, showing that complexity increase is not due to direct selection. Measurements of robustness and evolvability also showed that complex organisms are neither more robust nor more evolvable than simple ones, also excluding indirect selection from triggering the increase of complexity. Finally, a comparison of organisms that evolved in simple and complex environments showed that their complexities were of the same order of magnitude. This shows that complexity can accumulate even in conditions where it is not selected or even counter-selected! However, we also observed that in the absence of selection, complexity quickly drops in all situations showing that selection is an active force in the process of complexity accumulation. All these observations show that complexity increase is due to a "complexity ratchet": each event of gene acquisition creates the potential for the acquisition of more genes, ultimately pushing evolution towards complex solutions even when the environment is simple. More precisely, the complexity ratchet is due to sign epistasis: in a simple environment a simple organism can improve its fitness through different mutational paths, some (e.g., sequence divergence) leading to simple organisms and some (e.g., gene duplication-divergence) leading to complex ones. Both paths are initially favorable but in the genetic background of complex organisms, the first one becomes highly deleterious. Hence, complex organisms cannot switch back to simplicity, even if this would be ultimately highly favorable. On the opposite, they slowly improve by accumulating new genes, each gene acquisition making the ratchet click further and pushing complexity to an increase.

References

[1] V. Liard, D.P. Parsons, J. Rouzaud-Cornabas, G. Beslon (2020) The complexity ratchet: Stronger than selection, stronger than evolvability, weaker than robustness. *Artificial life*, 26(1), 38-57.

Modeling The Emergence Of Scale-Free Fire Outbreaks In Australia

G. Nicoletti[1]*, L. Saravia*[2]*, S. Suweis*[1]*.*
(1) Laboratory of Interdisciplinary Physics, Department of Physicsand Astronomy "G. Galilei", University of Padova, Padova, Italy giorgio.nicoletti.1@phd.unipd.it*. (2) Área Biología y Bioinformática, Universidad Nacional de General Sarmiento, Buenos Aires, Argentina*

Between 2019 and 2020 Australia experienced a dramatic bushfire season, with devastating ecological and environmental consequences. In this work we analyze the monthly evolution of the burned area in Australia from 2000 to 2020, obtained via satellite imaging with high spatial resolution. We find that the 2019-2020 peak of the burned area, the highest of the historical data, is associated with signatures that are typically found in phase transition. The spatial configuration of the burned area proves to be scale-invariant, with the size of the outbreaks that are distinctively power-law distributed and a dynamical synchronization that led to the presence at once of many smaller fires together with much larger fires. We propose a simple model that takes into account the combined effect of the fires spreading and of the vegetation regrowth in a spatial fashion, and we show that the features of 2019-2020 peak can be associated to a percolation transition where system-size outbreaks appear. We argue that the trajectory in the parameter space of the model can be related to the effects of climate change, which might eventually push the system close to a second critical point predicted by the model: an absorbing phase transition after which the system will be doomed to never recover.

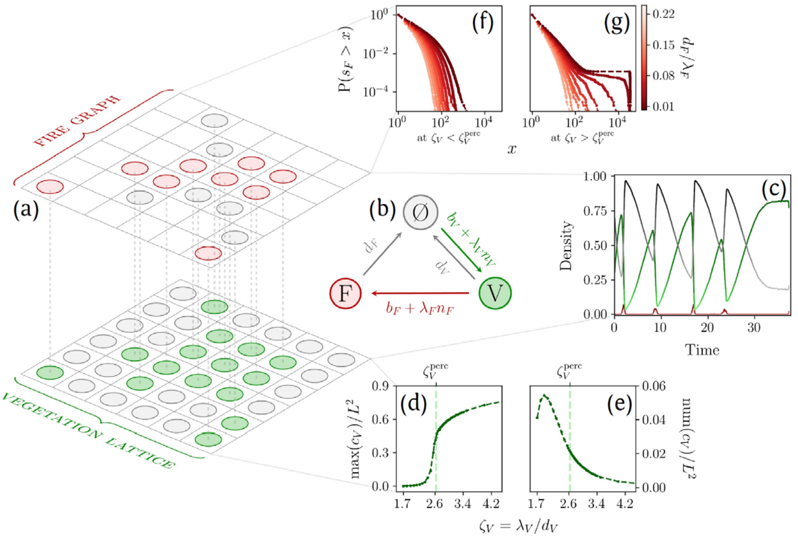


Figure 1: The fundamental properties of our model. (a) A pictorial depiction of the model as a multi-layer graph. The vegetation spreads on a 2D lattice and, on top of it, the fires spread on an arbitrary, time-varying graph that is induced by the sites occupied by the vegetation. In turn, the vegetation cannot spread in a site that is occupied by a fire. (b) The model can be cast as a single-layer model with three states and asymmetric rates. (c) On a 2D lattice the model displays a charge-discharge behavior if the vegetation dynamics is much slower than the fires one, and fires are relatively rare events. (d-e) The vegetation layer alone displays an isotropic percolation transition at ζ_V^c. (f-g) In a timescale separation scenario, below the critical threshold ζ_V^c fires are typically limited by small vegetation clusters and thus large outbreaks are not possible, no matter how effective the fire spreading is. Above ζ_V^c, on the other hand, it is the fire parameters that typically determine how large an outbreak can be, and we can have system-size outbreaks. In between these regimes lies a power-law distribution of the fire sizes.

Long- and Short-Term Effects of Cross-Immunity

I. Atienza[1,2], *L.F. Seoane*[1,2]. *(1) Departamento de Biología de Sistemas, Centro Nacional de Biotecnología (CSIC), C/ Darwin 3, Madrid, Spain. brigan@gmail.com. (2) Grupo Interdisciplinar de Sistemas Complejos (GISC), Madrid, Spain.*

Influenza strains have long afflicted humanity in seasonal waves. Similarity between successive variants can determine each year's epidemiological pattern [1]. In a much shorter time frame, we have seen SARS-CoV-2 mutants pose new threats as each new strain dominated the pandemic in different countries. While all variants are phylogenetically close, each new mutant raises the question of whether vaccines developed for an older target will still work. This issue is extended to naturally immunized individuals, whose trained immune system might be rendered ineffective. How does cross-immunity between successive variants (and loss thereof) change short-term epidemiological and large-scale evolutionary dynamics? Building on the SIR model, we show how cross-immunity deeply alters the combinations of infection and recovery rates leading to seasonal pandemics. In the short-term, our model constrains scenarios under which herd-immunity is lost after vaccination as new mutants escape cross-immunity. This provides us with a quantitative tool which (combined with empirical data about infectivity rates of each strain) allows us to quickly assess when new vaccines are needed (Figure 1**a**). At an evolutionary scale, cross-immunity strongly constrains the properties of variants capable of leading to epidemics (Figure 1**b**-**c**). This also affects the periodicity between pandemics of different magnitude. While simple, SIR equations capture the phenomenology that most matters in the long term. Alternatively, our approach can be extended to more detailed models, where similar drastic effects are expected.

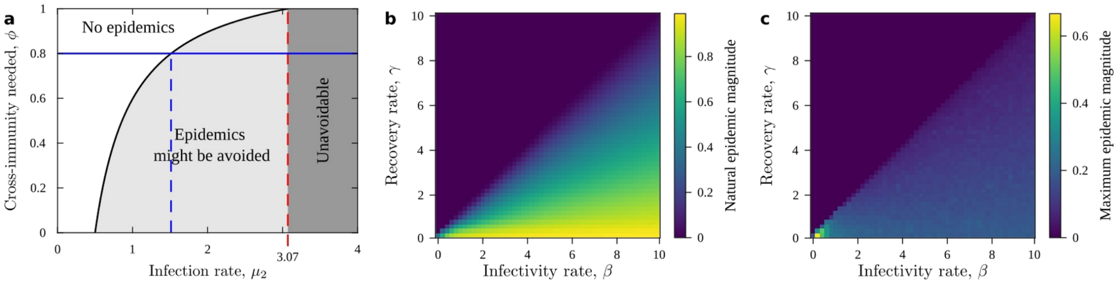


Figure 1: **a** Given a strain with infectivity rate $\beta_1 = 3.07$, the threat posed by a new mutant depends on its characteristics (new infection and recovery rates, and level of cross-immunity). A map charts when a new widespread outbreak is unavoidable or might be halted by enough cross-immunity. In the example, a new strain with infectivity rate $\beta_2 = 1.5$ (dashed vertical blue line) cannot produce widespread outbreaks if cross-immunity remains above an 80% (solid horizontal blue line). This can be assessed testing vaccine efficiency in the lab. **b** Magnitude of outbreaks according to the SIR model. Epidemics only ensue if recovery rate is slower than infection rate. **c** A population is affected by successive strains that are related to each other – hence, older epidemic processes provide some immunity in future ones. Lasting cross-immunization prevents epidemic outbreaks over extended periods of time. It also alters deeply the combinations of model parameters that can lead to large pandemic episodes.

References:

[1] K. Koelle, S. Cobey, B. Grenfell, M. Pascual. Epochal evolution shapes the phylodynamics of interpandemic influenza A. *Science*, **314**(5807), pp.1898-1903 (2006).

Constrained proteome allocation affects coexistence in models of competitive microbial communities

In this work, we reconsider one of the theoretical frameworks most commonly used to model population dynamics in competitive ecosystems, MacArthur's consumer-resource model, in light of experimental evidence showing how proteome allocation affects microbial growth. This new framework allows us to describe community dynamics at an intermediate level of complexity between classical consumer-resource models and biochemical models of microbial metabolism, accounting for temporally-varying metabolic networks subject to constraints on growth and protein synthesis in the presence of multiple resources, while preserving analytical insight into the dynamics of the system. We show that dynamic metabolic adaptation enables the community to self-organize, allowing several species to coexist even in the presence of few resources, and pointing how such adaptation could be an important mechanism for maintaining high levels of diversity even in environments with few energy sources

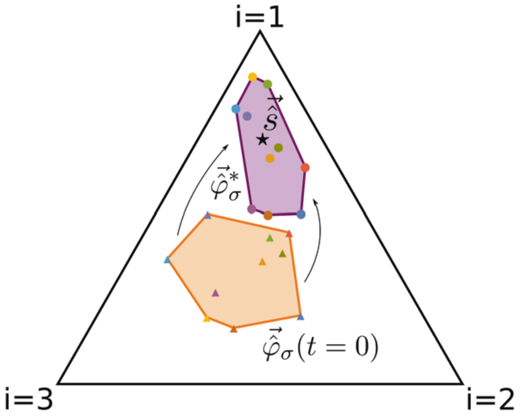


Figure. Graphical representation of the second condition necessary for coexistence. Here we consider a system with $N_S = 10$ species and $N_R = 3$ resources (for ease of representation). In this case, the system can be represented on a bi-dimensional simplex (i.e., a triangle) where each vertex corresponds to one of the available resources. On this simplex, we can draw the rescaled nutrient supply rate vector ***s*** (denoted by a star), and the rescaled initial proteome fractions ϕ allocated by the species σ to uptake and metabolism of the resources (colored triangles); their convex hull is drawn in orange. We have also drawn the stationary values ϕ^*_σ of the proteome fractions (colored circles), and their convex hull is drawn in purple. In this representation, if ***s*** lies on one on the sides of the simplex, it means that only two of the available resources are being externally supplied to the system, and analogously if one of the ϕ^*_σ lies on one of the sides of the simplex, it means that its corresponding species is uptaking and metabolizing only two of the available resources. All species coexist if the metabolic strategies adapt so that the corresponding convex hull include the supply vector **s**.

Reference Pacciani-Mori, L., Suweis, S., Maritan, A., & Giometto, A. (2021). Constrained proteome allocation affects coexistence in models of competitive microbial communities. *The ISME Journal*, *15*(5), 1458-1477.

Stationary distribution of node2vec random walk for brain connectivity networks discrimination

L. Carboni^{1,2}, *S. Achard*^{2}, *M. Dojat*^{1,2}. *(1)* Univ. Grenoble Alpes, Inserm, U1216, Grenoble Institut Neurosciences, 38000 Grenoble, Fr lucrezia.carboni@inria.fr *(2)* Univ. Grenoble Alpes, Inria, CNRS, LJK, 38000 Grenoble, Fr

The study of human brain functional connectivity under diverse conditions aims at analyzing the complexity of interactions between brain regions. A natural mathematical representation of brain connectivity is given by a graph model. From such a model, we may extract graph metrics and use them to define normal functional conditions and to detect pathological deviations. Unfortunately, there are no general graph metrics allowing to robustly detect such deviations independently on the underlying pathology. Instead of using idiosyncratic graph metrics and guided by our previous results [1], we propose a graph characterization based on the calculation of the stationary distribution (SD) of a random walk.

This random walk is proposed in the definition of a node proximity function in node2vec algorithm, it is a second order Markov-chain where equivalent nodes are visited infinity times with the same probability [2]. As previously assumed in [1], we consider, for the analysis of brain connectivity computed from resting state fMRI datasets, the role each region plays in the connectivity pattern, i.e. regions which are secondary or central nodes. Indeed, due to inter-subject variability, even under the same conditions, the individual network connection weights can differ. However, we conjecture that for individuals in the same class (e.g. healthy or pathological classes), the same regions behave similarly in the connectivity network. Thus, in this context of application the graph metric should capture the notion of structural equivalence: two nodes in a graph are said to be structurally equivalent if it exists a graph automorphism which maps the first node to the second. Moreover, detecting the structural pattern of each network may receive a neurological interpretation.

We propose to use the SD as graph descriptor to discriminate brain connectivity networks. The main advantage of our approach is the possibility of detecting equivalent nodes by the exact computation of the SD, which is only approximated when using the node2vec algorithm. Then, for graph discrimination the SD serves as input for a logistic classifier.

We applied our approach to two real clinical datasets, from coma patients (CHU Strasbourg, Fr) and Young & Elderly healthy subjects (Brain Imaging Centre Cambridge, UK), and to a synthetic dataset. For these three experiments, the classifier reached good accuracy results (see Table 1), demonstrating our approach feasibility in very different settings and supporting our assumption of the existence of a structural signature that requires future exploration.

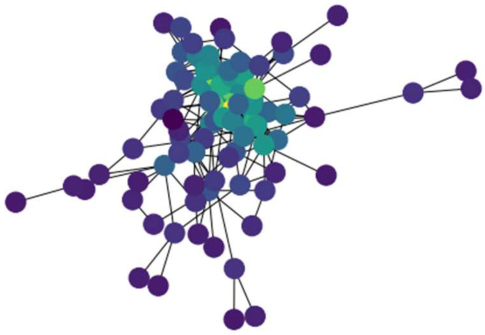


Figure 1 *(Left)Visualization of the stationary distribution on a real graph, same colors represent likely structurally equivalent nodes. Darker colors represent secondary nodes, while lighter colors represent central nodes.*

Synthetic	**Strasbourg**	**Cambridge**
0,73 ± 0,0050	0,95 ± 0,037	0,68 ± 0,016

Table 1 *Accuracy results*

Acknowledgements
LC is supported by a grant from by MIAI@Grenoble Alpes (ANR 19-P3IA-003).

References

[1] L. Carboni et al., *Network embedding for brain connectivity,* IEEE, *Proc.* ISBI 2021
[2] L. Meng and N. Masuda, *Analysis of node2vec random walks on networks,* Proc. of the Royal Society A, 2020

Identifying different risk factors for developing Diabetes mellitus in later life by using multilayer disease network

E. Dervic[1,2], *M. Leutner*[3], *A. Kautzky-Willer*[3,4,] *P. Klimek*[1,2,]. *(1) Section for Science of Complex Systems, CeMSIIS, Medical University of Vienna, Spitalgasse 23, Vienna, Austria, elma.dervic@meduniwien.ac.at (2) Complexity Science Hub Vienna, Josefstädter Straße 39, 1080 Vienna, Austria (3) Department of Internal Medicine III, Clinical Division of Endocrinology and Metabolism, Gender Medicine Unit, Medical University of Vienna, Waehringer Guertel 18–20, 1090, Vienna, Austria (4) Gender Institute, Gars am Kamp, Austria*

Multimorbidity, the co-occurrence of two or more diseases in a patient, is associated with decreased quality of life, increased need for hospitalisations, health care utilisation, mortality, and costs of care. Medical conditions, for example, comorbidities or risk factors such as arterial hypertension or cardiovascular disease, cannot be studied independently from each other, as many diseases tend to be comorbid. We aim to further our quantitative understanding of how multimorbidity develops over the life course of patients in a sex-specific way.

Here, we use a unique dataset containing 44 million records of almost all inpatient stays from 2003 to 2014 in Austria to investigate the effect of different risk factors on the development of diabetes mellitus in later life.

We developed a new, multilayer disease network approach to quantitatively analyse complex connections between two or more conditions and how they evolve over the life course of patients. Nodes represent diagnoses in specific age groups in intervals of ten years. Each layer is then a comorbidity network for one age group. Inter-layer links encode a significant correlation between diagnoses ($p < 0.001$, relative risk > 1.5), while intra-layers links present correlations between diagnoses across different age groups. Nodes of the introduced multilayer network are embedded into space by using the node2vec algorithm, which generates vector representations of nodes.

In the embedded space, we identified a cluster of risk factors for developing type two diabetes. Our results show that, for example, the occurrence of risk factors for cardiovascular disease, behavioural disorders, peripheral vascular disease, respiratory insufficiency or asthma bronchialis in the age group of 40-49 years were related to an increased risk of being firstly diagnosed with diabetes mellitus in the older ages between 50-59 years. These results point out that such cluster analyses can help to identify critical events that put patients at high risk for diabetes mellitus even decades later, according to different risk factor combinations - a new approach in personalised and individualised medicine.

The effective graph reveals redundancy, canalization, and control pathways in biochemical regulation and signaling

Alexander J. Gates[1*], *Rion Brattig Correia*[2,3], *Xuan Wang*[4], ***Luis M Rocha***[2,4,5*]. *(1) Network Science Institute, Northeastern Univ., Boston, MA, USA. (2) Instituto Gulbenkian de Ciência, Oeiras, Portugal. (3) CAPES Foundation, Ministry of Education of Brazil, Brasília, Brazil. (4) Center for Social and Biomedical Complexity, Luddy School of Informatics, Computing & Eng., Indiana Univ., Bloomington, IN, USA. (5) Dept. of Systems Science and Industrial Eng., Binghamton Univ., Binghamton, NY, USA. *a.gates@northeastern.edu, rocha@binghamton.edu*

The ability to map causal interactions underlying genetic control and cellular signaling has led to increasingly accurate models of the complex biochemical networks that regulate cellular function. These network models provide deep insights into the organization, dynamics, and function of biochemical systems: for example, by revealing genetic control pathways involved in disease. However, the traditional representation of biochemical networks as binary interaction graphs fails to accurately represent an important dynamical feature of these multivariate systems: some pathways propagate control signals much more effectively than do others. Such heterogeneity of interactions reflects canalization—the system is robust to dynamical interventions in redundant pathways but responsive to interventions in effective pathways. Here, we introduce the effective graph, a weighted graph that captures the nonlinear logical redundancy present in biochemical network regulation, signaling, and control. Using 78 experimentally validated models derived from systems biology, we demonstrate that 1) redundant pathways are prevalent in biological models of biochemical regulation, 2) the effective graph provides a probabilistic but precise characterization of multivariate dynamics in a causal graph form, and 3) the effective graph provides an accurate explanation of how dynamical perturbation and control signals, such as those induced by cancer drug therapies, propagate in biochemical pathways. Overall, our results indicate that the effective graph provides an enriched description of the structure and dynamics of networked multivariate causal interactions. We demonstrate that it improves explainability, prediction, and control of complex dynamical systems in general and biochemical regulation in particular.

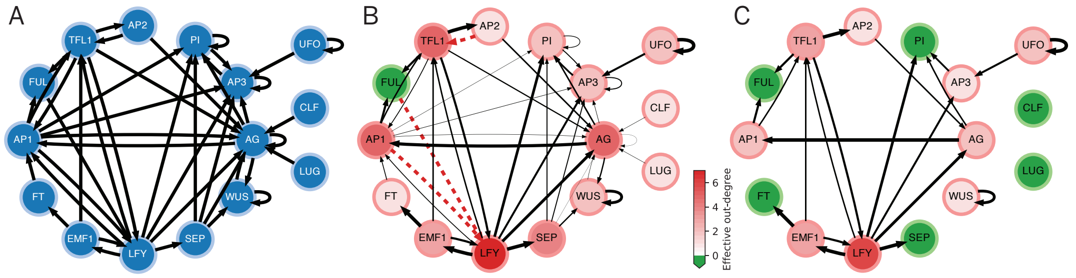


Figure 1: Study of the *A. thaliana* Boolean network model. (A) The interaction graph. (B) The effective graph. Edge thickness denotes effectiveness; dashed red indicates fully redundant edges; node color intensity denotes effective out-degree; and green nodes denote cases of null effective out-degree. (C) A threshold effective graph showing only edges with effective weight ≥ 0.4 to enhance visibility of the largest connected component that allows LFY to function as a master regulator and reveals that WUS functions simply as an autoregulator; green nodes denote cases of null effective out-degree at this threshold level.

References

[1] A.J. Gates, R.B. Correia, X. Wang, & L.M. Rocha. Proceedings of the National Academy of Sciences, 118 (12) e2022598118. doi: 10.1073/pnas.2022598118

Hubs, diversity, and synchronization in FitzHugh-Nagumo oscillator networks: Resonance effects and biophysical implications

*Stefano Scialla [1], Alessandro Loppini [1], **Marco Patriarca** [2], Els Heinsalu [2]*
(1) Department of Engineering, Università Campus Bio-Medico di Roma - Via À. del Portillo 21, 00128 Rome, Italy, (2) National Institute of Chemical Physics and Biophysics - Rävala 10, Tallinn 15042, Estonia marco.patriarca@kbfi.ee

Using the FitzHugh-Nagumo equations to represent the oscillatory electrical behavior of β-cells, we developed [1] a coupled oscillator network model with cubic lattice topology, showing that the emergence of pacemakers or hubs in the system can be viewed as a natural consequence of oscillator population diversity. The optimal hub to non hub ratio is determined by the position of the diversity-induced resonance maximum for a given set of FitzHugh-Nagumo equation parameters and is predicted by the model to be in a range that is fully consistent with experimental observations. The model also suggests that hubs in a β-cell network should have the ability to "switch on" and "off" their pacemaker function. As a consequence, their relative amount in the population can vary in order to ensure an optimal oscillatory performance of the network in response to environmental changes, such as variations of an external stimulus.

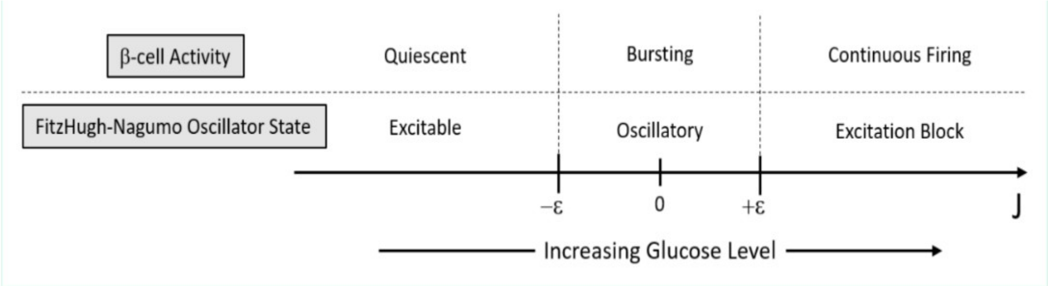


Figure 1 (from Ref. [1]): General correspondence between β-cell activity and the states of the FitzHugh-Nagumo oscillator network, used in the model studied.

Acknowledgements

A. L. acknowledges the support of Gruppo Nazionale per la Fisica Matematica (GNFM-INdAM). E. H. and M. P. acknowledge support from the Estonian Ministry of Education and Research through Institutional Research Funding IUT39-1, the Estonian Research Council through Grant PUT1356 and PRG1059, and the ERDF (European Development Research Fund) CoE (Center of Excellence) program through Grant TK133.

References

[1] Stefano Scialla, Alessandro Loppini, Marco Patriarca, Els Heinsalu. Phys. Rev. E 103, 052211 (2021). DOI: https://doi.org/10.1103/PhysRevE.103.052211

Multi population analysis of electronic health records reveal biases in the administration of known drug-drug interactions

Jon Sánchez-Valle[1*]*,* ***Rion Brattig Correia***[2*]*, Marta Camacho-Artacho, Rosalba Lepore*[1]*, Mauro M. Mattos*[3]*, Luis M. Rocha*[2,4]*, Alfonso Valencia*[1,5]*. (1) Barcelona Supercomputing Center, Barcelona, Spain. (2) Instituto Gulbenkian de Ciência, Oeiras, Portugal. (3) Univ. Regional de Blumenau, Blumenau, Brazil. (4) Dept. of Systems Science and Industrial Eng., Binghamton Univ., Binghamton, NY, USA. (5) ICREA, Barcelona, Spain. *jon.sanchez@bsc.es, *rionbr@gmail.com.*

The occurrence of drug-drug-interactions (DDI) from multiple drug dispensation is a serious problem, for patients' health and health-care systems alike, and previously highlighted by our own work [1]. Here we present preliminary results of a multi-population expanded analysis of the DDI phenomena. We characterize the age- and gender-associated biases in drug dispensation patterns from distinct health care systems, both public and private. We analyzed drug dispensations from population-wide electronic health records (EHR) in Blumenau (Brazil; pop. 330K), Catalonia (Spain; pop. 7.5M), and Indianapolis (USA; pop. 864K) with a temporal scale ranging from 1.5 to 10 years' worth of data. We investigate the role of polypharmacy in the observed DDI rates by building a statistical null model that shuffled drug labels while accounting for cohort specific drug availability. In total, 149 shared DDI were found in the three populations. The risk of such DDI as patient age was also characteristically similar in all three populations. We confirmed that in general women are at an increased risk of DDI–with the exception of males over 50 years-old in Indianapolis. Importantly, we found that the increased risk of DDI cannot be solely explained by polypharmacy or increased co-administration rates in the elderly. Finally, we simulated alternative drug regimens for patients dispensed DDI involving Omeprazole, a proton pump inhibitor with substitutes having less known DDI, to characterize the population effect of avoiding overly prescribed interactions. We found that alternatives to Omeprazole can reduce the number of patients affected by DDIs by up to 21% in Blumenau and Catalonia. Interestingly, Omeprazole is less used in Indianapolis, therefore the same simulation reduces only 2% of affected patients, which demonstrates social and/or economic factors at play in the global DDI phenomena.

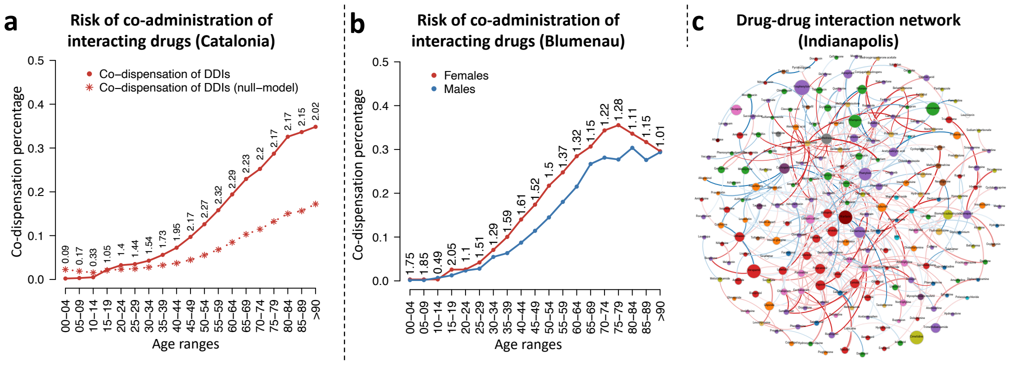


Figure 1: (**A**) Risk of DDI dispensation (circles) and the associated null model (stars) stratified by age in Catalonia. (**B**) Risk of DDI dispensation stratified by age and gender in Blumenau, where women (men) are shown in red (blue). (**C**). A network visualization of the DDI found in Indianapolis; nodes are drugs (size denote the probability of interaction; color their class); edges weights represent the DDI association strength; edge color the gender associated risk.

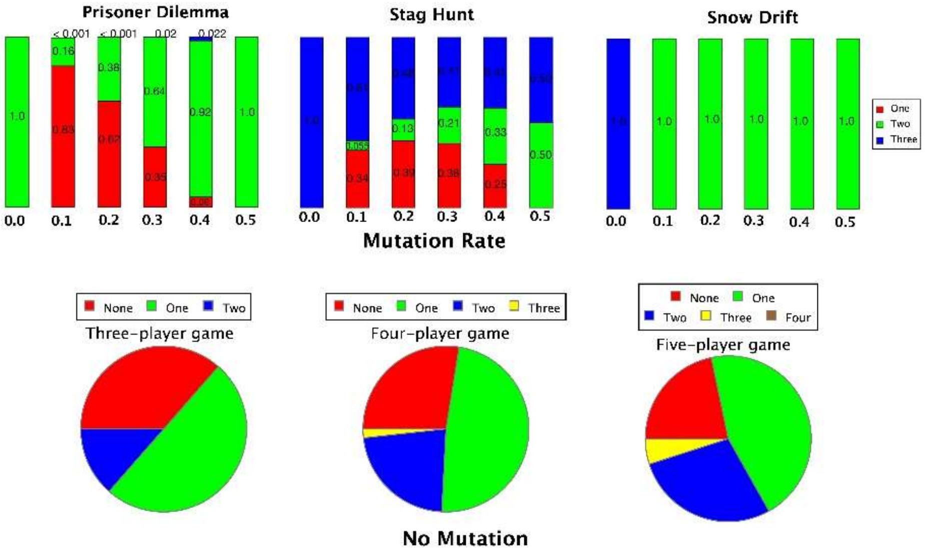
References:

[1] R.B. Correia, et al. *npj Digital Medicine*. 2 (**74**), 2019. doi: 10.1038/s41746-019-0141-x

Statistics of equilibria points in pairwise social dilemmas under replicator-mutator dynamics

M.H. Duong[1], *T. A. Han*[2] *(1) University of Birmingham, Birmingham, UK, Email: h.duong@bham.ac.uk (2) Teesside University, Middlesbrough, UK.*

The replicator-mutator equation is a set of differential equations describing the evolution of frequencies of different strategies in a population that takes into account both selection and mutation mechanisms. It is a fundamental mathematical framework for the modelling, analysis and simulation of complex biological, economical and social systems and has been utilized in the study of, just to name a few, population genetics, autocatalytic reaction networks, language evolution and the evolution of cooperation, see e.g. [1]. In this talk we discuss our recent works [2,3] on the statistics of the equilibria of the replicator-mutator equation. For two-player social dilemmas (namely the Prisoner's Dilemma, Snow Drift, Stag Hunt and Harmony), we characterize (stable) equilibrium points and analytically calculate the probability of having a certain number of equilibria when the payoff entries are uniformly distributed. For multi-player random games whose pay-offs are independently distributed according to a normal distribution, by employing techniques from random polynomial theory, we compute the expected or average number of internal equilibria. In addition, we perform extensive simulations by sampling and averaging over a large number of possible payoff matrices to compare with and illustrate analytical results. Our results provide new insights into the behavioural diversity of dynamical systems, including biological, social and artificial life ones.



(Upper row) impact of mutation rate on the probability of having a certain number of internal equilibrium points for three popular social dilemmas (two-player two-strategy games). (Lower row) probability of having a certain number of internal equilibrium points for two-strategy games with three, four and five players without mutation.

References

[1] Komarova NL (2004) Replicator-mutator equation, universality property and population dynamics of learning. J Theor Biol 230(2):227–239.

[2] Duong MH, Han TA (2020) On Equilibrium Properties of the Replicator–Mutator Equation in Deterministic and Random Games. *Dyn Games Appl* 10, 641–663

[3] Duong MH, Han TA (2021) Statistics of equilibria of social-dilemmas with mutation, In preparation.

Graph topology and habitat assortativity drive phenotypic differentiation in an eco-evolutionary model

V. Boussange[1,2], *L. Pellissier*[1,2], *(1) ETH Zürich, Universitätstrasse 16, 8092 Zürich, Switzerland, bvictor@ethz.ch.*
(2) WSL, Zürcherstrasse 111, 8903 Birmensdorf, Switzerland

Biodiversity results from differentiation mechanisms that are largely influenced by the features of the landscape over which populations are distributed. Notably, landscape connectivity and habitat heterogeneity constrain the movement and survival of individuals, thereby promoting differentiation through drift and local adaptation. Nonetheless, the complex topology of landscapes can blur our understanding on how they regulate differentiation. Here, we construct from first principles a stochastic, individual based model where the landscape is explicitly represented as a graph. We investigate both analytically and with simulations how the graph topology and the underlying spatial habitat configuration affect differentiation at the population level. Differentiation emerges from differences in local populations' phenotypes, which consist of neutral and adaptive traits that are co-evolving. Using a macroscopic description of the dynamics developed in ref. [1] we demonstrate that a simple quantity, the habitat assortativity of the graph, conditions local adaptation. This has non-trivial consequences on differentiation mechanisms, and we show that habitat assortativity can amplify or dampen neutral differentiation depending on the migration regime. The approach undertaken provides a rigorous framework to study macroscopic patterns emerging from microscopic processes developing over complex structures.

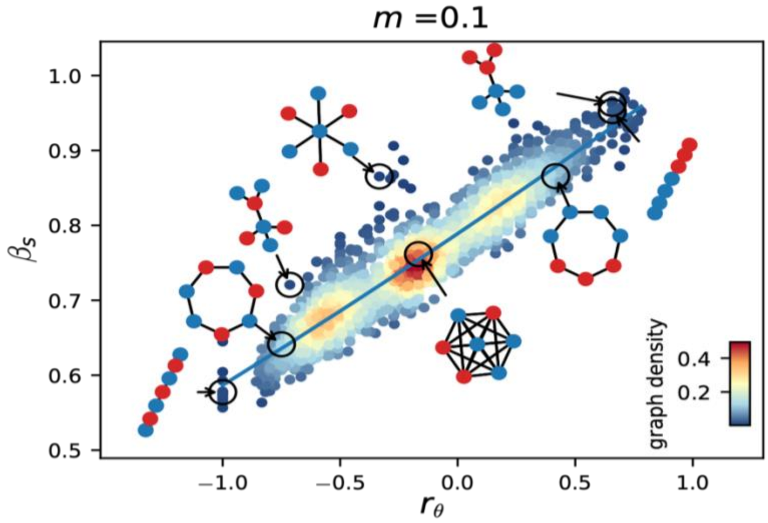


Figure 1: Habitat assortativity r_θ conditions adaptive differentiation β_s in graphs with heterogeneous habitats.

References

[1] N. Champagnat, R. Ferrière, S. Méléard, *Theoretical Population Biology*, (2006) *69*(3) 297–321.

Fishing at the High Seas: the Global Network Connecting Ports to Fishing Areas

J.P. Rodríguez-García[1], *J. Fernández-Gracia*[2], *C.M. Duarte*[3], *X. Irigoien*[4], *V.M. Eguíluz*[2]. *(1) Instituto Mediterráneo de Estudios Avanzados IMEDEA (UIB-CSIC), Esporles, Spain, jorgeprodriguezg@gmail.com. (2) Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (UIB-CSIC), Palma de Mallorca, Spain. (3) Red Sea Research Centre (RSRC) and Computational Bioscience Research Centre (CBRC), King Abdullah University of Science and Technology, Thuwal, Saudi Arabia. (4) AZTI Marine Research, Pasaia, Spain.*

Human activities and natural resources are coupled through biological, economical, political and social aspects. In particular, fishing is a complex activity due to the interaction of multiple factors. For example, there is an interplay between the specific interests of extracting industries and those of nations, having the right to regulate fishing within their Exclusive Economic Zones (EEZs). The problem of sustainable fishing becomes harder at the Areas Beyond National Jurisdiction (high seas), where the fishing stocks are shared resources and the regulation is weaker. We analyse one year of fishing vessels' trajectories, obtained from the Automatic Identification System (AIS), to understand how fishing is distributed at the high seas. We find an accumulation of fishing effort at the external borders of the most productive EEZs, hindering the sustainable fishing within those waters, such that 47% of the fishing effort in the high seas concentrates in a 200 km strip adjacent to the EEZ borders. To quantify this observation, we obtain 14 marine provinces from the trajectories of the vessels and link these provinces to the ports that give support to the vessels fishing on them. This leads to a global network where typically the ports are specialized on one or two provinces, and those located in low and middle-income countries have a key role, becoming potential candidates to receive rewards for control and catch verification tasks.

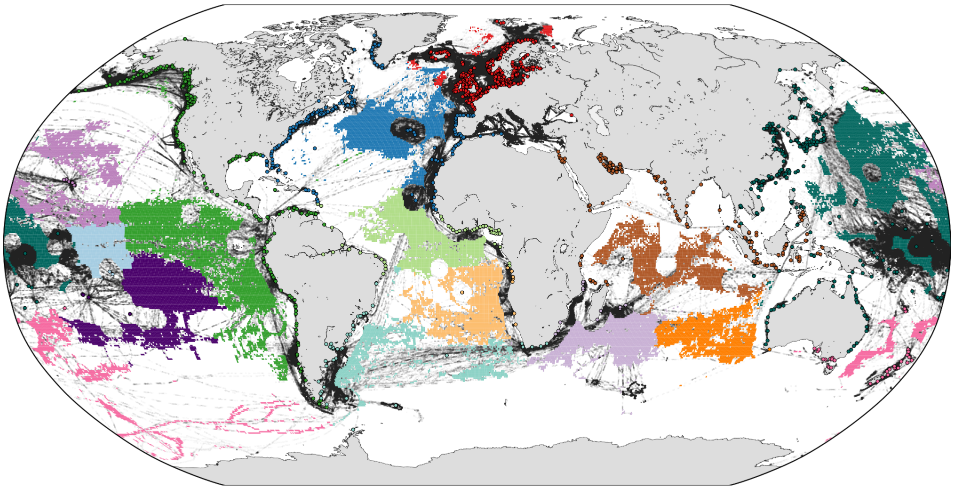


Figure 1: Marine provinces according to the trajectories of fishing vessels in the high seas in 2014. Different colours stand for different provinces. Circles at the shore depict ports, with their colour identifying the province that they are linked to, and black lines represent the trajectories between ports and fishing areas.

References

[1] J.P. Rodríguez, J. Fernández-Gracia, C.M. Duarte, X. Irigoien, V.M. Eguíluz, *Science Advances*, 7 (2021) eabe3470.

Inferring cell cycle phases from a partially temporal network of protein interactions

M. Lucas *[1,2,3,†] , A. Townsend-Teague [4] , A. Morris [4] , L. Tichit [3] , B. H. Habermann [2] , A. Barrat [1,] (1) Aix Marseille Univ, CNRS, CPT, Turing Center for Living Systems, Marseille France, (2) Aix Marseille Univ, CNRS, IBDM, Turing Center for Living Systems, Marseille France, (3) Aix Marseille Univ, CNRS, I2M, Turing Center for Living Systems, Marseille France, (4) University of Oxford, Oxford, UK, † maxime.lucas.1@univ-amu.fr*

The temporal organisation of biological systems into phases and subphases is often crucial to their functioning. Identifying this multiscale organisation can yield insight into the underlying biological mechanisms at play. To date, however, this identification requires a priori biological knowledge of the system under study. Here, we recover the temporal organisation of the cell cycle of budding yeast into phases and subphases, in an automated way. To do so, we represent the cell cycle as a partially temporal network of protein-protein interactions (PPIs) by combining a traditional static PPI network with protein concentration or RNA expression time series data. Then, we cluster the snapshots of this temporal network to infer phases, which show good agreement with our biological knowledge of the cell cycle. We systematically test the robustness of the approach and investigate the effect of having only partial temporal information. Our results show that a temporal network with only partial temporal information, i.e. for some of the PPIs, is sufficient to infer the temporal organization of a system. The generality of the method makes it suitable for application to other, less well-known biological systems for which the temporal organisation of processes plays an important role.

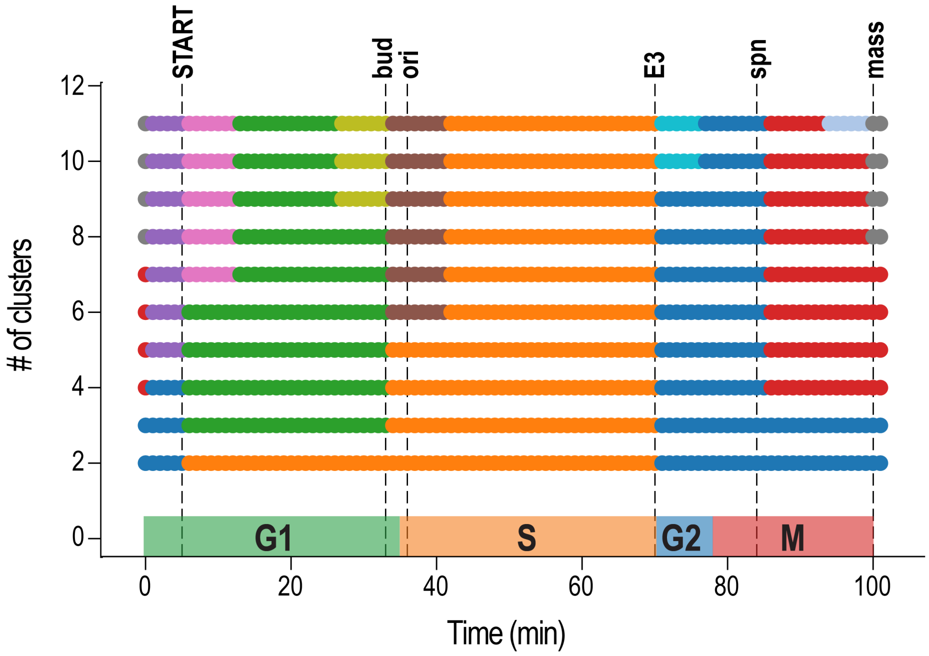


Figure 1: Inferred phases of the cell cycle , over a range of time scales (number s of clusters).

References

[1] M. Lucas, A. Morris, A. Townsend-Teague, L. Tichit, B. H. Habermann, and A. Barrat. Inferring cell cycle phases from a partially temporal network of protein interactions. bioRxiv:10.1101/2021.03.26.437187, 2021.

Nonmedical Opioid Consumption on Reddit: Patterns of Routes of Administration and Drug Tampering

D. Balsamo[1], *P. Bajardi*[2], *A. Salomone*[1], *R. Schifanella*[1,2]. *(1) University of Turin, Turin, Italy, duilio.balsamo@unito.it. (2) ISI Foundation, Turin, Italy*

Despite the efforts of traditional pharmacology research in understanding the health-associated risks of opioid usage through various routes of administration (ROA) and methods of drug tampering [1], no large-scale empirical evidence has been found to unveil the relationships between substance manipulation, unconventional ROA, and nonmedical substance consumption. This work aims at characterizing the patterns of the nonmedical consumption of opioids in the US, as discussed online by users on Reddit, with a focus on ROA, and drug tampering methods. First, leveraging a recent query expansion methodology, we identified a large cohort of over 86,000 Reddit users potentially involved in firsthand opioid usage from 2014 to 2018. Second, we developed a vocabulary expansion procedure based on word embeddings to identify and catalogue a large set of terms describing practices of nonmedical opioids consumption (Figure1(a)): these terms are crucial to perform exhaustive analyses of user-generated content from social media, as they include colloquialisms and slang terminology hardly used in the medical literature. Third, we modeled the temporal evolution of interest in opioid consumption by evaluating the popularity of opioid-related terms for 5 years, observing relevant trends such as the surge of synthetic opioids like fentanyl, and an increasing interest in rectal administration. Lastly, by measuring odds ratios based on co-mentions, we quantified the strength of association between opioid substances, ROA and drug-tampering methods to better characterize emerging abusive practices. We found quantitative evidence of known nonmedical opioid consumption behaviors, like the intranasal administration of crushed pills, as well as understudied abusive behaviors, like chewing fentanyl patches and dissolving buprenorphine sublingually (Figure 1(b)).

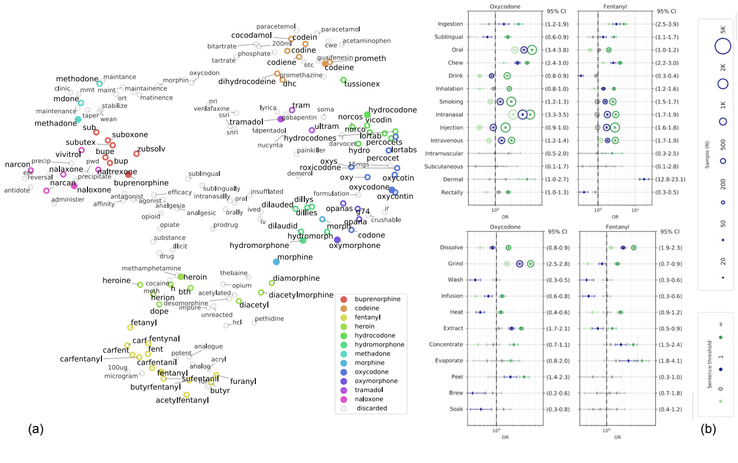


Figure 1: a) UMAP projection of the embedding, modeling the semantic relationships among terms. b) Odds ratios of opioid and ROA (top) and drug-tampering methods (bottom).

References

[1] Katz N, et al., *The American journal of drug and alcohol abuse. 2011;37(4):205–217.*

Emergent Order in a Continuous State Adaptive Network Model of Living Systems

G. Dadunashvili[1], C.T van de Kamp[1], T. Idema[1], ***J.L.A. Dubbeldam***[1]*,(1) Delft University of Technology, Mekelweg 4, 2628CD Delft, The Netherlands, j.l.a.dubbeldam@tudelft.nl.*

Swarming is a collective phenomenon which is encountered in many biological systems varying from schools of fiish to flocks of birds. Network models, such as those put forward by Vicsek *et al.* [1] in the context of opinion formation can also be used to find heuristical models for swarming behavior. In this paper we generalize a discrete model of Chen *et al.* [2] to a *continuous setting,* which is more relevant to most biological systems, and investigate how *spontaneous order* -corresponding to swarming behavior- arises. The interplay between the changing network topology and the physical movement of agents leads to a phase transition from a disordered to an ordered swarming state. We show [3] that the the nature of this transition changes from *first-order to second-order* when the time scale of the physical movement of the agents and that of the update of the network topology become further separated.

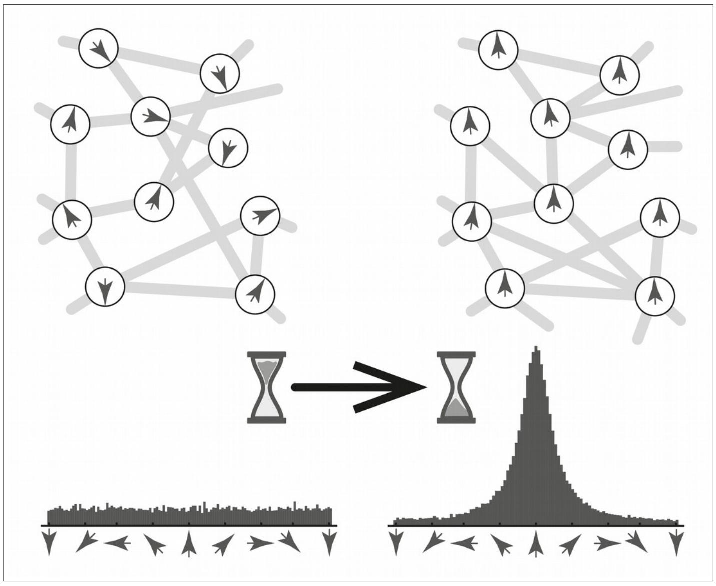


Figure 1:An illustration of a swarming phase transition in an adaptive network.***Top***: Each node represents an agent that can point in anydirection on a plane. These agents are connected to each other through links, which represent mutual awareness.***Top left:*** The agents do not have a common heading direction and the system is disordered.***Top righ***t:Through interactions the agents can form a coherently moving swarm and choose an overall heading direction.***Bottom***:How often a certain state occurs in the network tells us if the network is ordered or disordered. ***Bottom left***: Different states occur with roughly the same frequency, whereas on the ***Bottom right***: one direction is dominant.

References

[1] T. Vicsek and A. Zafeiris, *Physics Reports* 517, (2012), 71.

[2] L. Chen, C. Huepe, and T. Gross, *Physical Review E,* 94, (2016) 022415.

[3] C.T. van de Kamp, G. Dadunashvili, J.L.A. Dubbeldam, T. Idema, *arXiv:2103.04312*

Minimizing Communication Costs with Asymmetric Weight Codes

Pau Vilimelis Aceituno. *Institute of Neuroinformatics, ETH Zürich, pau@ini.uzh.ch*

Motivation

When Shannon presented his theory of information transmission, he assumed that the cost of sending any of the symbols {0,1} was the same, and that the number of bits per codeword could be changed. While these assumptions made sense for reliable computation over electronic circuits, many applications are not compatible with them. For instance, DNA has four different bases and their energy bonds are different, and in neuroscience the cost of emitting an action potential is higher than the cost of remaining idle. In this project we present a framework which accounts for such optimization.

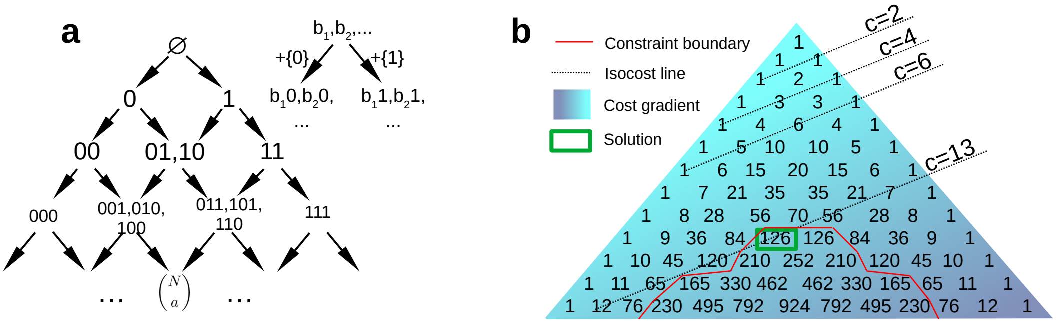


Figure 1: Graphical explanation of minimizing the cost $C(a,N) = a+N$ **subject to** $S = 125$: In **a** we construct the number of possible codewords for a fixed number of active units by starting with an empty string and adding one unit, either active (right-low arrow) or inactive (left-low arrow). In **b** we show that this way of generating Pascal's triangle gives us a graphical interpreration of our problem: the boundary of our constraint (red line) marks the values of N and a that have enough codewords while the blue gradient gives the cost. The minimum cost is 13 at N=9 with a=4.

Approach

As shown in Figure 1, the starting point of our derivations is to study fixed weight codes and find the parameters that minimize a given cost function. This calculations are easy if we use Stirling's approximation on the combinatorial constraint, which yields $N \geq log(S)/H(f)$, where S is the number of required codewords and $f=a/N$. To account for noise we compute how much distance between codewords is required to guarantee decodability, which is a well studied problem in fixed weight codes. However, to the best of our knowledge previous works focused on deterministic codes, while our interaction with Neuromorphic Engineers suggests that random codes would be a more interesting scenario. We thus follow an approach similar to the one used in *Mezard & Montanari, 2009* to prove the Noisy Channel Theorem with Random Codes using the so-called distance enumerator, but on fixed weight codes.

Applications

Our framework is extremely flexible, but it requires knowing the cost function. Although allometric scaling laws for metabolism can be applied to the brain and the stability of DNA bases is well know and will be discussed, to the best of our knowledge we lack data to probe the effective cost function, at least in neurobiology. However, such knowledge is readily available for engineered systems such as neuromorphic systems, delay insensitive circuits or DNA-based information storage. Furthermore, the codes presented here can be used to find bounds in partitioning of networks into subnetworks with predefined overlaps, helping design complex systems where different subsystems must retain independence.

[1] P.V.A, Minimizing costs of communication with random constant weight codes, arxiv.org

A recurrence study of the effects of early vocal development on the birdsong of a juvenile canary

Rabindev Bishal[1], Gabriel Mindlin[2,3] , Neelima Gupte[1,3] (1) Department of Physics, IIT Madras, Chennai, India, rabindevbishal321@gmail.com, (2) Dept of Physics, University of Buenos Aires, Buenos Aires, Argentina, (3) Complex Systems and Dynamics Group, IIT Madras, Chennai, India

The early vocal development of a canary has important consequences for its bird song. These effects are characterised via a recurrence study. The songs of a canary bird are recorded at different stages in the early development part of its vocal system. The syllabic sequences from each of the recorded songs are extracted and quantified using the chroma vector, where the chroma vector is defined as a finite elements representation of spectral energy. The recurrence plot of chroma vector of the syllabic sequences shows a characteristic self similar pattern which gets more organized and shows the gradual development of recurrent structures in the syllabic sequences with increase in vocal development (see Fig. 1). Quantities such as the laminarity (defined from the vertical recurrent lines in the recurrence plot) and the determinism (defined from the diagonal recurrent lines in recurrence plot) show a increase in the vocal development of the bird with increasing age. Thus, our recurrence plots and the associated computed quantities demonstrate the increasing control on the syllabic sequences which emerges with increase in vocal development at different stages.

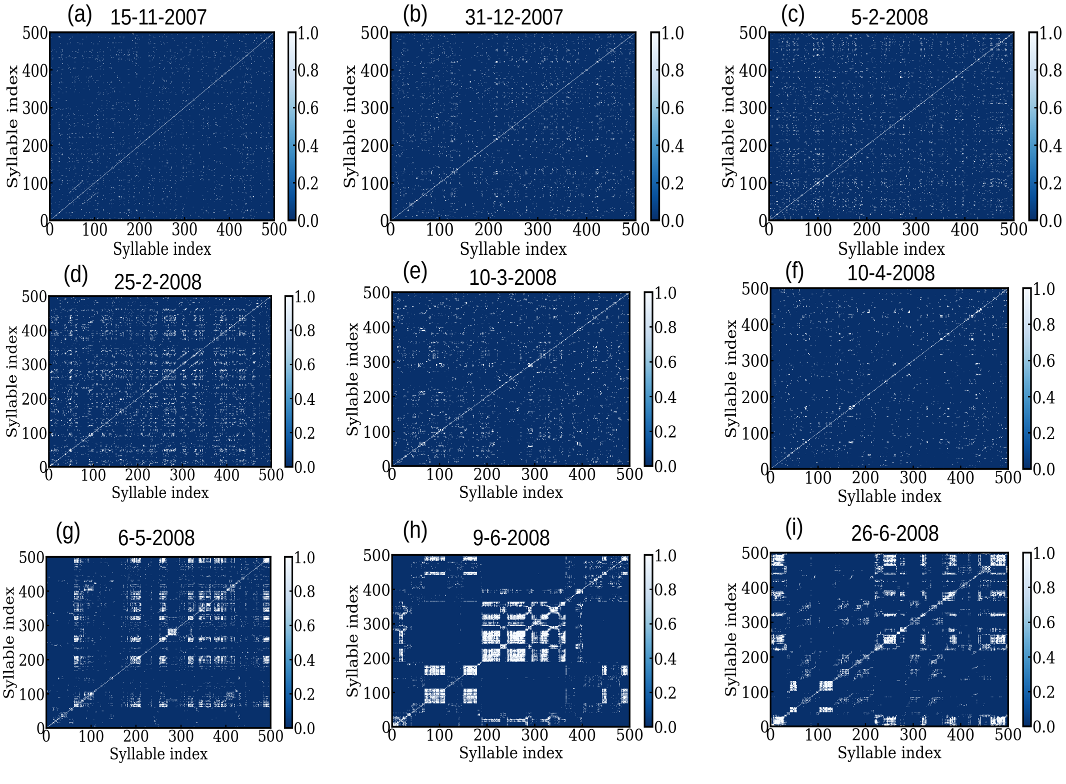


Figure 1: Recurrence plots of the syllabic sequence at a different stage of the vocal development.

References:

[1] C. Webber and N. Marwan, Recurrence quantification analysis, Theory and Best Practices (2015).

[2] Jorge A Alliende, Jorge M M´endez, Franz Goller, and Gabriel B Mindlin. Hormonal acceleration of song development illuminates motor control mechanism in canaries. *Developmental neurobiology*, 70(14):943–960, 2010.

[3] T. Giannakopoulos and A. Pikrakis, *Introduction to audio analysis: a MATLAB® approach* (Academic Press, 2014).

Flocking in complex environments—Attention trade-offs in collective information processing

P. Rahmani[1,2]*, F. Peruani*[3]*,* ***P. Romanczuk***[2,4]*. (1) Department of Physics, Institute for Advanced Studies in Basic Sciences (IASBS), Zanjan 45137-66731, Iran. (2) Institute for Theoretical Biology, Department of Biology, Humboldt Universität zu Berlin, Germany. (3) Laboratoire de Physique Théorique et Modélisation, CNRS UMR 8089,CY Cergy Paris Université, F-95302 Cergy-Pontoise Cedex, France. (4) Bernstein Center for Computational Neuroscience, Berlin, Germany. pawel.romanczuk@hu-berlin.de*

The ability of biological and artificial collectives to outperform solitary individuals in a wide variety of tasks depends crucially on the efficient processing of social and environmental information at the level of the collective. Here, we model collective behavior in complex environments with many potentially distracting cues [1]. Counter-intuitively, large-scale coordination in such environments can be maximized by strongly limiting the cognitive capacity of individuals, where due to self-organized dynamics the collective self-isolates from disrupting information. We observe a fundamental trade-off between coordination and collective responsiveness to environmental cues. Our results offer important insights into possible evolutionary trade-offs in collective behavior in biology and suggests novel principles for design of artificial swarms exploiting attentional bottlenecks.

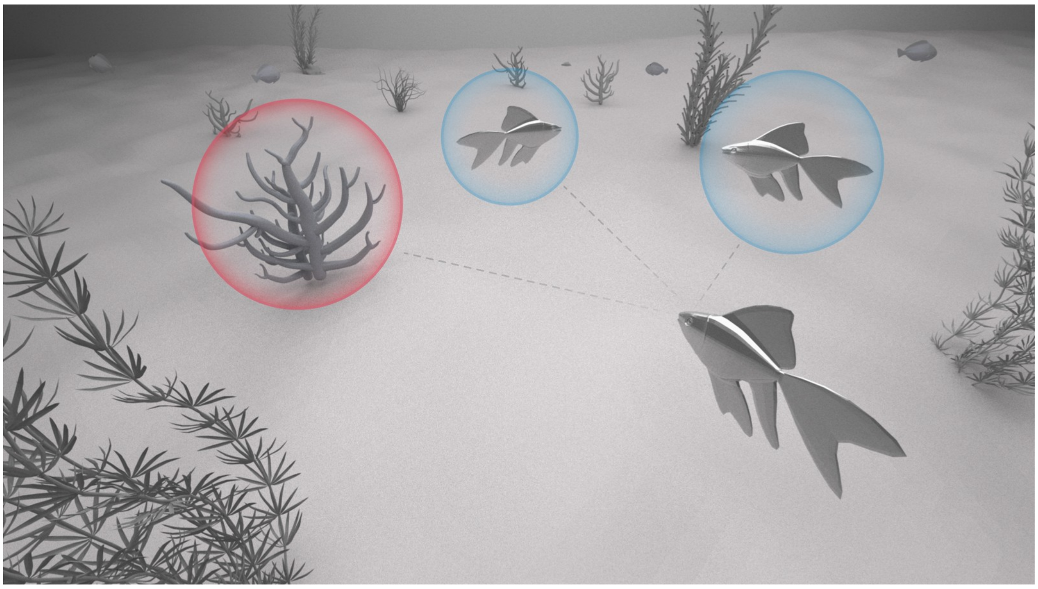


Figure 1: Model motivation via a schematic visualization of attention trade-off in collective behavior in complex environments. The focal individual can only pay attention to $k = 3$ nearest objects—other agents or non-social environmental features—simultaneously.

Acknowledgements

We thank Ana Sofía Peruani for providing the artwork in Fig 1.

References

[1] P. Rahmani, F. Peruani, P. Romanczuk, *PLoS Computational Biology*, 16 (2020) e1007697.

A model of collective behavior based purely on vision

*R. Bastien[1], **P. Romanczuk**[2,3]. (1) Institut de Recherche en Informatique de Toulouse, 31400 Toulouse, France. (2) Institute for Theoretical Biology, Department of Biology, Humboldt Universität zu Berlin, Germany. (3) Bernstein Center for Computational Neuroscience, Berlin, Germany. pawel.romanczuk@hu-berlin.de*

Classical models of collective behavior often take a “bird’s-eye perspective,” assuming that individuals have access to social information that is not directly available (e.g., the behavior of individuals outside of their field of view). Despite the explanatory success of those models, it is now thought that a better understanding needs to incorporate the perception of the individual, i.e., how internal and external information are acquired and processed. In particular, vision has appeared to be a central feature to gather external information and influence the collective organization of the group. Here, we show that a vision-based model of collective behavior is sufficient to generate organized collective behavior in the absence of spatial representation and collision [1]. Our work suggests a different approach for the development of purely vision-based autonomous swarm robotic systems and formulates a mathematical framework for exploration of perception-based interactions and how they differ from physical ones.

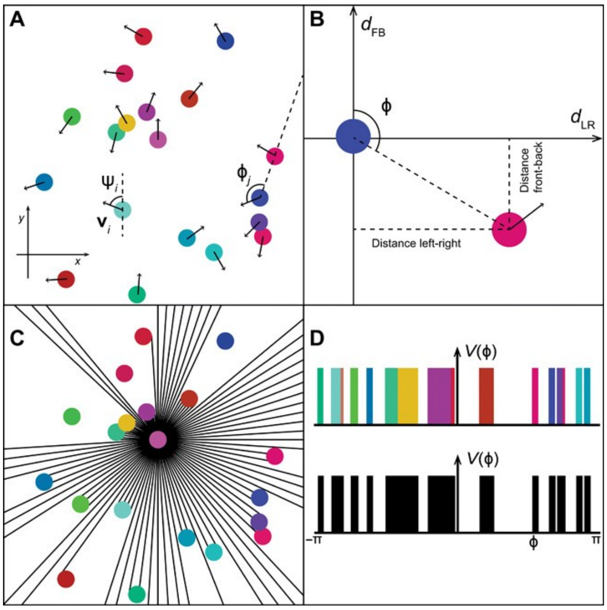


Figure 1: **Model setup** (adapted from [1]) (A) A set of disk-shaped agents with diameter BL (body length) is considered. Each disk is propelled in the direction ψ_i with a velocity $v_i(t)$. (B) A co-moving referential can be defined following the movement of the disk j. This referential is centered on the position of the disk, (x_j,y_j), and oriented so that the vertical axis is aligned with the direction ψ_j, ϕ represents the polar angle in the co-moving frame. (C) Representation of the visible field of the pink disk through ray casting. The position of the eye is considered to be at the center of the disk with a fully circular point of view. (D) The projection of the visual field in 2D is given by a 1D function. On top, objects can be represented by their colors. On the bottom part, a minimal binary visual field is given, lacking the ability to distinguish individuals.

References

[1] R. Bastien, P. Romanczuk, *Science Advances*, 6 (2020) eaay0792.

Structural Predictors for Species Survival in Ecological Networks

Violeta Calleja-Solanas[1], *Emilio Hernández-García*[1], *Sandro Meloni*[1]. *(1) Institute of Cross-Disciplinary Physics and Complex Systems (IFISC, UIB-CSIC), Palma de Mallorca, Spain , violeta@ifisc.uib-csic.es*

Understanding the characteristics that promote ecosystems' resilience is crucial to explain and protect biodiversity [1]. In this picture, environmental changes may alter the equilibrium of species interactions, leading to cascade extinctions. For this reason, understanding which species are more prone to initiate cascades is of paramount importance [2].

To do so, here [3], we study the structural features of ecological network that better predict species' survival when the strength of their interaction is varied. Employing the replicator equation as a proxy for species dynamics, we find that a high eigenvector centrality is a key factor for survival in mutualistic networks. This result reinforces the importance of the species's neighbors identity and the existence of positive feed-backs that prevent ecosystem collapse. On the contrary, this panorama changes with competitive interactions. In that case, the species that go extinct by extinctions cascades are determined not by their low eigenvector centrality but by a high degree. When both competitive and mutualistic interactions are considered, the behavior of the system is richer. In this case, not only networks are more resilient, eigenvector centrality and degree alone cannot fully predict species fate. Other individual properties such as the dynamical ratio between mutualistic and competitive interactions also plays a fundamental role. Finally and looking at the big picture, our work demonstrates that the delicate interplay between different types of interactions shapes ecosystems resilience and the factors that can lead to their collapse.

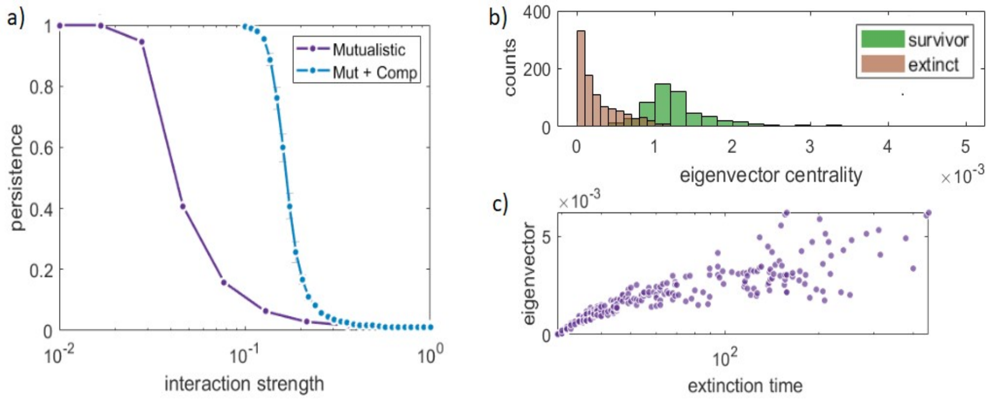


Figure 1: a) Species' persistence vs interactions strength: considering both mutualistic and competitive interactions increases ecosystems' resilience. The two networks are empirical pollination networks. b) Distribution of the eigenvector centrality for extinct and surviving species: an initial low eigenvector centrality is a good predictor of species extinction probability. c) Species' eigenvector centrality as function of their extinction time: not only species with a low eigenvector centrality are more likely to go extinct, but they also are the ones starting the extinction cascades.

References

[1] E. Thébault, C. Fontaine, *Science*, 329, 852-856 (2010).

[2] F. Morone, G. Ferrano, H. Makse, Nature Physics, 15, 95-102 (2019).

[3] V. Calleja-Solanas, Emilio Hernández-García and Sandro Meloni, *submitted* (2021).

6 Social complexity

Contents

Reputation and punishment sustain cooperation in the Optional Public Goods Game

S. Righi[1], *S. Podder*[2], *F. Pancotto*[3]. *(1) Ca'Foscari University of Venice, Fondamenta S. Giobbe, 873, Venezia, Italy, simone.righi@unive.it (2) University College London, Gower Street, London, UK (3) University of Modena and Reggio Emilia, Viale Antonio Allegri, 9, Reggio Emilia, Italy.*

Cooperative behaviour has been extensively studied, in both evolutionary biology and the social sciences, as a choice between cooperation and defection. However, in many cases, the possibility to not participate or to exit a situation is also available. This type of problem can be studied through the optional public goods game (OPGG). The introduction of the 'Loner' strategy, allows players to withdraw from the public goods game, radically changing the dynamics of cooperation in social groups and leading to a never-ending cooperator-defector-loner cycle. While pro-social punishment has been found to help increase cooperation, anti-social punishment - where defectors punish cooperators - causes the downfall of cooperation in both experimental and theoretical studies. In this paper, we extend the theory of the optional public goods game, introducing reputational dynamics in the form of social norms that allow agents to condition both their participation and contribution decisions to the reputation of their peers. We benchmark this setup both with respect to the standard optional public goods game and to the variant where all types of punishment are allowed. We find that a social norm imposing a more moderate reputational penalty for opting out than for defecting, increases cooperation (**AD** in Figure 1). When, besides reputation, punishment is also possible, the two mechanisms work synergically under all norms that do not punish loners too harshly. Under this latter setup, the high levels of cooperation are sustained by conditional strategies, which largely reduce the use of pro-social punishment and almost eliminate anti-social punishment.

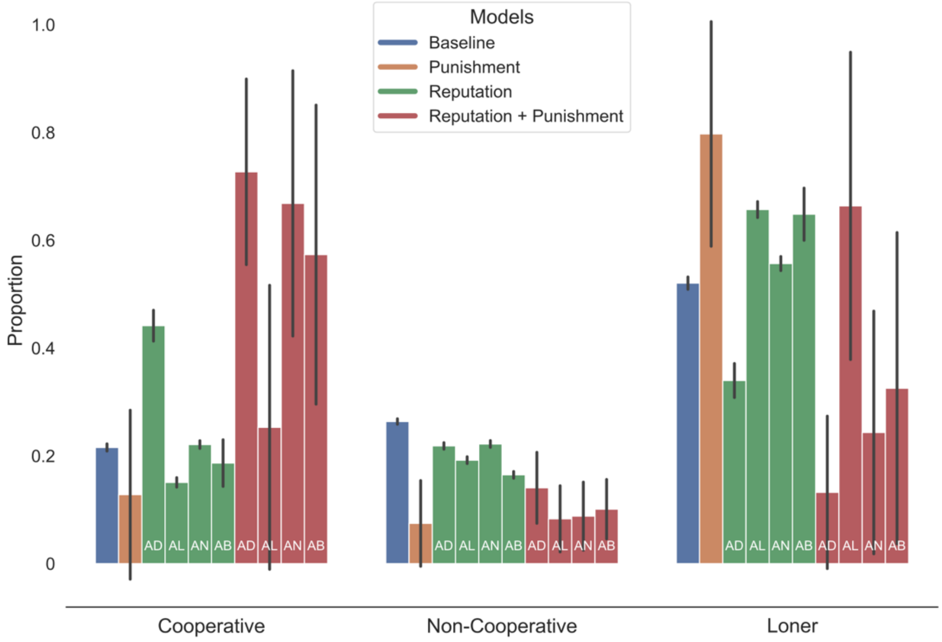


Figure 1: Distribution of cooperative, non-cooperative and loner actions in the OPGG without reputation or punishment (Baseline), with reputation, with punishment, and with both reputation and punishment under social norms punishing equally (AN and AB) and differently (AD and AL) loners and defectors.

Vulnerabilities of Democratic Electoral Systems

Mateusz Wiliński[1], Jarosław Klamut[2], Paul Bouman[3],
Juan Fernández-Gracia[4], Roger Cremades[5,6], **Tomasz Raducha**[2,4,*]

[1] Theoretical Division, Los Alamos National Laboratory, USA
[2] Faculty of Physics, University of Warsaw, Poland
[3] Erasmus School of Economics, Erasmus University, Rotterdam, The Netherlands
[4] Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Palma, Spain
[5] Wageningen University and Research, Wageningen, The Netherlands
[6] Climate Service Center Germany (GERICS), Hamburg, Germany
*tomasz@ifisc.uib-csic.es

The vulnerability of democratic processes is under scrutiny after scandals related to Cambrige Analytica (2016 U.S. elections, the Brexit referendum, and elections in Kenya [1]). The deceptive use of social media in the US, the European Union and several Asian countries, increased social and political polarization across world regions. Finally, there are straightforward frauds like Crimea referendum and Belarus elections. These challenges are eroding democracy, the most frequent source of governmental power, and raise multiple questions about its vulnerabilities [2].

Democratic systems have countless ways of performing elections, which create different electoral systems (ES). It is therefore in citizens' interest to study and understand how different ESs relate to different vulnerabilities and contemporary challenges. These systems can be analyzed using network science in various layers – they involve a network of voters in the first place, a network of electoral districts connected by commuting flow for instance, or a network of political parties to give a few examples. The electoral system together with the underlying voting processes and opinion dynamics can be seen as a complex system [3].

We study electoral systems in a dynamical framework. We look at the volatility of the election results, analyzing how much they vary over time. However, the term volatility is frequently used in relation to the Pedersen index of volatility. In this meaning it has been studied and even linked to the party system instability [4, 5]. Our approach goes far beyond two-point volatility. We analyze the vulnerability of an ES based on a long run of opinion dynamics process with many elections performed during the evolution. In this context, we consider that a system is more vulnerable, if it has a larger variance of the election results and if it magnifies the influence of extremism and media. We can further identify which voting system is more sensitive to fluctuations, and which one is more vulnerable to internal/external influences, like zealots or propaganda. This allows us to construct a probability distribution of election results under every electoral system.

It is essential to provide new tools and arguments to the discussion on the evaluation of electoral systems. We aim at comparing different ESs in a dynamical framework. Our novel approach of analyzing electoral systems in such way with all its aspects included, from opinion dynamics in the population of voters to inter-district commuting patterns to seat appointment methods, will help answering questions like: Which electoral systems are more predictable/stable under fluctuations? Which electoral systems are the most robust (or vulnerable) under external and internal influences? Which features of electoral systems make them more (less) stable?

References

1. Laterza, V.: Cambridge analytica, independent research and the national interest. Anthropology Today **34**(3) (2018)
2. Bennett, W.L., Livingston, S. European journal of communication **33**(2) (2018) 122–139
3. Ladyman, J., Lambert, J., Wiesner, K.: What is a complex system? European J. for Philosophy of Science **3**(1) (2013)
4. Tavits, M. European Journal of Political Research **47**(5) (2008) 537–555
5. Marinova, D.: A new approach to estimating electoral instability in parties. Available at SSRN 2636806 (2015)

Local Reputation, Local Selection, and the Leading Eight Norms

S. Podder[1], *S. Righi*[2], *K. Takács*[3]. *(1) University College London, Gower Street, London, UK, ucabpod@ucl.ac.uk. (2) Ca'Foscari University of Venice, Fondamenta S. Giobbe, 873, Venezia, Italy. (3) Linköping University, Kopparhammaren 7, Linköping, Sweden.*

Humans are capable of solving cooperation problems following social norms. Social norms dictate appropriate behaviour and judgement on others in response to their previous actions and reputation. Recently, the so-called *leading eight* norms [1] have been identified from many potential social norms that can sustain cooperation through a reputation-based indirect reciprocity mechanism. Despite indirect reciprocity being claimed to extend direct reciprocity in larger populations where direct experiences cannot be accumulated, the success of social norms has been analyzed in models with global information and evolution. This study is the first to analyse the leading eight norms with local information and evolution. We find that the leading eight are robust against selfish players within most scenarios and can maintain a high level of cooperation also with local information and evolution. In fact, local evolution (Le) sustains cooperation under a wider set of conditions than global evolution (Ge), while local reputation (Lr) does not hinder cooperation compared to global reputation (Gr). Four of the leading eight norms that do not reward justified defection offer better chances for cooperation with quick evolution, reputation with noise, larger networks, and when unconditional defectors enter the population.

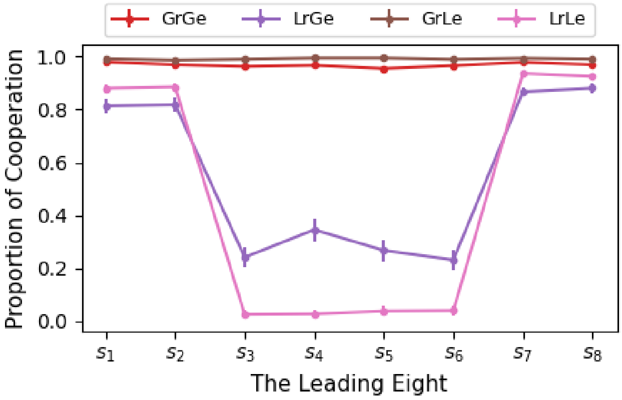


Figure 1: Social norms that do not reward justified defection (s_3-s_6) are the least resilient towards harsher environments and scarce or unreliable reputational information.

References

[1] H. Ohtsuki & Y. Iwasa (2004, November). 'How Should We Define Goodness? – Reputation Dynamics in Indirect Reciprocity. *Journal of Theoretical Biology* (*231), 107-120.*

A comparison of community-aware centrality measures in online social networks

S. Rajeh[1]*, M. Savonnet*[1]*, E. Leclercq*[1]*, H. Cherifi*[1]*.* [1]*Laboratoire D'Informatique de Bourgogne, University of Burgundy, 21000 - Dijon, France. E-mail: stephany.rajeh@u-bourgogne.fr.*

Identifying influential nodes is a fundamental task for amplifying or hindering any spreading phenomena. Community-aware centrality measures exploit the community structure to address this issue. More precisely, local measures based on intra-community links focus on the impact of a node in its community. In contrast, global ones rely on inter-community links to characterize a node's global influence. These measures differ in the way they handle the two link types. A community-aware centrality measure could emphasize the ability of nodes to link multiple communities, while another might target local hubs rather than bridges between communities. Previously, we looked at the correlation between classical and community-aware centrality measures and their relation to network topology [1]. Results show that the correlation between community-aware and classical centrality measures is very sensitive to the community structure strength. This work goes one step further. We compare the effectiveness of various community-aware centrality measures in an epidemic-spreading scenario. Considering the Susceptible-Infected-Recovered (SIR) spreading process, we perform extensive simulations on a set of ten online social networks and seven influential community-aware measures (Comm Centrality, Community-based Mediator, Community Hub-Bridge, K-shell with Community-, Community-based Centrality, Participation Coefficient, and Modularity Vitality). Results show that K-shell with Community and Community-based Centrality are the best performers in a single-spreader setting. Indeed, they have the lowest imprecision function (Figure 1 - Top). In other words, they predict more accurately the influential nodes compared to the nodes' ground truth spreading power. Additionally, Community-based Mediator performs better in networks with a weak community structure. Finally, results are generally insensitive to the variation of the SIR transmission rate (Figure 1 - Bottom). This study gives clear indications about the effectiveness of prominent community-aware centrality measures in a single-spreader scenario.

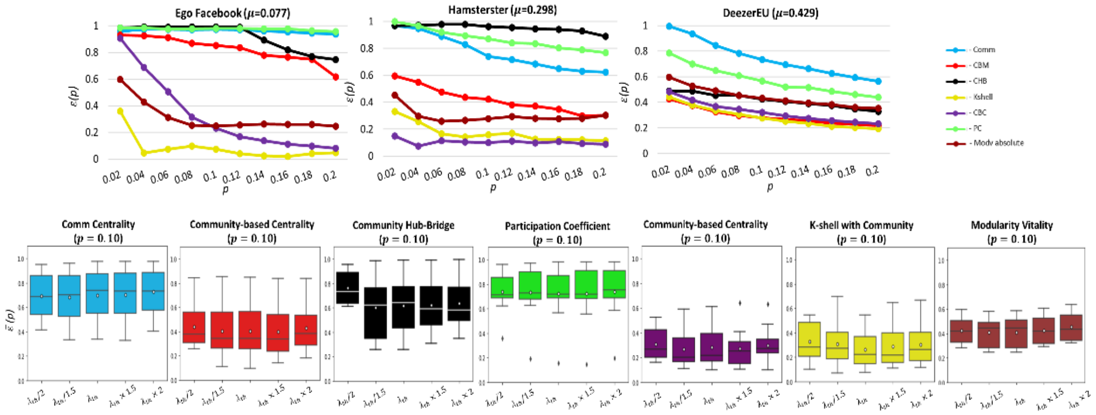


Figure 1: The top figures show the performance of the seven community-aware centrality measures based on the imprecision function $\varepsilon(p)$ which quantifies their predictive power in identifying influential nodes of top p nodes on 3 online social networks. The bottom figures show the average imprecision function $\bar{\varepsilon}(p)$ of each of the seven community-aware centrality measures as a function of different transmission rates set based on the epidemic threshold of the ten online social networks (λ_{th}).

References

[1] Rajeh, S., Savonnet, M., Leclercq, E. et al. Characterizing the interactions between classical and community-aware centrality measures in complex networks. Sci Rep 11, 10088 (2021).

Characterising different communities of Twitter users: Migrants and natives

Jisu Kim[1], Alina Sîrbu[2], Giulio Rossetti[3], and Fosca Giannotti[3]. (1) Scuola Normale Superiore, Pisa, Italy and Max Planck Institute for Demographic Research, Rostock, Germany. kim@demogr.mpg.de. (2) University of Pisa, Italy. alina.sirbu@unipi.it. (3) Institute of Information Science and Technologies - Italian National Research Council, Italy. fosca.giannotti,{fosca.giannotti,giulio.rossetti} @isti.cnr.it

Today, many users are actively using Twitter to express their opinions and to share information. Thanks to the availability of the data, researchers have studied behaviours and social networks of these users. International migration studies have also benefited from this social media platform to improve migration statistics. Although diverse types of social networks have been studied so far on Twitter, social networks of migrants and natives have not been studied before. This paper aims to fill this gap by studying characteristics and behaviours of migrants and natives on Twitter. To do so, we perform a general assessment of features including profiles and tweets, and an extensive network analysis on the network. We find that migrants have more followers than friends. They have also tweeted more despite that both of the groups have similar account ages. More interestingly, the assortativity scores showed that users tend to connect based on nationality more than country of residence, and this is more the case for migrants than natives. Furthermore, both natives and migrants tend to connect mostly with natives.

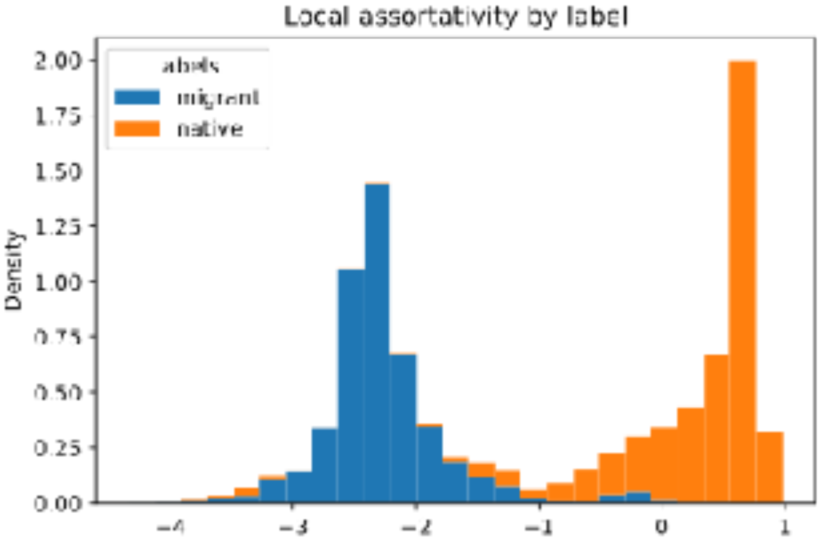


Figure 1: Stacked histogram of local assortativity by migrant/native label. Please note that the histogram is stacked, therefore there is no overlap between the plot bars

Behavioural patterns of hate speech and misinformation on Twitter

R. Gallotti[1]*, E. Leonardelli*[1]*, S. Menini*[1]*, P. Sacco*[1,2]*, S. Tonelli*[1]*, M. De Domenico*[1]*.*

(1) Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo (TN), Italy, rgallotti@gmail.com. (2) IULM University, Via Carlo Bo 1, 20143 Milano, Italy

Online communication through social media is mediated by complex behavioral processes involving the production and the diffusion of textual content. The linguistic features of the messages circulating online can be characterized by either focusing on the content shared or the communicative intent behind the message. In this study, we attempt to provide a comprehensive view by examining at the same time both communication content and intent in 4 datasets describing the Twitter discussion around the Covid-19 pandemics between January and May 2020 in USA, UK and Italy and the discussion around the George Floyd protest movement in the USA (80M+ tweets in total, all in the local language). In this study, we characterize the content in term of reliability using the analysis method developed for the Covid-19 Infodemic Observatory [1]. The intent is instead measured as the level of hate speech classified with Google's BERT model [2], fine-tuned on domain-specific datasets, able to identify tweets written with the intent of offending or attacking someone. We characterize the behavioral patterns of hate speech observing the dependency of the hate level on other features of the tweet (text length, URL shared) and of the users (verified status, number of followers). We report significant positive correlations between the tendency of sharing fake news and to use of offending language in all Covid19 datasets but not in the one centered around the Floyd movement. We propose evidence of a behavioral change in the way users having a large social reputation are systematically less prone to write message with an offensive intent, as users with larger following, and verified users, display smaller values of hate across all datasets (Fig. 1). More heterogeneous patterns emerge when studying the level of misinformation shared by verified and unverified of varying number of followers, as in some cases influencers might decide to invest their social reputation in pushing their (often politically biased) agenda.

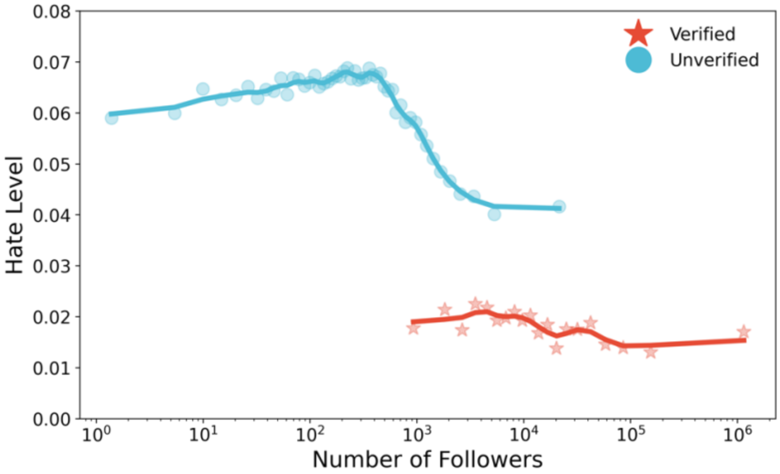


Figure 1: Hate level of verified and unverified users with different following in the UK.

References

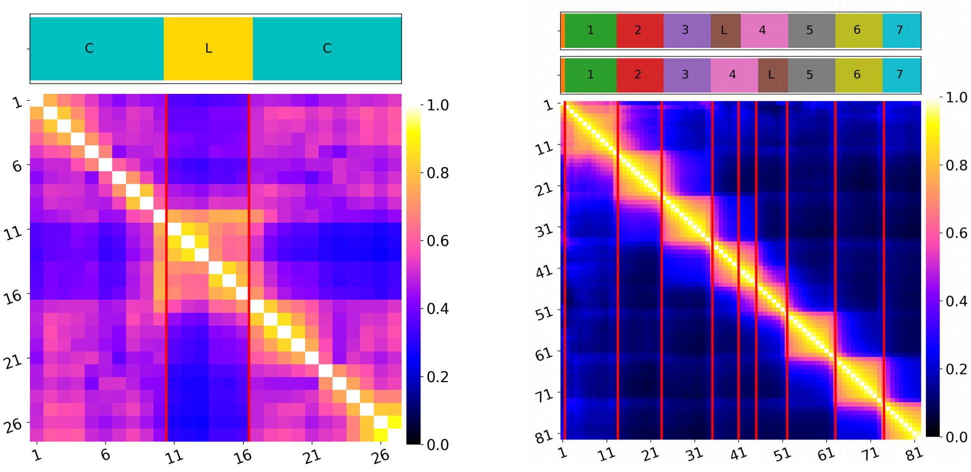
[1] R. Gallotti et al., Nature Human Behavior (2020).

[2] J. Devlin, et al., Proceedings of NAACL (2019).

From temporal network data to the dynamics of social relationships

Valeria Gelardi,[1, 2] Alain Barrat,[1] and Nicolas Claidière[2]
[1]*Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France*
[2]*Aix Marseille Univ, CNRS, LPC, FED3C, Marseille, France*

Networks are well-established tools to represent social systems, and temporal networks are now widely used to describe their dynamics. Temporal network data are often presented as a succession of static networks aggregated on successive time windows, whose length however is somehow arbitrary, not necessarily corresponding to any intrinsic timescale of the system. Moreover, aggregating on increasing time window lengths does not bring a satisfactory view of the social network evolution, as short time windows contain too little information whereas aggregating over large time windows blurs the dynamics. How to transform a temporal network into a sensible continuously evolving representation of the social network remains therefore a challenge. Here we introduce a framework to that purpose: it transforms temporal network data of discrete dyadic interactions into an evolving weighted network in which the weights of the links between individuals are updated at every event of the temporal network. The main assumption is that each individual has a limited capital of ”attention” or available time that has to be divided among peers: as a result, if individual i has repeated interactions with a peer j and not with others, the tie between i and j strengthens (rate controlled by a parameter α) while the ties between i and other peers deteriorates (at rate β). The parameters α and β determine the sensitivity of the representation to the underlying interaction dynamics. We illustrate the interest of our framework using real data describing interactions of (i) groups of humans, namely in a French and a US school and (ii) a group of baboons. We show how computing the matrix of similarities between the evolving networks at different times unveils the differences in the organizations of the schools and the various periods of importance in the school schedules (see Figure 1). Moreover, we use the baboons' data to simulate a perturbation in a group's social structure by switching the identity of 2 animals in the data at a certain time t_0. We then investigate the capacity of our data representation to automatically detect the time of the perturbation by a hierarchical clustering of the matrix of similarities between representations computed at different times. We show that the perturbation is correctly identified with a robust performance on a broad range of parameter values. Our framework could be extended in various directions, and provides a simple and natural way to describe the dynamic evolution of social networks, with far reaching consequences for the study of social networks and social evolution.



(a) (b)

FIG. 1. **Similarity matrices and school schedules.** Matrices of similarities between the evolving networks built from one day of data collected in a French elementary school (a) and a US middle school (b). Here $\alpha = \beta = 0.1$, and the evolving networks are observed every $\Delta = 20$ minutes for the French school and every $\Delta = 5$ minutes for the US school. The horizontal bars give information about the schedule of a school day. The different colors in the bar correspond to the different class times (indicated by the letter C in (a) and with different numbers in (b)) and lunchtimes (indicated by the letter L), the length of each colored interval representing the duration of the corresponding period. In (b) there are two bars because the students were split into two groups for their lunchtime and fourth class period and therefore have slightly different schedules.

Iacopo Iacopini (iacopiniiacopo@gmail.com)

Vanishing critical mass for tipping points in social conventions

I. Iacopini[1], G. Petri[2], A. Baronchelli[3], A. Barrat[1]

[1]Aix Marseille Univ, Université de Toulon, CNRS, CPT, Marseille, France
[2]Mathematics and Complex Systems Research Area, ISI Foundation, Turin, Italy
[3]Department of Mathematics, City, University of London, EC1V 0HB, London, United Kingdom

How can minorities of regular individuals overturn social conventions? The theory of *critical mass* argues that apparently stable social conventions can be overturned by a minority of committed individuals if such minority reaches a critical size[1]. Several studies have focused on identifying this critical size. Qualitative studies of gender conventions in corporate leadership roles have hypothesised that a critical mass of 30% of the population is necessary in order for the tipping point to be reached[2]. Related observational work has proposed a higher critical mass size approaching 40% of the population[3]. Quantitative support to the tipping points dynamics has come from the naming game model for social convention[4], where a critical mass of 10% was shown to be able to induce norm change[5]. Subsequent generalisations showed how the threshold can vary between 10% and 40% of the population. Finally, empirical evidence for tipping points in social conventions showed that the theoretically predicted dynamics of critical mass do in fact emerge as expected within an empirical system of social coordination, finding a critical threshold of 25% of the population[6].

Adhering to the most common formulation of the theory, all theoretical models and empirical studies existing so far have considered minority groups of fully committed individuals, i.e., individuals who will not adopt majoritarian conventions under any circumstances. In this light, even 10% of the population, the smallest threshold obtained thus far, can hardly be considered a “small” group. This raises the classical problem of explaining where the critical mass itself comes from. Numerous observations suggest that even groups counting just tens of committed individuals, and not significant fractions of the population, may trigger abrupt social and normative change. Social movements offer several examples in this sense[7;8]. In this context, the problem is not explaining how a large critical mass is created in the first place, but rather clarifying whether the critical size of a committed group can be much smaller than what the current theories and models predict.

Here[9], we show that the critical mass is dramatically reduced when non-committed members of the population are–counter-intuitively–less prone to social influence than previously assumed. We adopt the naming game framework, already used to determine the range of 10% to 40% critical mass thresholds, and admit the possibility that individuals may be reluctant to let go of alternative conventions even when they successfully manage to coordinate with one another on a specific norm. We control this through a global communication efficiency parameter $\beta \in [0, 1]$. This generalises the standard naming game model where a successful coordination is followed by a certain and exclusive adoption of the norm that allowed the coordination[4;10]. We inform the model with real-world data concerning the structure of the social networks and of the microscopic interactions, considering both pairwise and group meetings[11]. The example of a temporal evolution leading to a minority takeover is shown in Fig. 1. We show that results hold in a significant region of the parameters and on all the considered real and artificial social structures. In addition, we clarify the role of group interactions showing a non-trivial dependency of the long-term dynamical output with the group size (non-monotonicity is observed). Our findings reconcile the numerous observational accounts of rapid change in social conventions triggered by committed minorities with the apparent difficulty of establishing such large minorities.

[1] M. Granovetter, Am. J. Sociol. **83**, 1420 (1978).
[2] D. Dahlerup, Scandinavian Political Studies **11**, 275 (1988).
[3] S. Grey, Politics & Gender **2**, 492 (2006).
[4] A. Baronchelli, M. Felici, V. Loreto, E. Caglioti, and L. Steels, J. Stat. Mech. Theory Exp. **2006**, P06014 (2006).
[5] J. Xie, S. Sreenivasan, G. Korniss, W. Zhang, C. Lim, and B. K. Szymanski, Phys. Rev. E **84**, 011130 (2011).
[6] D. Centola, J. Becker, D. Brackbill, and A. Baronchelli, Science **360**, 1116 (2018).
[7] M. Gladwell, The New Yorker **4**, 42 (2010).
[8] E. Chenoweth, M. J. Stephan, and M. Stephan, *Why civil resistance works: The strategic logic of nonviolent conflict* (Columbia University Press, 2011).
[9] I. Iacopini, G. Petri, A. Baronchelli, and A. Barrat, arXiv preprint arXiv:2103.10411 (2021).
[10] A. Baronchelli, Belgian J. of Linguistics **30**, 171 (2016).
[11] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, Phys. Rep. **874**, 1 (2020).

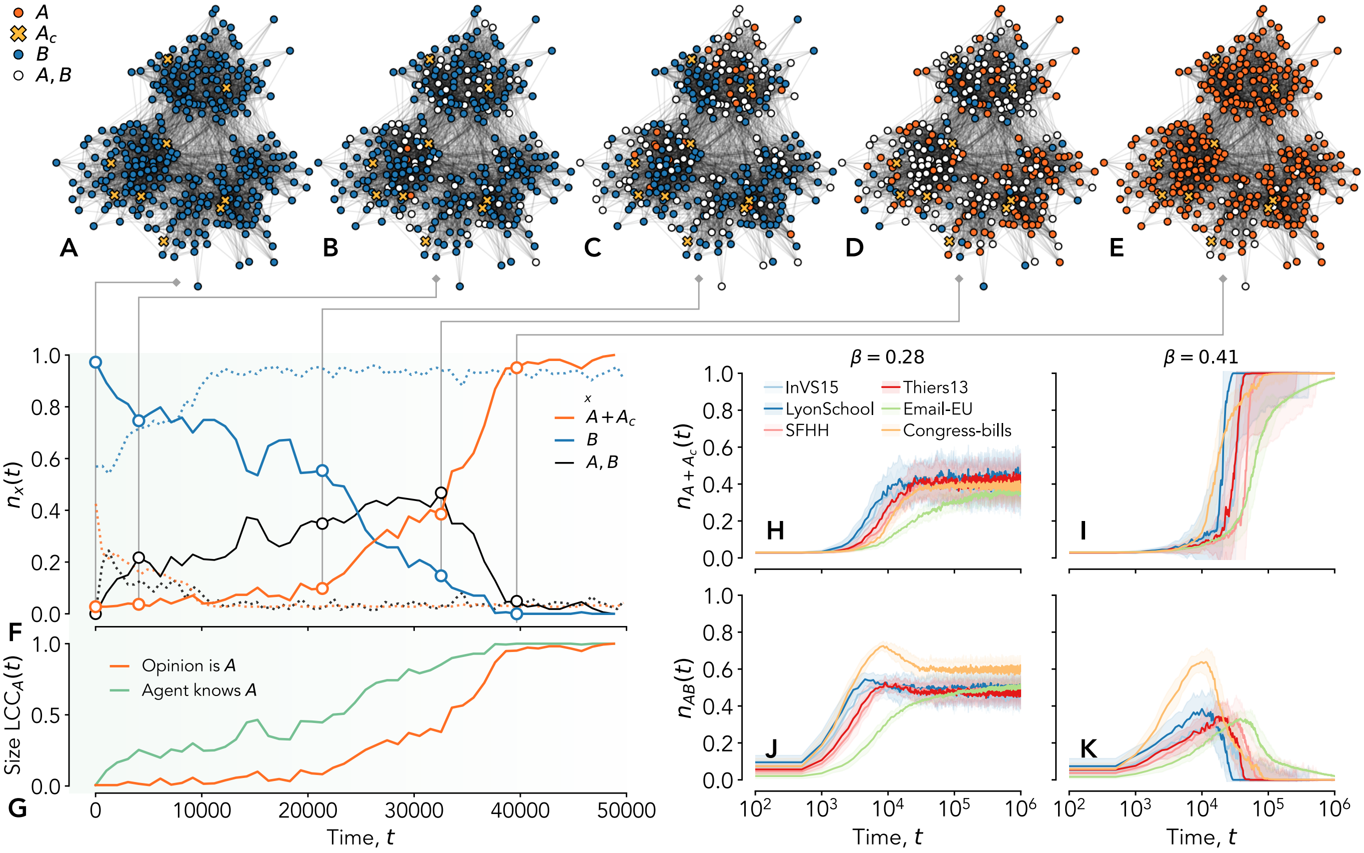


Figure 1: (*A*-*G*) Illustrative example of a simulation on an empirical social structure (*Thiers13*) where a minority consisting of $p = 3\%$ of the population (supporting A and denoted as A_c) overturns the stable social norms and reaches global consensus (with a communication efficiency set to $\beta = 0.41$). (*F*) Temporal evolution of the density $n_x(t)$ of nodes holding opinion x. Different solid lines correspond to different opinion values $x = \{A + A_c, B, (A, B)\}$. Dotted lines are reported as a benchmark, representing a naming game without committed minorities and with 40% of the initial population supporting opinion A ($p = 0$, $n_A(t = 0) = 0.4$). (*G*) Temporal evolution of the normalised size of the largest connected component (LCC) of nodes holding opinion A (red curve) and nodes that have A (but not necessarily A only) in their vocabulary (green curve). (*H*-*K*) Temporal evolution of the dynamics with committed minorities on empirical higher-order structures. The social structures (hypergraphs) are constructed from empirical data sets collected in six different context: workplace (*InVS15*), a primary school (*LyonSchool*), a conference (*SFHH*), a high school (*Thiers13*), email communications (*Email-EU*) and a political congress (*Congress-bills*). The temporal evolution of the density of nodes holding opinion A and A, B (with 3% of committed individuals) is reported in panels (*H*,*I*) and (*J*,*K*), respectively, for two different values of the communication efficiency, namely $\beta = 0.28$ (*H*,*J*) and $\beta = 0.28$ (*I*,*K*). The results averaged over different runs of stochastic simulations are reported as solid curves and shaded areas, representing median values and standard deviations.

Capturing the Diversity of Multilingual Societies

T. Louf[1], *D. Sánchez*[1], *J.J. Ramasco*[1]. *(1) Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (UIB-CSIC), Palma, Spain, thomaslouf@ifisc.uib-csic.es.*

Cultural diversity encoded within languages of the world is at risk, as many languages have become endangered in the last decades in a context of growing globalization. To preserve this diversity, it is first necessary to understand what drives language extinction, and which mechanisms might enable coexistence. In this work [1], we consider the processes underlying language shift through a conjunction of theoretical and empirical perspectives. A large-scale empirical study of spatial patterns of languages in multilingual societies using Twitter and census data yields a wide diversity. It ranges from an almost complete mixing of language speakers, including multilinguals, to segregation with a neat separation of the linguistic domains and with multilinguals mainly at their boundaries. To understand how these different states can emerge and, especially, become stable, we propose an agent-based model in which coexistence of languages may be reached when learning the other language is facilitated and when bilinguals favor the use of the endangered language. A mean-field analysis for a single population uncovers interesting stable states of extinction and coexistence, including with bilinguals alone sustaining a minority language. As shown in Figure 1, simulations carried out in a metapopulation framework highlight the importance of spatial interactions arising from population mobility to explain the stability of a mixed state (Fig. 1C) or the presence of a boundary between two linguistic regions (Fig. 1B). Changes in the parameters regulating the relation between the languages can destabilize a system, which undergoes global transitions. According to our model, the evolution of the system once it undergoes a transition is highly history-dependent. It is easy to change the status quo but going back to a previous state may not be simple or even possible.

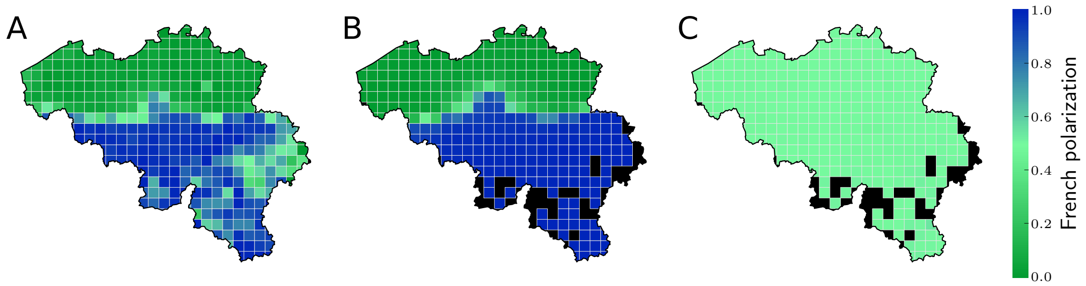


Figure 1: Polarization of French in Belgium, (A) from Twitter data, and stable states reached with our model (B) for a relatively low and (C) high ease to learn the other language. Polarization equals 1 when only French is spoken, and 0 when only Dutch is.

References

[1] T. Louf, D. Sanchez and J.J. Ramasco (2021). "Capturing the diversity of multilingual societies". arXiv:2105.02570

The evolution of moral rules in a model of indirect reciprocity with private assessment

C. Perret[1,2]*, M. Krellner*[2]*, T.A. Han*[1,2]*. (1) University of Exeter, Penryn TR10 9FE, United Kingdom, cedric.perret.research@gmail.com (2) Teeside University, Southfield Rd, Middlesbrough TS1 3BX, United Kingdom*

Humans follow moral rules which dictate how they judge others' actions. Moral rules have been shown to foster cooperation by indirect reciprocity, that is cooperators are rewarded by receiving future cooperation from third party [1]. Yet, it is not clear which moral rules best implement indirect reciprocity, and which should be favoured by natural selection. Do individuals use only a few rules among all the possible rules, or do a wide variety of rules coexist? Are the most successful rules complex, or are they simple as observed in direct reciprocity?

Previous studies either considered only public assessment, that individuals are considered good or bad by all others in the same manner [2], or compared a subset of of all possible rules or strategies [3]. Here we present a recently submitted work [4] in which we develop a method to determine the reputation of individuals and the frequency of cooperation in the long run for any strategies up to second-order assessment rules. We use this method to determine analytically if a strategy as mutant can invade another strategy as resident, for every pair of mutant and resident strategies.

Our results identify which rules are evolutionary stable strategies (ESS), that are moral rules that remain once adopted by the whole population. Our findings show that successful rules share few crucial features, but also that a diversity of rules can be observed. We also show that the presence of errors breaks down the evolutionary stability of these ESS strategies. Ultimately, this model contributes to better understand how morality emerges and its role in the evolution of cooperation.

Acknowledgements

C.P. and T.A.H. were supported by Future of Life Institute (grant RFP2-154). T.A.H. is also supported by Leverhulme Research Fellowship (RF-2020-603/9).

References

[1] R. D. Alexander, *The biology of moral systems*. Somerset, USA: Taylor & Francis Inc, 1987.

[2] H. Ohtsuki and Y. Iwasa, "How should we define goodness? - Reputation dynamics in indirect reciprocity," *J. Theor. Biol.*, vol. 231, no. 1, pp. 107–120, 2004, doi: 10.1016/j.jtbi.2004.06.005.

[3] I. Okada, "Two ways to overcome the three social dilemmas of indirect reciprocity," *Sci. Rep.*, no. 0123456789, pp. 1–9, 2020, doi: 10.1038/s41598-020-73564-5.

[4] C. Perret, M. Krellner, and T. A. Han, "The evolution of moral rules in a model of indirect reciprocity with private assessment," *Submitted*, 2021.

A minimalistic model of bias, polarization and misinformation in social networks

Orowa Sikder,[1] Robert E. Smith,[1] Pierpaolo Vivo,[2] and Giacomo Livan[1]
[1] *University College London, Department of Computer Science, 66-72 Gower Street, WC1A 6EA London, UK*
[2] *Department of Mathematics, King's College London, Strand, WC2R 2LS London, UK*

Online social networks provide users with unprecedented opportunities to engage with diverse opinions. At the same time, they enable confirmation bias on global scales, both directly by empowering individuals to self-select narratives they want to be exposed to, and indirectly through algorithmic filter bubbles. A precise theoretical understanding of such tradeoffs is still largely missing. We introduce a social learning model where agents in a network receive and exchange information about a binary event (e.g., a ground truth vs untruth). The majority of agents update their beliefs unbiasedly (via Bayes' rule) when they receive information, whereas a minority of participants display confirmation bias and therefore reject – with a certain probability – new information that is in contrast with their current beliefs.

We provide a full analytical solution of the model, and find that the interplay of the two above mechanisms results in very rich dynamics. Our main findings (published in Sikder et al., *Scientific Reports* 10, 5493) are:

1. In the long-run, the population of agents converges towards one of three possible steady states: consensus on the ground truth, consensus on the untruth, or permanent polarisation. The long-run steady state is *entirely* determined by two parameters that characterise the sub-population of biased agents (see Fig. 1); this result is extremely robust with respect to changes in the underlying network's topology.

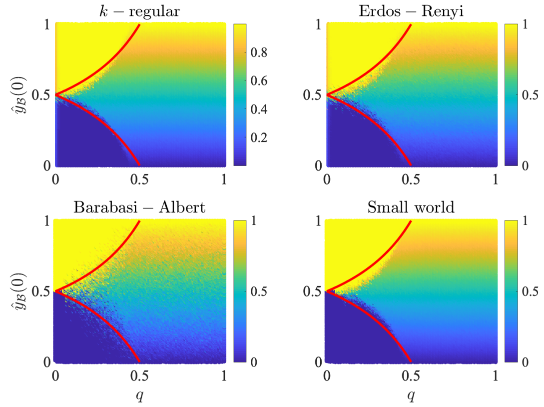


FIG. 1: Simulation results on different network topologies for the long-run steady state of the population of agents. Yellow (blue) denotes consensus on the ground truth (untruth), intermediate colours denote polarized states. Red lines denote analytical predictions for the transitions between different regimes. The q parameter on the x-axis quantifies the extent of confirmation bias in the population, while the parameter on the y-axis quantifies the average belief of the biased population at the beginning of time.

2. The population's accuracy (quantified as the fraction of agents whose beliefs are aligned with the ground truth in the steady state) is a non-monotonic function of the fraction of biased agents. This is due to the fact that confirmation bias functions as a double-edged sword, as it both "protects" the accurate beliefs of those whose opinions are aligned with the ground truth and prevents those with inaccurate beliefs from changing their mind. The balance between those two competing forces results in an optimal amount of confirmation bias in a society, which maximises its overall accuracy.

3. The model displays a clear tradeoff between bias and connectivity. We demonstrate that low connectivity effectively results in clusters of unbiased agents ultimately ending up "surrounded" by biased agents, a phenomenon reminiscent of the formation of echo chambers in online social media. Conversely, increases in connectivity break up echo chambers and allow a better diffusion of information throughout the network. We empirically validate this result against US county-level data about beliefs on global warming.

Our model only provides a stylised representation of real-world complexity. Yet, its value lies in its ability to generate testable predictions and in the clarity of its assumptions, which ultimately allow to assign a precise mathematical meaning to blanket terms that are often adopted in the debate around misinformation and opinion polarisation.

What science can do for democracy – A complexity science approach

K. Wiesner*[1] *et al. (1) Institute of Physics, University of Potsdam, Germany. karoline.wiesner@uni-potsdam.de.

The idea that democracy is under threat, after being largely dormant for at least 40 years, is looming increasingly large in public discourse. What makes a democracy stable? And which processes potentially lead to instability of a democratic system? We offer a complex systems perspective on this question, informed by areas of the mathematical, natural, and social sciences. We explain the meaning of the term 'stability' in different disciplines and discuss how laws, rules, and regulations, but also norms, conventions, and expectations are decisive for the stability of a social institution such as democracy.

Political scientists have conventionally assumed that achieving democracy is a one-way ratchet. Only very recently has the question of “democratic backsliding” attracted any research attention. We argue that democratic instability is best understood with tools from complexity science. The explanatory power of complexity science arises from several features of complex systems. Their relevance in the context of democracy is discussed. Several policy recommendations are offered to help (re)stabilize current systems of representative democracy.

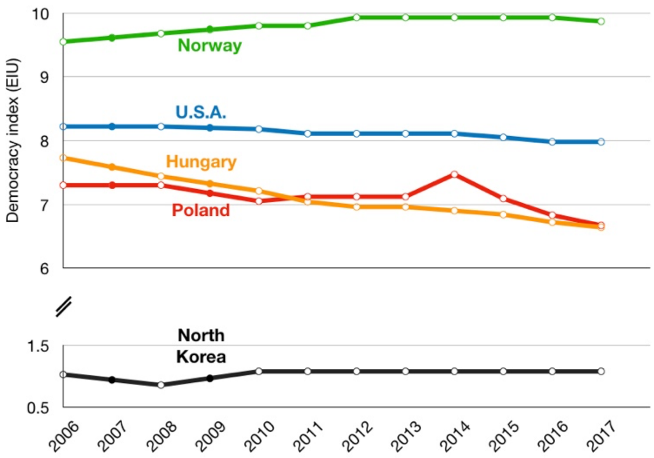


Figure 1: Democracy index for selected countries, as measured by the Economist Intelligence Unit for 2006–2017. The values for 2007 and 2009 are linear interpolations as no indices were reported for these two years

Acknowledgements

This work was carried out in collaboration with Tina Eliassi-Rad, Henry Farrell, David Garcia, Stephan Lewandowsky, Patricia Palacios, Don Ross, Didier Sornette, Karim Thébault.

References

[1] Wiesner, Karoline, et al. "Stability of democracies: a complex systems perspective." *European Journal of Physics* 40.1 (2018): 014002.

[2] Eliassi-Rad, Tina, et al. "What science can do for democracy: a complexity science approach." *Humanities and Social Sciences Communications* 7.1 (2020): 1-4.

“We follow the science”: Studying information dynamics across the science-policy interface

N.Li[1], *M. Luczak-Roesch*[1,2], *F. Donadelli*[3]. *(1) School of Information management, Victoria University of Wellington, New Zealand, nancy.li@vuw.ac.nz (2) Te Pūnaha Matatini – New Zealand’s Centre of Research Excellence for Complex Systems & Networks (3) School of Government, Victoria University of Wellington, New Zealand*

Evidence-informed policymaking has seen wide adoption by governments around the world in response to the ongoing global crisis caused by the COVID-19 pandemic. Yet, little is known about what it actually means when various politicians state “we follow the science”. Our work looks at the science-policy interface through the lens of a complexity science framework. Based on contemporary political science theory [1] we conceptualise the science-policy interface as multiple conversational spheres in which various actors (e.g., scientists, politicians, general public) produce, re-produce and disseminate (scientific) arguments that then may diffuse through the different spheres. To understand what promotes or prevents scientific arguments from diffusing we pay specific attention (via modelling and analysis) to argumentative coalitions that may amplify or dampen the diffusion processes.

We performed a series of initial experiments to understand the linguistic and temporal alignment between two of the aforementioned spheres, namely scientific discourse and parliamentary debates. By using natural language processing techniques we extracted key terms and their co-occurrence relationship from a focused bibliographic corpus retrieved from the Web of Science database, and from the Hansard transcripts of debates held at New Zealand’s parliament since the beginning of the COVID-19 pandemic. We find five distinct topical clusters in the set of scientific articles published related to COVID-19 that feature a distinct temporal attention shift (see Figure 1). Comparing the linguistic and temporal profile of the scientific sphere with the one we created for the parliamentary debates shows that both language and attention are neither trivially aligned across spheres nor stable across time. Our experiments provide essential insights for the development of robust methods to link the distinct spheres of the science-policy interface, and to model diffusion processes with not perfectly aligned and morphing information.

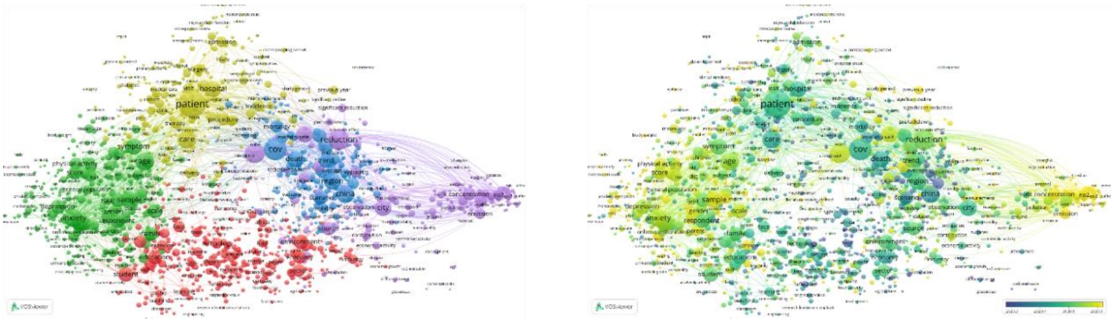


Figure 1: Term co-occurrence network of scientific publications, coloured by cluster (left) and time of publication (right)

References

[1] Nutley, Sandra M., Isabel Walter, and Huw TO Davies. *Using evidence: How research can inform public services.* Policy press, 2007.

Biased voter model: How persuasive a small group can be?

A. Czaplicka[1,2]*,* ***C. Charalambous***[2]*, R. Toral*[2]*, M. San Miguel*[2]*. (1) Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza 50009, Spain (2) IFISC, Instituto de Fisica interdisciplinar y Sistemas Complejos (CSIC-UIB), Campus Universitat de les Illes Balears, 07122, Palma de Mallorca, Spain, christos@ifisc.uib-csic.es.*

We study the voter model dynamics in the presence of confidence and bias. We assume two types of voters. Unbiased voters (UV) whose confidence is indifferent to the state of the voter and biased voters (BV) whose confidence is biased towards a common fixed preferred state. We study the problem analytically on the complete graph using mean field theory and on an Erdos-Renyi (ER) random network topology using the pair approximation, where we assume that the network topology is independent of the type of voters. We verify our analytical results through numerical simulations. We find that for the case of a random initial setup, and for sufficiently large number of voters N, the time to consensus increases proportionally to log(N)/γv, with γ the fraction of biased voters and v the bias of the voters. Finally, we study this model on a biased-dependent topology. We examine two distinct, global average-degree preserving strategies (model I (MI) and model II (MII) in Fig.1(b)) to obtain such biased-dependent random topologies starting from the biased-independent random topology case as the initial setup. In M1 we find that if we simply vary μ_{BU} , the average number of links among the two types of voters (BV-UV), no significant effect was observed (blue line Fig.1). Instead, in MII, increasing μ_{BB}, the average number of links among only biased voters (BV-BV) at the expense of μ_{UU} that of only unbiased voters (UV-UV), while keeping μ_{BU} constant, resulted in a significant decrease in the average time to consensus in the group (red line Fig.1). Hence, persuasiveness of the biased group depends on how well its members are connected among each other, compared to how well the members of the unbiased group are connected among each other.

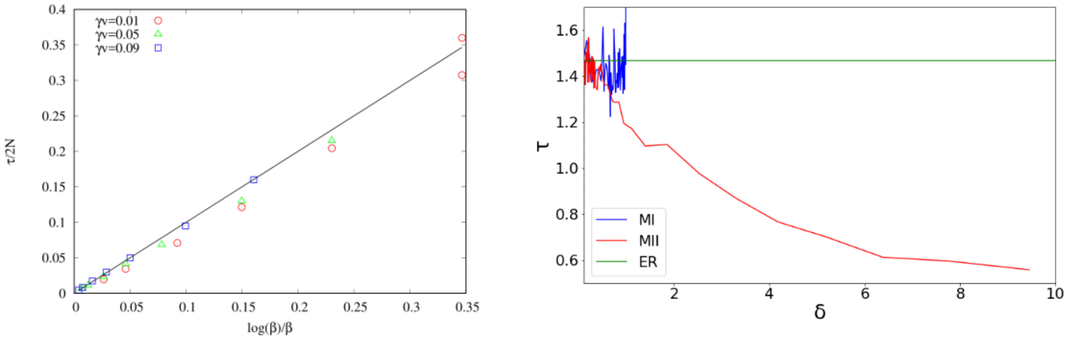


Figure 1: (a) we see the logarithmic scaling of the consensus time as a function of the effective bias $\beta = \gamma v N$ for the complete graph topology (analytical vs numerical results). (b) we plot the consensus time as a function of $\delta = \mu_{BB}/\mu_{UU}$ with μ_{XX} the average degree among only type X voters, for the case of a biased-dependent ER topologies (blue and red lines) vs the biased-independent ER topology case (green line).

Acknowledgements

This study was financially supported by the Spanish State Research Agency, through the Severo Ochoa and Marıa de Maeztu Program for Centers and Units of Excellence in R&D (MDM-2017-0711).

Attention Ecology

C. Pilgrim[1], *W. Guo*[2,3], *T. Hills*[1,2]. *(1) The University of Warwick, UK*
charlie.pilgrim@warwick.ac.uk. (2) The Alan Turing Institute, UK (3) Cranfield University, UK

Modern media competes for human attention in what has been termed the attention economy. Humans direct their attention in order to gather information in a similar way to how animals forage for food, and food foraging models have been used to describe information foraging in humans[1]. We develop this model to describe how people choose between information types and different media.

In the animal kingdom prey often adapt to keep themselves off the menu by making themselves harder to catch or digest. With information we expect the opposite - with media adapting to make itself more accessible and easier to consume. Given our information foraging model we ask how media must adapt to compete for our limited attention in a world with increasingly powerful communication systems and high information prevalence. We assume that media competes by providing higher information utility rates, which we can partly measure with the proxy of word entropy. With this in mind we compare our model to empirical results of language change over time and between media categories.

References

[1] P. Pirolli, S. Card, Information Foraging. *Psychological review*, 106(4):643, 1999.

Quantifying Knowledge Spillovers from *Negative Emissions* Research within and beyond Science

G.Tripodi[1,2], F. Lamperti[2], A. Mina[2], F. Chiaromonte[2,3], F. Lillo[1,4], R. Mavilia[5]. (1) Scuola Normale Superiore, Pisa, Italy, giorgio.tripodi@sns.it, (2) Institute of Economics and EMbeDS, Scuola Superiore Sant'Anna, Pisa, Italy, (3) Department of Statistics, The Pennsylvania State University, University Park, PA, USA (4) Department of Mathematics, University of Bologna, Bologna, Italy, (5) ICRIOS - Bocconi University, Milan, Italy.

According to a large majority of Integrated Assessment Models (IAMs), negative emission technologies (NETs) represent a pivotal element to meet the requirements of the Paris Agreement, and thus to tackle global warming. However, as of today, they are hardly fully developed technologies ready for large-scale deployment as soon as many models assume [1]. Different NETs have been evaluated according to three main dimensions: negative emissions potential, energy and natural resources requirements and economic costs. Such estimates vary significantly and each of the proposed options comes with unique features and challenges [2]. This work takes an innovation networks perspective with the aim to quantify knowledge spillovers generated from different NETs within and beyond the scientific realm, taking into account marketplace inventions. By analyzing 20 years of academic literature via citation networks and regression techniques, we find that (i) knowledge spillovers play a non-negligible role in NETs research, accounting for about 50% of knowledge flows (Figure 1), (ii) the impact within the scientific realm varies greatly among NETs, with Direct Air Capture (DAC), Blue carbon and Biochar being particular impactful in terms of both citations and scope, and (iii) DAC is substantially more linked to marketplace inventions (measured as citations from patents). To limit possible sources of bias, we construct two control groups (general and climate specific) by matching similar articles (same year/same journal). In addition, we identify cities and countries that can serve as hubs for supporting collaborations. Innovation might play an essential role in dealing with the climate change crisis, but we might need to move beyond the conventional innovation policy toolkit.

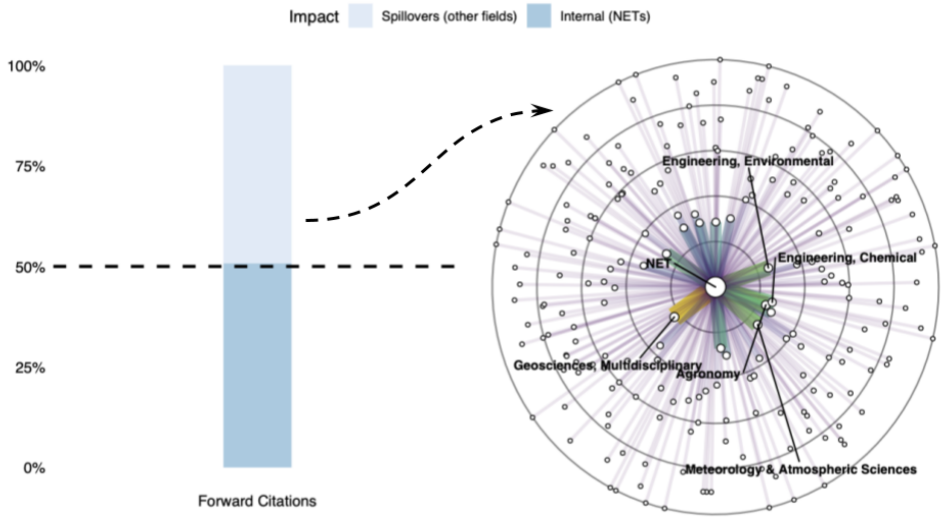


Figure 1: Knowledge spillovers generated by NETs research account for about 50% of total forward citations (bar), spanning several different fields from agronomy to chemical engineering (Radar).

References

[1] Anderson, K. & Peters, G. *Science* 354, 182–183 (2016).

[2] Smith, P. et al. *Nat. climate change* 6, 42–50 (2016).

Analytical Derivations of Critical Tolerance Thresholds in the Schelling Model

__A. Lucquiaud__[1,2], J. Randon-Furling[1,3,4]. (1) SAMM, Université Paris 1 Panthéon-Sorbonne, Paris, France. lucquiaud@crans.org (2) Modal'X, Université Paris Nanterre, Nanterre, France (3) Department of Mathematics, Columbia University, New York, USA (4) MSDA, UM6P, Ben Guerir, Morocco

Schelling's is one of the oldest and most famous agent-based models in the social sciences [1]. It reveals the emergence of macroscopic segregation in a system of agents belonging to two groups and interacting at a microscopic level on a 2D lattice. The system's behaviour has been studied extensively through numerical simulations, leading to phase diagrams [2,3] that show which stationary macrostate (mixed or segregated) is reached depending on the values of a handful of parameters. The most important one is the so-called "tolerance" T: an agent will move to another location if a fraction strictly larger than T of their neighbours belongs to the other group.

Consistent with Schelling's counter-intuitive observations, a first phase transition – from low-tolerance mixed equilibria to segregated macrostate – occurs around T=1/2 when vacant sites are scarce, or when unsatisfied agents are allowed to trade positions, in a Kawasaki-type exchange process. We derive this transition in a dynamical setup, using Markov chain aggregation techniques [4]: to some extent, the system's behaviour resembles that of an urn from which pairs of agents would be drawn and swapped. This analogy enables us to compute the transition matrix and invariant distribution, which we compare to the empirical study of attractor states [5].

Since Schelling's seminal work, the existence of a second phase transition has also been confirmed at a value T_c of the tolerance with $1/2 \leq T_c \leq 3/4$ – the exact value varies with the local neighbourhood size [6]. We build upon previous works [7,8] that focus on the formation of homogeneous clusters and use asymptotic results of percolation theory in infinite systems with infinitely large neighbourhoods. In a simpler, finite setup, mimicking nucleation models from statistical physics, we derive analytically the value of T_c from simple conditions on the stability of small homogeneous clusters, that act as nuclei in the transition toward segregation. For $1/2 < T < T_c$, the stationary state is one of separation between homogeneous clusters, while for $T > T_c$ it is one of integration, as observed on the wealth of existing simulation results.

Acknowledgements

This research has been conducted within the FP2M federation (CNRS FR 2036).

References

[1] T.C. Schelling, *Journal of Mathematical Sociology*, 1 (1971) 143–186.
[2] L. Gauvin, J. Vannimenus, J.P. Nadal, *European Physical J B*, 70 (2009) 293–304.
[3] N. Goles Domic, E. Goles, S. Rica, *Physical Review E*, 83 (2011) 056111.
[4] S. Banisch, *Markov chain aggregation for agent-based models,* Springer (2015).
[5] V. Cortez, P. Medina, E. Goles, R. Zarama, S. Rica, *European Physical J B*, 88 (2015) 25.
[6] A.J. Laurie, N.K. Jaggi, *Urban Studies*, 40.13 (2003), 2687–2704.
[7] G. Barmpalias, R. Elwes, A. Lewis-Pye, *Nonlinearity*, 31 (2018) 1593.
[8] H. Omidvar, M. Franceschetti, *Journal of Statistical Physics*, 170 (2018) 748–783.

Investigating the changing structure of research collaborations between institutions in Latin America & the Caribbean using multilayered networks

J. Fitzgerald[1]. *(1) Mathematical Institute, University of Oxford, Radcliffe Observatory, Andrew Wiles Building, Woodstock Rd, Oxford, OX2 6GG, UK, fitzgeraldj@maths.ox.ac.uk.*

Which disciplines have similar or dissimilar patterns of collaboration, and to what extent are the patterns for publications in languages other than English different? Further, how do these factors impact the research communities that are formed? Using Scopus data on publications between 1996-2018 with authors affiliated to an institution based in Latin America or the Caribbean, we apply a multilayer hierarchical stochastic blockmodel (SBM) approach first outlined in [1] to the institutional co-authorship network to investigate these questions. Layers correspond to combinations of disciplines and languages. In total, we consider around 1.4 million unique publications across 822 distinct institutions.

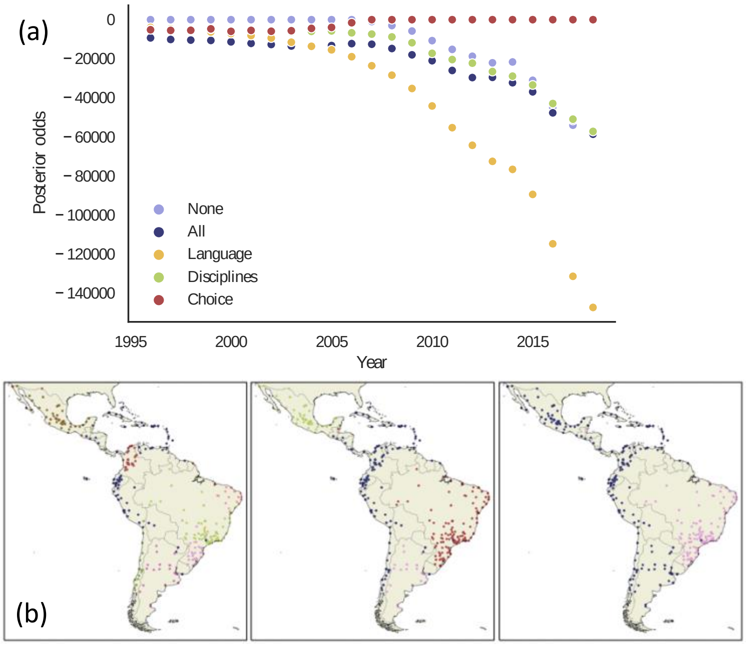


Figure 1: Subfigure (a) displays the log posterior odds ratio between the best layer structure and each other for each year. In subfigure (b) we display research institutions coloured by community for each of the three non-trivial levels of the hierarchical partition of our best model in 2018, ascending the hierarchy from left to right.

By evaluating posterior odds ratios between layered models inferred for a variety of layer structures (discipline-language combinations), we can ascertain whether the pattern of collaborations between institutions is sufficiently different when disaggregated by these categories to justify the additional model complexity. If so, communities detected by models applied to collapsed datasets only see part of the picture, as they do not account for this further information.

We find that publication discipline and language play an important structural role in patterns of collaborations between institutions, and importantly that the two are interrelated, in that publication language matters more in some disciplines (and institutions) than others. We also observe strong geographic organisation, that displays how borders still play a fundamental role in research collaboration today, even as many herald the ‘death of distance’.

Acknowledgements

This work uses Scopus data provided by Elsevier through the International Center for the Study of Research (ICSR) Lab. This publication is based on work partially supported by the EPSRC Centre For Doctoral Training in Industrially Focused Mathematical Modelling (EP/L015803/1) in collaboration with Elsevier.

References

[1] Peixoto, T. P. (2015). Inferring the mesoscale structure of layered, edge-valued, and time-varying networks. *Phys Rev E*, *92*(4), 042807.

Trust-Distrust Model on Multiple Network of Society as New Type of Opinion Dynamics

A. Ishii[1,2], *N. Okano*[1]. *(1)Tottori University, Koyama, Tottori 680-8552, Japan, ishii.akira.t@gmail.com. (2) Center for Computational Social Science, Kobe University, Kobe, Japan*

The connections between people in a society can be seen as a network. However, it is not a single network, but a multiplicity of networks, such as social network connections, neighborhood connections, and workplace connections. In this study, we extend the new opinion dynamics named Trust-Distrust Model, proposed by Ishii et al. in 2018, to multiple networks and apply it to various social phenomena[1,2,3]. The Trust-Distrust Model differs from the conventional Bounded Confidence Model in that it introduces distrust in addition to trust among people in society.

If the networks connecting people in a society are multiple, the Trust-Distrust Model can be extended as follows.

$$\varDelta I_i(t) = c_i A(t)\varDelta t + \sum_{\alpha}\sum_{j=1}^{N} D_{ij}^{(\alpha)}\varPhi(I_i(t), I_j(t))(I_j(t) - I_i(t))\varDelta t$$

The matrix of trust coefficients connecting people, $D_{ij}^{(\alpha)}$, is the sum of multiple networks, and a negative trust coefficient means distrust. $I_i(t)$ is the opinion of an agent “i” at time t. As can be seen from this equation, this opinion dynamics equation is linear with respect to the matrix $D_{ij}^{(\alpha)}$ of coefficients representing the connections of people in a multiplex network, and the multiplex network can be easily calculated as the sum of this matrix.

The Trust-Distrust Model has also been used to simulate social divisions and presidential elections [3], which can also be understood with this multiple network model. Here, as an example of the multiple network model, we have dealt with two synthetic networks, the network of the whole society and the network of neighbors. In the network of neighbors, we assumed that each agent constitutes a complete network with 20 surrounding people, and investigated the conditions for consensus formation. The consensus formation is much stricter in the neighbors' network than in the society-wide network.

Acknowledgements

This work was supported by JSPS KAKENHI Grant Number JP19K04881..

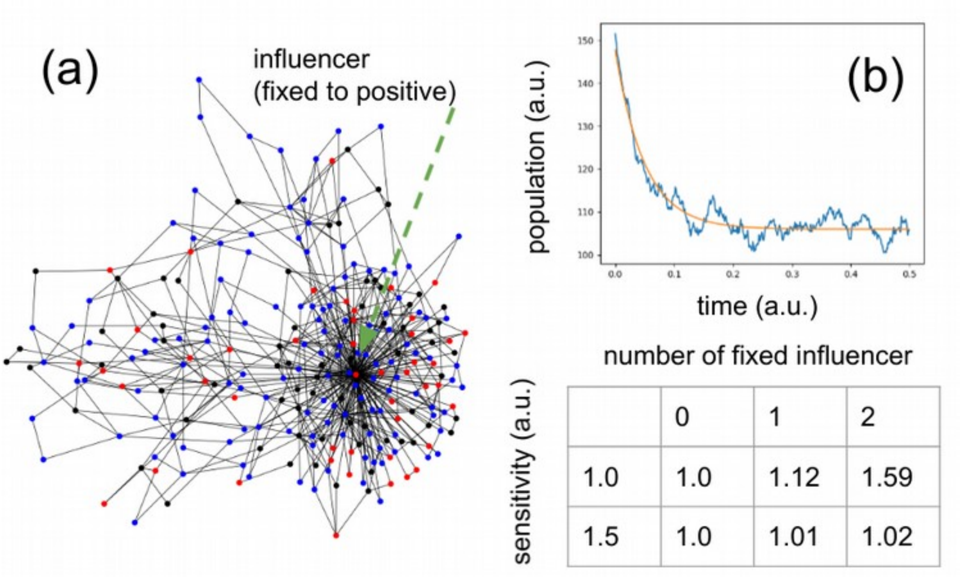
References

[1] Ishii, A. and Kawahata, Y. : Opinion Dynamics Theory for Analysis of Consensus Formation and Division of Opinion on the Internet. : Proceedings of The 22nd Asia Pacific Symposium on Intelligent and Evolutionary Systems (IES2018), 71-76; arXiv:1812.11845 [physics.soc-ph] (2018)
[2] Ishii A. : Opinion Dynamics Theory Considering Trust and Suspicion in Human Relations. In: Morais D., Carreras A., de Almeida A., Vetschera R. (eds) Group Decision and Negotiation: Behavior, Models, and Support. GDN 2019. Lecture Notes in Business Information Processing, vol 351. Springer, Cham (2019).
[3] Ishii A, Okano N and Nishikawa M, Social Simulation of Intergroup Conflicts Using a New Model of Opinion Dynamics. Front. Phys. 9:640925, (2021).
doi: 10.3389/fphy.2021.640925

Ising Model Simulation of Epidemic Control by Influencers on Social Networks

Ryohto Sawada *Preferred Networks (attend as indivisual), Otemachi Bldg., 1-6-1 Otemachi, Chiyoda-kuokyo, Tokyo, Japan, shi9_lo_knight@yahoo.co.jp*

Controlling the relaxation process of the epidemic in social networks is an important issue to prevent the spread of fake news. Recent studies suggest that the relaxation process can be expressed as a Poisson process where the views of the information decay almost exponentially [1]. Additionally, information exchange mainly occurs through personal connections such as social network service, where influencers, who advertise specific opinions, are thought to play a critical role [2]. In this study, I reproduce the relaxation process of the epidemic in the social networks by the relaxation of Ising models on an undirected graph and simulate the effect of influencers by fixing the state of the influencer nodes that have the most adjacent nodes. Figure 1(a) is an example of the Ising model. The state of each node represents the individual's opinion (red: positive, blue: negative, black: not interested) and the state of an influencer fixed to positive. Figure 1(b) shows the temporal variation of the number of nodes in the negative (blue) state where the number of nodes relaxes according to the exponential. The table under Figure 1(b) compares the relative relaxation time by changing the number of influencer nodes and coupling constant. One can see that the relaxation process can be accelerated by fixing the state of the influencer nodes. One can also see the smaller the coupling constant, in other words, the smaller susceptibility to the other people's opinions, causes the greater the relative influence of the influencers.



	0	1	2
1.0	1.0	1.12	1.59
1.5	1.0	1.01	1.02

Figure 1 (a) : An example of Ising model in a social network. (b) : Temporal variation of the number of nodes with a negative opinion. Table: Comparison of the relative relaxation by changing the number of influencer nodes and coupling constant

References

[1] Riley Crane and Didier Sornette, PNAS October 14, 2008 105 (41) 15649-15653

[2] Jan van Dijk, *The Network Society,* SAGE Publications Ltd; 2020

The Role of Chance in Fencing Tournaments: an Agent-Based Approach

***C. Zappalà*[1]*, A. E. Biondo*[2]*, A. Pluchino*[1]*, A. Rapisarda*[1,3]*, P. Sobkowicz*[4]*. (1) Department of Physics and Astronomy, University of Catania and INFN sezione di Catania, Via S. Sofia 64, 95123 Catania, Italy, chiara.zappala@phd.unict.it. (2) Department of Economy and Business, University of Catania, Corso Italia 55, 95129 Catania, Italy, (3) Complexity Science Hub Vienna, Josefstadter Strasse 39, 1080 Vienna, Austria, (4) National Centre for Nuclear Research, A Soltana 7, 05-400 Otwock, Polonia*

It is a widespread belief that success is mainly due to innate qualities, rather than to external forces. This is particularly true in sport competitions, where individual talent is considered the only ingredient in order to reach success. In this study, inspired by recent papers on this topic [1, 2], we explore the relative weight of talent and luck in fencing, a combat sport involving a weapon. Fencing competitions are structured as direct elimination tournaments, where randomness is explicitly present in some rules (for example, in case of tie). Our dataset covers the last decade of international events and consists of both single competition results and annual rankings for male and female fencers under 20 years old (Junior category). We first perform a statistical analysis of these data, provided by *fie.org*, looking for correlations between the initial ranking at the start of each event of a fixed season and the final placement at the end of that event. Then, through an agent-based model, we reproduce fencing tournament dynamics in a virtual environment, in a given year and for several seasons, with just one free parameter a which describes the importance of talent in competitions (a = 1 indicates the ideal scenario where only talent matters, a = 0 the complete random one). We find that simulations approximate very well the real data when talent weights slightly more than luck, i.e. when a is around 0.55. We conclude that the role of chance in fencing is heavily underestimated, although it creates a huge disparity between comparable talents and it consequently influences not only sports results but also future career development.

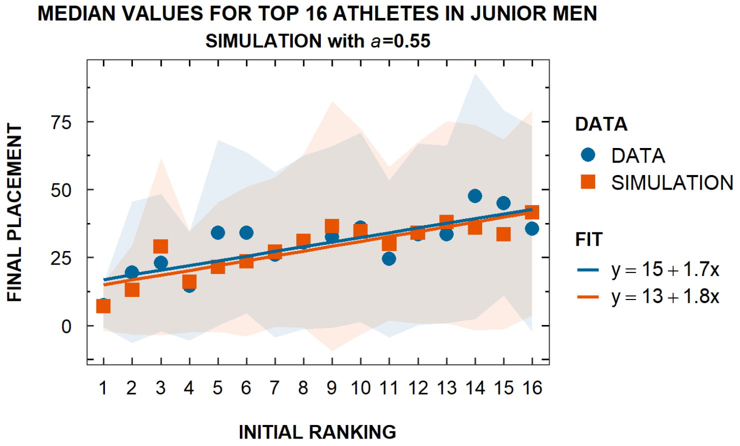


Figure 1: Comparison between single tournaments' data and simulations for the first sixteen male fencers, with talent importance a = 0.55.

References

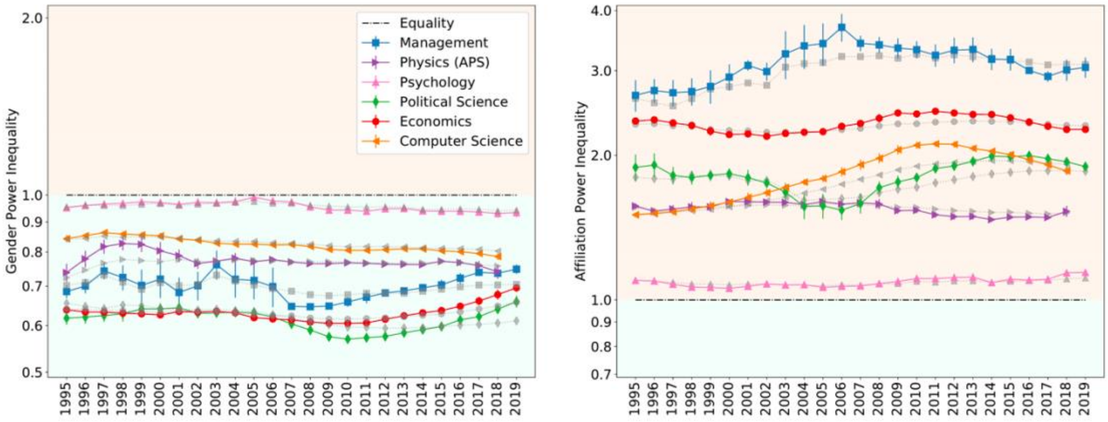
[1] A. Pluchino, A. E. Biondo, A. Rapisarda, *Talent vs Luck: the role of randomness in success and failure*, *Advances in Complex Systems*, 21 (2018) 03n04.
[2] P. Sobkowicz, R. H. Frank, A. E. Biondo, A. Pluchino, A. Rapisarda, *Inequalities, chance and success in sport competitions: Simulations vs empirical data, Physica A: Statistical Mechanics and its Applications*, 557 (2020) 124899.

Emergence of Structural Inequalities in Scientific Citation Networks

B. Nettasinghe[1], *N. Alipourfard*[2], *V. Krishnamurthy*[1], *K. Lerman*[2] *(1) Cornell University, Ithaca, USA, dwn26@cornell.edu. (2) USC Information Sciences Institute, Los Angeles, USA*

Structural inequalities persist in society, conferring systematic advantages to some people at the expense of others, for example, by giving them substantially more influence and opportunities. Using bibliometric data about authors of scientific publications, we identify two types of structural inequalities in scientific citations. First, female authors, who represent a minority of researchers, receive less recognition for their work (through citations) relative to male authors; second, authors affiliated with top-ranked institutions, who are also a minority, receive substantially more recognition compared to other authors. We present a model (*Directed Mixed Preferential Attachment model*) for the growth of directed citation networks and show that citations disparities arise from individual preferences to cite authors from the same group (homophily), highly cited or active authors (preferential attachment), as well as the size of the group and how frequently new authors join. We analyze the model and show that its predictions align well with real-world observations. Our theoretical and empirical analysis also suggests potential strategies to mitigate structural inequalities in science. In particular, we find that merely increasing the minority group size does little to narrow the disparities. Instead, reducing the homophily of each group, frequently adding new authors to a research field while providing them an accessible platform among existing, established authors, together with balanced group sizes can have the largest impact on reducing inequality.

Our work generalizes previous work which were limited to either a directed network with a homogeneous population or an undirected network with non-homogenous population. Thus, our work also highlights additional complexities of mitigating structural disparities stemming from asymmetric relations (e.g., directed citations) compared to symmetric relations (e.g., collaborations). More details about the proposed model and the results can be found in [1], [2].



(a) Gender Power-inequality **(b)** Affiliation Power-inequality

Figure 1: Power-inequality in (a) gender-partitioned and (b) affiliation-partitioned networks.

References (optional)

[1] B Nettasinghe, N Alipourfard, V Krishnamurthy, K. Lerman, *arXiv preprint arXiv:2103.12149* (2021)

[2] B Nettasinghe, N Alipourfard, V Krishnamurthy, K. Lerman, *arXiv preprint arXiv:2103.10944* (2021)

Heterogeneous Interactions in Artificial Intelligence Development Races

T. Cimpeanu[1], *T.A. Han*[1], *F. C. Santos*[2,5], *L.M. Pereira*[3], *T. Lenaerts*[4,5] *(1) Teesside University, Middlesbrough, UK, t.cimpeanu@tees.ac.uk (2) Universidade de Lisboa, Porto Salvo, Portugal (3) Universidade NOVA de Lisboa, Caparica, Portugal (4) Vrije Universiteit Brussel, Brussels, Belgium (5) Université Libre de Bruxelles, Brussels, Belgium*

The field of Artificial Intelligence (AI) is introducing a certain level of anxiety in research, business, and policy making. Tensions are further heightened by an AI race narrative which makes many stakeholders fear that they might be missing out. Whether real or not, a belief in this narrative may be detrimental, as some stakeholders will feel obliged to cut corners on safety precautions or ignore societal consequences just to "win". Starting from a baseline model describing an idealised technology race in a well-mixed world, we investigated how different spatial structures underlying the network of contacts among the race participants can alter the evolutionary outcomes and requirements for regulatory actions. Our findings indicate that, when participants portray a strong diversity in terms of connections and peer-influence (e.g., when scale-free networks shape interactions among parties), the conflicts that exist in homogeneous settings are significantly reduced, thereby lessening the requirement for regulatory actions.

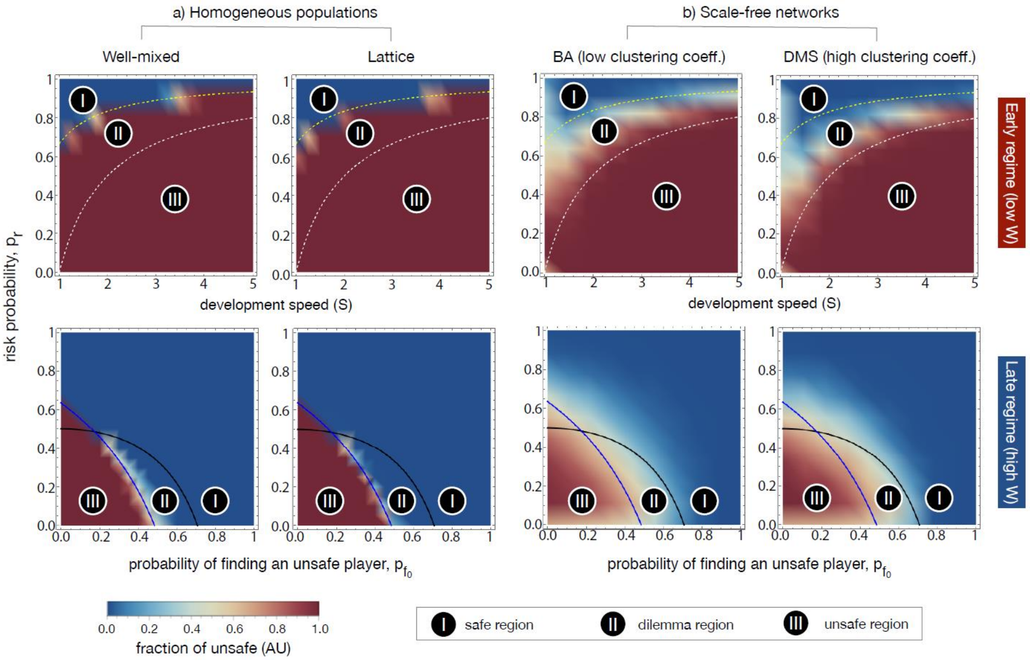


Figure 1: Color gradients indicating the average fraction of AU (unsafe strategy) for different networks and regimes. Dotted and full lines indicate the phase diagram regions obtained analytically. Parameters: $p_{fo} = 0.5$ and $W = 100$ (top panels); $s = 1.5$ and $W = 10^6$ (bottom panels); $c = 1$, $b = 4$, $B = 10^4$, $\beta = 1$, in all panels.

References

[1] T. Cimpeanu, F. C. Santos, L. M. Pereira, T. Lenaerts, T. A. Han. (2020). “AI Development Race Can Be Mediated on Heterogeneous Networks (working title)”. *Preprint Journal Submission.* URL: https://arxiv.org/abs/2012.15234.

Sign of a common pattern in social explosions

Mariano G. Beiró,[1, 2] Ning Ning Chung,[3] Lock Yue Chew,[4, 5]
J. Stephen Lansing,[6, 7] Stefan Thurner,[8, 9] and Yérali Gandica[10]
[1] *CONICET – Universidad de Buenos Aires. INTECIN, Argentina*
[2] *Universidad de Buenos Aires, Facultad de Ingeniería, Argentina*
[3] *Centre for University Core, Singapore University of Social Sciences, Singapore 599494.*
[4] *School of Physical & Mathematical Sciences, Nanyang Technological University, Singapore 637371.*
[5] *Data Science & Artificial Intelligence Research Centre, Nanyang Technological University, Singapore 639798.*
[6] *Santa Fe Institute, 1399 Hyde Park Road, Santa Fe, NM 87501, USA.*
[7] *Stockholm Resilience Centre, Kraftriket 2B, 10691 Stockholm, Sweden.*
[8] *Section for Science of Complex Systems, Medical University of Vienna, Spitalgasse 23, A-1090 Vienna, Austria.*
[9] *Complexity Science Hub Vienna, Josefstädterstraße 39, A-1080 Vienna, Austria.*
[10] *Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS, CY Cergy Paris Université, France*[*]

The success of an on-line movement could be defined in terms of the shift of scale and the later massive off-line protest. The role of the social media in this process is to facilitate the transformation between small or local feelings of disagreement into large-scale social actions. The way how social media achieves that effect is growing clusters of people and groups with similar effervescent feelings, which in other case would be never in communication by several constraints, as for instance, geographical distance.

Society is the emergent result from the interaction between their individuals. It is natural to think that any abrupt change in the structure, led by the communication between all the constituents, could be related to the percolating geometrical structures at the transition points. Independently of the kind of transition (in a physics sense), at the critical point the correlation length is expected to growth attaining a power-law dominance. The latter will, then, give rise to power-laws functionalities on the cluster's statistics forms.
In this respect, we propose that the transition online-offline protest is characterised by some signs of universality, as the expected consequences from the divergence on the correlation function at the critical point.

We show that the frequency-distribution of hashtags for several social explosions are power-law shaped. We have normalised our data-windows in terms of both the time and the number of hashtags. We defined as the transition point the day of the off-line massive actions in the streets.

The analysis has been performed on four different large-scale spontaneous manifestations. The first data set is the Spanish Indignados movement. The critical point for this data-set was the 15th of May 2011. The second and third data sets were taken in Argentina during the year 2019 around two popular movements: a protest against high taxes taking place between January 4th and 6th, and a mobilization asking for justice on November 9th. The last dataset is related to the "Occupy Wall Street" massive demonstrations taking place around May 2012.

Each data-set has been divided into the periods before, during and after the event. The period during the event is the most important, as it sets the time-scale. This is regularly of one day, except for the movement 9-N, which took place during two-days, and the Occupy Wall Street, which is a long-term event. Then, we divided our analysis into two parts, according to how we define the periods before and after the event: for the first analysis, we use the same amount of time to designate the other two periods, namely, before and after. In a second analysis, we extend the timeline before and after the event as to cover the same number of hashtag usages. In figure 1 we plot the number of hashtags used in each period, each segment having the same time-window. Blue, yellow and red designate the periods before, during and after, respectively. We can see completely different kinds of behaviors for the four movements. No similarities seem to appear beyond the highest activity during the event. As an illustration, in figure 2 we show the distributions for the frequency of hashtags on the the data set in figure 1 and the fit for the power-laws exponents.

Our results show that the lowest values for the exponents occur on the data-set "during" for all the events. That consequence of the highest heterogeneity on the distribution is expected for the case of the same-time data-sets, then varying the number of hashtags. However, it starts to be interesting the fact that this is maintained also when the data is divided by having the same number of hashtags, then varying the time. Here there are not more hashtags in that window of time.

---

[*] ygandica@gmail.com

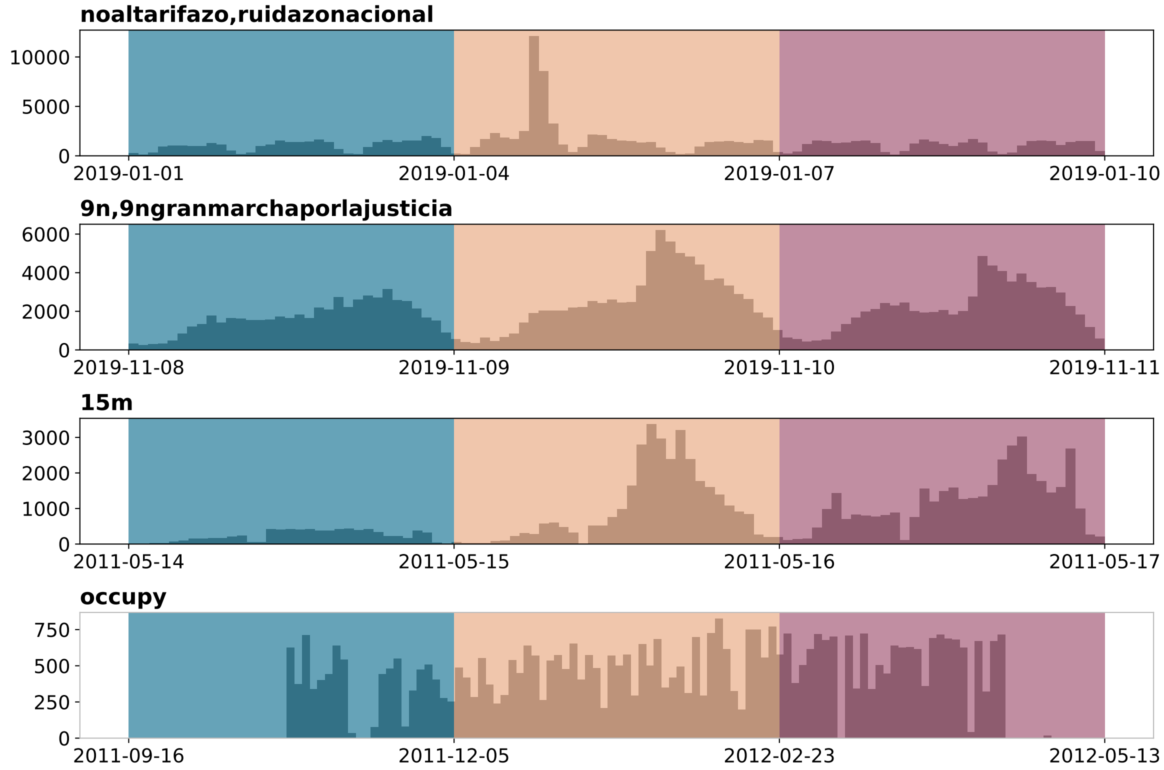


FIG. 1. **Daily activity in each dataset in terms of number of tweets, classified as before, during, and after the online-street-action switch-event date. Each segment has the same time-window.** No similarities were found on the four events beyond the higher activity on the period during.

One of the most interesting results is that the exponents of the periods ”before” and ”during” are the same, within the error bars. That happens when the periods have the same time windows but also when they have, instead, the same number of hashtags. The most surprising finding is that the last phenomena is robust against counting one or all the usages of the hashtags.

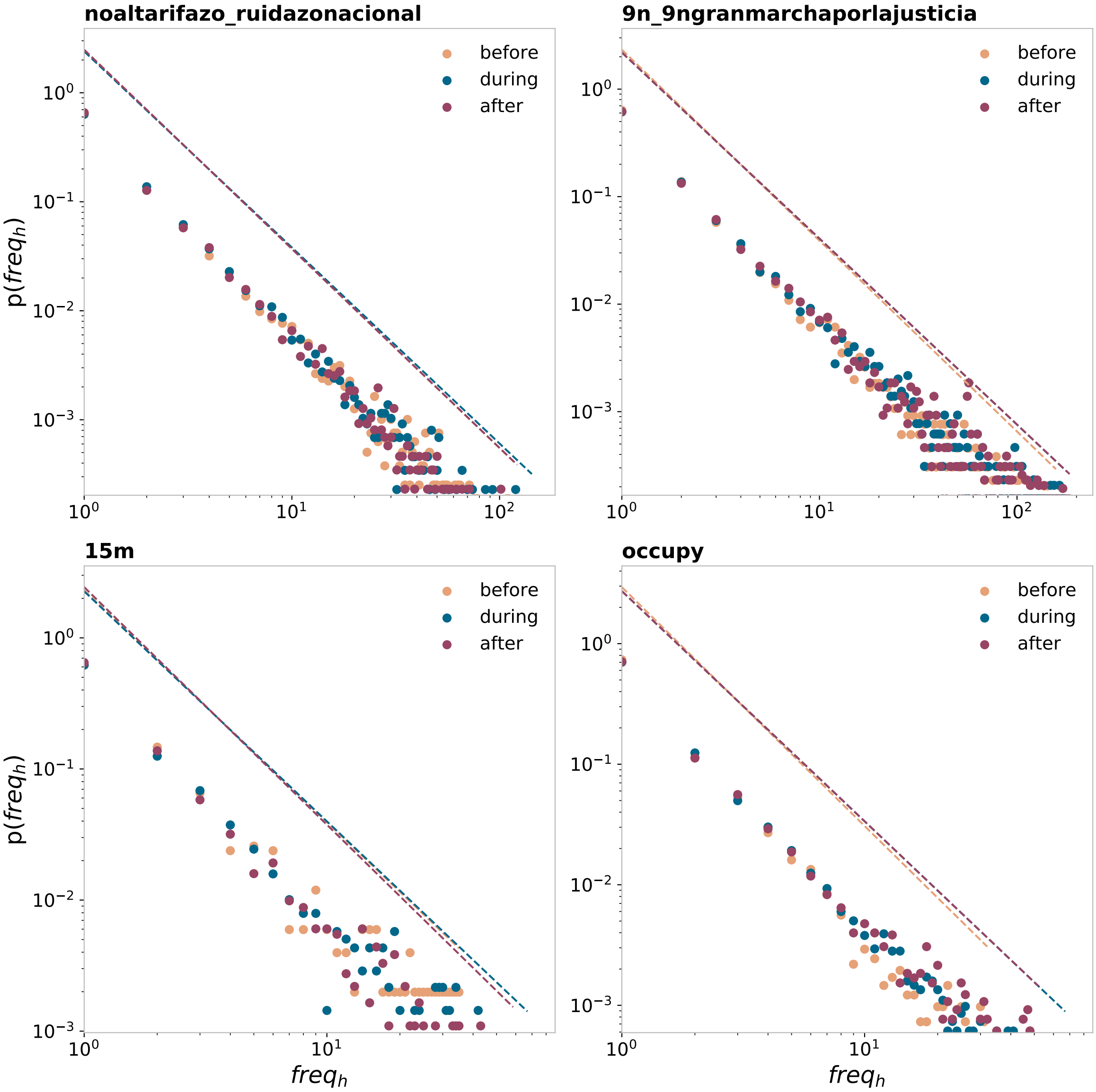


FIG. 2. **Probability distributions of hashtag usage for each data-set**, classified as before, during, and after the events (using a similar time extension for the 3 intervals), with their respective fitted exponents.

The impact of Ethnic Linkage in Global Trade

J. Jun[1], *T. Mizuno*[1,2]. *(1) SOKENDAI, 2-1-2 Hitotsubashi Chiyoda, Tokyo, Japan, joomi@nii.ac.jp. (2) National Institute of Informatics, 2-1-2 Hitotsubashi Chiyoda, Tokyo, Japan*

Global trade is based on a wide variety of factors. In this study, we focus on the ethnicity of corporate executives among the factors. To this end, we extract the factors of ethnicity that have been difficult to grasp so far using the surname-ethnicity classifier built in a prior study and analyze the impact of ethnic linkage on trading.

We built a surname-ethnic classification model [1] using Recurrent Neural Network (RNN) and surname data from more than 35 million corporate executives provided by ORBIS. By applying this classification model to corporate data registered in FactSet, we analyze U.S. customs data from 2014 to 2018 focusing on ethnic factors.

We classify the ethnicity of executives of companies in the U.S. with our surname-ethnicity classifier. And we analyze the trading dataset categorized by country and trade sectors. Figure1 shows trade between companies separated by trade exporting countries and ethnicity as a heatmap. A clear diagonal line in the heatmap means that trade between the same ethnic group occurs more than trade between other ethnic groups.

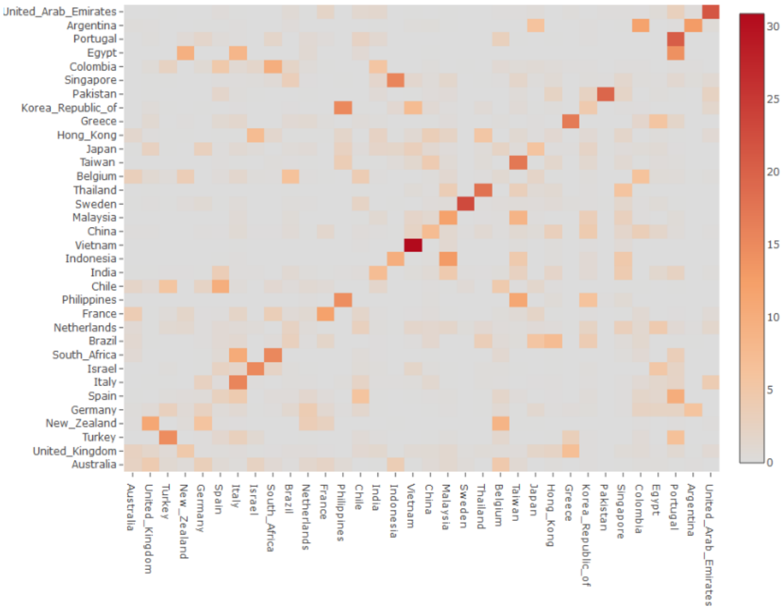


Figure 1: Ethnic self-loop in global trading

We analyze the factors that ethnic-clustering occurs in global trade. Through this, we want to estimate the impact of ethnic linkage and find the key to the solution on issues such as international trade disputes.

References

[1] J. Jun, T. Mizuno, *Detecting Ethnic Spatial Distribution of Business People Using Machine Learning*, *Information*, 11(2020), 197.

Hybrid Structural Arrangements Mediate Stability and Feasibility in Mutualistic Networks

A. Lampo[1], *M. Palazzi*[2], *J. Borge-Holthoefer*[2], *A. Solé-Ribalta*[1,3]. *(1) Universitat Rovira I Virgili, Tarragona, Catalonia, nello.lampo@gmail.com. (2) Universitat Oberta de Catalunya, Barcelona, Catalonia. (3) University of Zurich, Zurich, Switzerland.*

The debate about ecosystems persistence, namely how ecological communities sustain a large number of species, holds considerable importance since the middle of the last century. It has been recognized that the persistence of a system requires stability, but also feasibility, defined as the capability to sustain positive abundances for all biotas. The current abstract focuses on the role of mutualistic interaction architecture jointly considering both properties. So far, the relationship between the emergence of ecological variables and the interspecific interaction networks has been studied by restricting to single structural patterns, such as nestedness or modularity. In our work [1], we propose a new approach relying on hybrid architectural configurations, motivated by the intuition that the co-occurrence of multiple variables cannot be properly grasped by considering different network configurations distinctly. We launch in-block nestedness [2] in the field of ecology, compound by disjoint subsets of species (modules) with internal nested character, and prove that it trades off stability and feasibility. Remarkably, we analyze a large amount of empirical communities and find that a relevant fraction of them exhibits a marked in-block nested structure. We elaborate on the implications of these results, arguing that they provide useful insights about the key properties ruling community assembly.

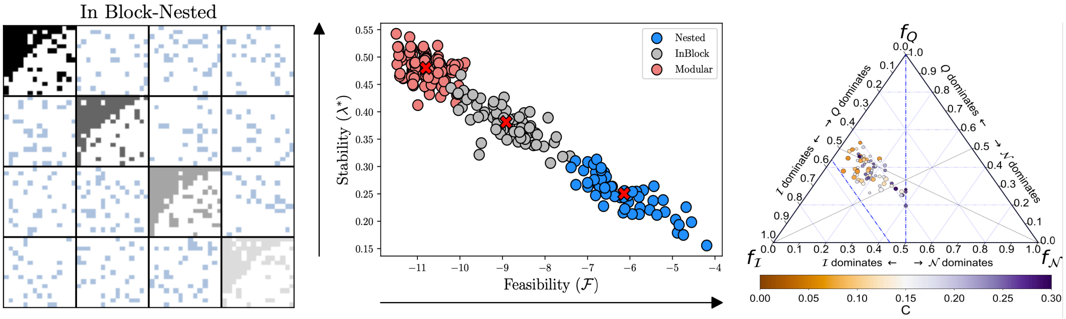


Figure 1: Left: Adjacency matrix of an in-block nested network. Middle: Feasibility vs. stability scatter plot for nested, in-block and modular networks. Stability is quantified through the eigenvalue of the Jacobian matrix, while feasibility as the volume of the intrinsic growth vector space associate to positive abundances. The plot shows the trade-off role of in-block nestedness with respect to the two properties. Right: Ternary plot for the empirical communities networks, highlighting the prevalence of in-block nested configuration.

References

[1] A. Lampo, M. Palazzi, J. Borge-Holthoefer, A. Solé-Ribalta, *arXiv:2105.09191 (2021)*.
[2] M. Palazzi, J. Borge-Holthoefer, C. Tessone, A. Solé-Ribalta. *J.R. Soc. Interface 16*, 20190553 (2019).

An evolutionary model for the spread of information on Twitter

L. Perfeito[1], J. Gonçalves-Sá[1,2]. (1) LIP, Lisboa, Portugal, lperfeito@lip.pt, *(2) Instituto Superior Técnico, Lisboa Portugal*

In the same way that populations of organisms compete for scant resources, memes [1], or ideas also compete for our attention. Here, we use the gene/meme analogy to tackle how information spreads on social media, using Twitter as a case-study.

On Twitter, public short messages (tweets) can be copied (re-tweeted) several times giving rise to long cascades. These cascades have been shown to have a power-law tail, with a few, very large cascades, and the large majority of tweets not being retweeted at all (fig. 1A and [2,3]). This observation is often attributed to the ubiquitous preferential attachment phenomenon, whereby new connections are more likely to be formed with individuals who are already well connected [4] (a form of exponential growth). However, even though tweets travel through such networks, they also possess their own likelihood of being copied. This likelihood, also known as fitness, is key to determine the success of the tweet [5,6] and the distribution of fitnesses among cascades of a given topic is key to determine the fate of the system [7].

Here we took a set of tweets on Covid-19, in four different languages, and fitted a simple model of cascade growth where the success of a cascade is determined by the number of followers of the users that participate in it, its intrinsic fitness, and an aging parameter similar to [8]. We further introduce the concept of Twitter time, a “relativistic” measure that uses user activity as a normalizer. When we simulate cascade growth using different distributions of popularity, we recover a power law tail of cascade sizes. In addition, the exact shape of the distribution is not power-law, but in fact is highly dependent on the fitness distribution.

Our simple model captures well the dynamics of over 98 % of the tweet cascades larger than 100 (examples in fig. 1B). This means that we do not need to invoke any property of the social network to observe a scale-free distribution of Twitter cascade sizes. Popularity varies substantially between cascades, but our results suggest the resulting distributions are better fitted by using a Gamma distribution of fitnesses, than a simple power law (fig. 1A).

These results have two important consequences: 1) we can interpret the distribution of cascade sizes in light of the distribution of popularity and without any knowledge on the underlying network structure; 2) more powerful statistical analyses can be performed on popularity because, unlike the power law distribution, the moments of the Gamma distribution are well defined.

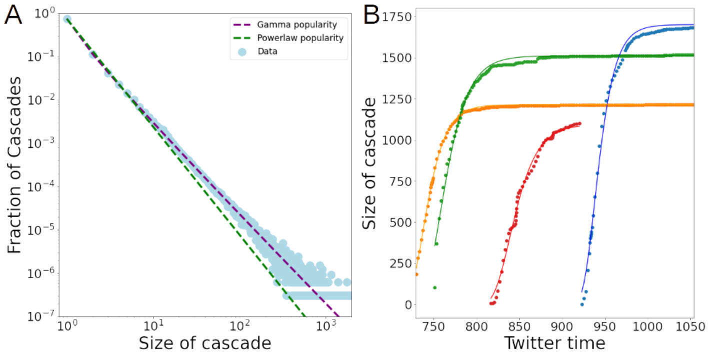


Figure 1*: **A -** Distribution of cascade sizes for tweets in German with the words Covid-19 or Corona between March and September 2020. The points show the data while the lines represent the best fit of a powerlaw (purple) or using a popularity model with a gamma distribution of fitnesses (purple) **B -** Examples of four tweets and their fits to the popularity model.*

References

[1] Dawkins, R. (2016). The selfish gene. Oxford university press.[2] Weng, L., et al (2012). Competition among memes in a world with limited attention. Scientific reports. [3]Gleeson, J. P., et al. (2014). Competition-induced criticality in a model of meme popularity. Physical review letters. [4] Barabási, A. L., & Albert, R. (1999). Emergence of scaling in random networks. science. [5]Bianconi, G., & Barabási, A. L. (2001). Competition and multiscaling in evolving networks. EPL (Europhysics Letters. [6] Kong, J. S., et al. (2008). Experience versus talent shapes the structure of the Web. PNAS. [7] Lera, S. C., et al (2020). Prediction and prevention of disproportionally dominant agents in complex networks. PNAS. [8]Wang, D., et al. (2013). Quantifying long-term scientific impact. Science.

How social interactions influence the discovery of novel music in online platforms

G. Di Bona[1], *E. Ubaldi*[2], *I. Iacopini*[3,4], *B. Monechi*[2], *V. Latora*[1,3,5,6], *V. Loreto*[2,6,7]. *(1) Queen Mary University of London, London E1 4NS, United Kingdom, g.dibona@qmul.ac.uk. (2) SONY Computer Science Laboratories, Paris, 6, rue Amyot, 75005, Paris, France. (3) Centre for Advanced Spatial Analysis, University College London, London, W1T 4TJ, United Kingdom. (4) The Alan Turing Institute, The British Library, London NW1 2DB, United Kingdom. (5) Università di Catania, I-95123 Catania, Italy. (6) Complexity Science Hub, Josefst äadter Strasse 39, A 1080 Vienna, Austria. (7) Sapienza Univ. of Rome, Physics Dept., Piazzale Aldo Moro 2, 00185 Rome, Italy.*

In this work, we characterize the collective exploration of a particular space of concepts, namely the space of music artists, finding that the typical exploitation/exploration strategies of individuals are affected by the structure of their network of contacts. In particular, we analyze a data set containing the whole listening history of a sizeable, socially connected sample of users from the online music platform Last.fm. We find that users connect in an assortative way: those who are keen to explore more tokens of the content space (higher discovery rate) tend to interact with people featuring a similar tendency to explore new contents and vice versa (see Fig. 1A). We propose a new model where a population of socially connected users explore a network of semantically related contents (artists in the data) via a modified random walk. Specifically, we extend the urn model with semantic triggering (already used for individual exploration) to account for interactions between single explorers on the same semantic network. By imposing the same set of parameters and evolution rules for each user, we find that our model reproduces several key features of the empirical data, such as: *i)* the innovators' propensity to cluster together, *ii)* the emergence of communities of individuals sharing the same musical tastes, and, *iii)* a heterogeneous inclination of individuals to explore new contents (see Fig. 1B-C). These results prove that the real-world heterogeneous exploration dynamic can be reproduced without inserting differences in the individuals' behavior, but rather by taking into account the different positioning of the users in the social network. By highlighting the crucial role of social networks, our model represents an important step toward the characterization of collective exploration processes.

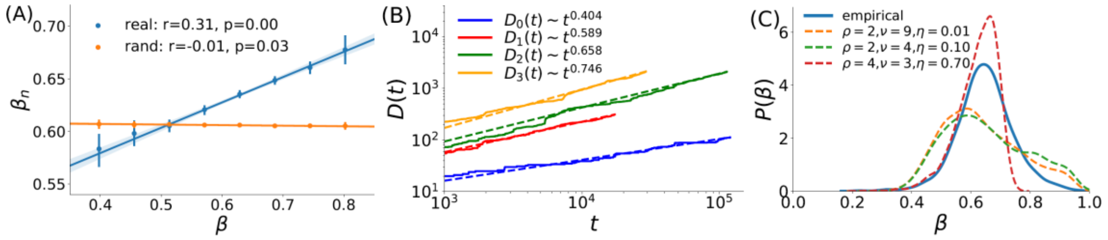


Figure 1: Statistical analysis of Last.fm dataset.

Panel (A): Heaps' exponent assortativity. We plot the average exponent β_n associated with the nearest neighbors (on the social network) of each user i as a function of the exponent β estimated for i. We compare the results from the real network (blue) to those of a randomized social network using a configuration model (orange). Panel (B): Example of Heaps' laws estimation from the sequence of artists listened by four users in the dataset. We report the number of unique artists in a sequence of songs of length t as a function of t. Curves of different users are fitted with power-laws (Heaps' laws) with different exponents β. Panel (C): Distribution of Heaps' exponents estimated from all the users in the empirical data set (continuous line), compared to three simulations of the proposed model with different parameters (dashed lines).

Effect of Interaction Network Structure in Response Threshold Model on Task Allocation

M. Shiraish[1]*, O. Yamanaka, H. Nishimori. (1) Meiji Institute for Advanced Studies of Mathematical Sciences, Meiji University, 4-21-1 Nakano, Nakano-ku, Tokyo 164-8525, JAPAN, shiraishi.mu@gmail.com*

Task allocation in social insects such as ants is an important mechanism to maintain the productivity to live in a large group, so-called a colony. To allocate tasks in the case of ants, worker ants, which are individuals in a colony, spontaneously access the situation and do the task without a leader. The task allocation mechanism is mathematically explained by the Fixed Threshold Response Threshold (FTRT) model [1]. In the FTRT model, the individuals stochastically decide whether to do their tasks. In the FTRT model, a hypothesis is required that all workers could access the global stress precisely. This strong hypothesis is represented as the global coupling of the stress. However, the global coupling hypothesis is not confirmed experimentally, and the worker usually recruits their colleague through local interactions such as antennation or pheromone on the ground. Moreover, in a large colony of ants, the workers construct the networks based on the local interactions [2].

In this study, we introduced the interaction network structure to the FTRT model to overcome the unrealistic global coupling hypothesis of the original FTRT model. We call the new model so as the Network Interaction Response Threshold (NIRT) model, in which stress information is shared by individual interactions. We analyzed the effects on the division of labor mechanism given by the different NW structure of the interaction network (NW) between individuals in the NIRT model. We examined the effects of the network's edge coupling structure and the number of individuals constituting the colony on changes in the number of activities and the amount of stress of individuals. As the number of individuals in the colony increased, the distribution of labor frequency became more localized. As a result, it was difficult to maintain the division of labor structure by local interactions on the lattice-connected NWs, but local interactions on the NW structure with scale-free property enabled the distribution of labor quantity showing the division of labor structure, indicating that the structure of the interacting NWs is important to maintain the division of labor mechanism in practice.

Acknowledgements

This study was financially supported by Grant-in-Aid for Scientific Research (B) No. 20H01871 and the promotion project for development of a characteristic joint research base of the Ministry of Education, Culture, Sports, Science and Technology (JPMXP0620335886).

Reference

[1] Bonabeau, E., Theraulaz, G. & Deneubourg, J.-L. “Fixed response thresholds and the regulation of division of labor in insect societies.”, Bulletin of Mathematical Biology 60, (1998).

[2] Mersch, D. P., Crespi, A. & Keller, L. “Tracking Individuals Shows Spatial Fidelity Is a Key Regulator of Ant Social Organization.”, Science 340, (2013).

Coevolution of Novelty: Scientific Collaboration Networks and Knowledge Diffusion during the COVID-19

Chuqing Zhao[1]. *(1) The University of Chicago, 5801 S Ellis Ave, Chicago, IL, US, cqzhao@uchicao.edu.*

Science can be regarded as a dynamic and complex network. There are many reasons why scientific collaboration could happen. Collaborating efforts across various aspects of science can be vital to fight against the global spread of COVID-19. The changing structure of collaboration network and diffusion of knowledge make it possible to reshape science. Moving beyond the analysis of bipartite networks and single-mode view of scientific collaboration, this project tries to capture the dynamic of collaboration networks during the pandemic and analyze network information from the multiple networks. Findings show that the collaboration network becomes "smaller" with denser connections between scientists, and still follows the "small world effect". The analysis reveals that even under pandemic, geographic location and personal connection between scientists remain significant factors influencing the collaboration behavior. Most biomedical knowledge is stable, but central and novel concepts expand semantic meanings across time. This could be because a field specific concept could diffuse to other disciplines due to scientific collaboration and social learning.

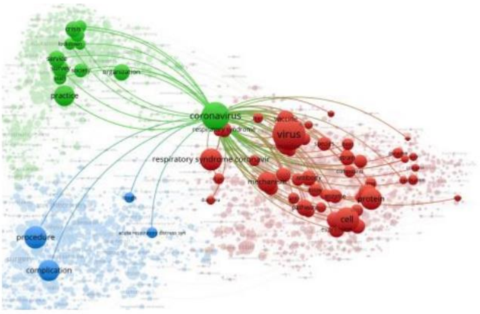


Figure 1: Co-occurrence of "Coronavirus" in the Semantic Network. Each Point Represents a Specific Concepts and Colors Represents Clusters (red: biological term, blue: clinical term, green: public health term).

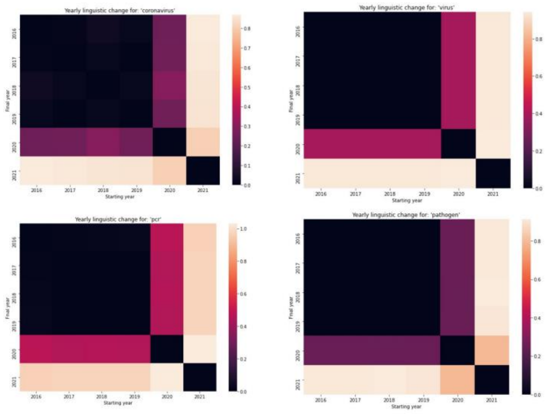


Figure 2: Semantic Shifts of Key Concepts Before and After the Pandemic –"Coronavirus", "Virus", "PCR" and "Pathogen"

Social Network Analysis in Second Language Acquisition

***Michał B. Paradowski*[*]**, *Andrzej Jarynowski*[†], *Jeremi Ochab*[‡], *Chih-Chun Chen*[||], *Karolina Czopek*[§], *Magdalena Jelińska*[*], *Agnieszka Cierpich*[¶]. [*]*Institute of Applied Linguistics, University of Warsaw,* m.b.paradowski@uw.edu.pl; [†]*Interdisciplinary Research Institute in Wrocław/Głogów;* [‡]*Theory of Complex Systems Department, Jagiellonian University;* [||]*Engineering Design Centre, University of Cambridge;* [§]*Institute of English Studies, University of Warsaw;* [¶]*Institute of English Studies, Jesuit University Ignatianum in Kraków*

We demonstrate how computational Social Network Analysis can illuminate the influence of peer interactions on L2 outcomes among immersion/Study Abroad (SA) sojourners in Germany (n=40) and Poland (n_1=181; n_2=210). Unlike in most of the existing research that only tends to identify *ego networks*, we focus on the *full network* approach, because such networks of interactions of individuals over time allow to better apprehend processes such as SLA [1, 2]. Established metrics were used such as node degree, closeness, betweenness and other centrality measures as well as local clustering coefficients, using generalisations of these metrics to weighted graphs. Additionally, we used community detection algorithms and stochastic blockmodeling.

In the German course, we find among others i) that the best predictor of progress is reciprocal interactions between students *in the TL*, ii) that outgoing interactions in the TL are a stronger predictor than incoming interactions, iii) a negative relationship between performance and interactions with same-L_1 speakers (r_{outgoing}= -0.31, CI [-0.63, 0.00], p=.048; r_{incoming}= -0.38, CI [-0.68, -0.08], p=.15) and iv) more intense interactions taking place *across* proficiency groups.

In the Polish course, participants' patterns of social embeddedness in TL communication are significantly moderated by their i) *individual* entry TL competence (positively) and ii) psycho-situational portrait, while iii) negatively by competence in *lingua-franca* English. iv) The influence of the network is strongest in the domains of pronunciation and lexis, where degree centrality *in TL* positively correlates with progress ($r_{\text{outdegree}}$=0.258, p=.001 for pronunciation; $r_{\text{outdegree}}$=0.304, p=.0002 and r_{indegree}=0.263, p=.001 for vocabulary), while betweenness *in total communication* is significantly anticorrelated (r=-0.242, p=.003 and r=-0.204, p=.01, respectively). v) This mirrors the influence direction—on global TL progress—of closeness centrality. Combined with the detrimental impact on language acquisition of a high in-degree, this suggests that for language acquisition, the structural properties of the network matter more than processes such as information flow.

Acknowledgements

The authors are supported by SONATA-BIS Grant № 2016/22/E/HS2/00034 from the National Science Centre of Poland.

References

[1] Paradowski, M.B., Jarynowski, A., Czopek, K. & Jelińska, M. Peer interactions and second language learning: The contributions of Social Network Analysis in Study Abroad vs At-Home environments. In: R. Mitchell & H. Tyne (Eds.), *Language, Mobility and Study Abroad in the Contemporary European Context*. New York: Routledge 2021, 99–116. DOI: 10.4324/9781003087953-8.

[2] Paradowski, M.B., Jarynowski, A., Jelińska, M. & Czopek, K. Out-of-class peer interactions matter for second language acquisition during short-term overseas sojourns: The contributions of Social Network Analysis. *Language Teaching*, *54*(1) (2021), 139–143. DOI: 10.1017/S0261444820000580

A new framework for understanding resilience and change in long-term trajectories of social complexity using archaeological data

D. Daems[1], (1) Middle East Technical University, Universiteler Mh, Çankaya, Ankara, Turkey, daems@metu.edu.tr

Human societies have changed dramatically over the last 12,000 years, from small hunter-gatherer bands to highly urbanized, industrialized, and globally connected societies. These changes are often interpreted as increasing social complexity. Yet, what exactly constitutes social complexity and how it changes over time is not always specified. It is all too easy to perceive this change as an almost teleological sequence of progress and advancement towards a better society. No doubt modern societies have produced many benefits worthy of being characterized as progress. However, by recasting all of history in this light, we run the risk of discarding the past as an inherently inferior and altogether different world. In doing so, we lose out on the benefits of understanding long-term patterns of change (and stability) to inform our present and future.

In this paper, I present a novel framework of social complexity formation centered on causal mechanisms of decision-making, driving pushing and pulling forces that shape flows of energy, resources and information. This framework draws from complex systems thinking as a suitable ontological framework to provide a bottom-up approach to social complexity as an emergent property in human societies. In this sense, human societies can be conceptualized as complex systems aggregating and/or dispersing flows of energy, resources and information to sustain their population and maintain/expand their organisational structures (Daems 2021). I apply this framework to a case study of southern Anatolia (modern-day Turkey) from Bronze Age to Hellenistic times (3000-100 BCE). In this paper, I use archaeological data such as settlement patterns, central place formation, artisanal production and distribution as proxies for the capture, processing and usage of flows of energy, resources and information to study resilience and change in long-term trajectories of social complexity.

This work will not only improve our knowledge of the past *sui generis*, but also elucidates how understanding long-term trajectories of social complexity can inform trade-offs in decision-making and pathways of complexity formation in the present. This study shows that shifts and trade-offs in decision-making can have long-lasting consequences in social complexity trajectories expressed in human-environment interactions. It also raises an important point of reflection for the long-term consequences of decision-making in the present and the initiation of future pathways of development. The framework proposed here provides a suitable approach to use the past to inform such decision-making processes through the lens of complex systems thinking.

References

[1] D. Daems, *Social Complexity and Complex Systems in Archaeology*, London, 2021.

Non-normality driven polarization in social dynamics

B. Hennessy[1], *M. Asllani*[1]. *(1) School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland* brian.hennessy1@ucdconnect.ie

Over the past years, network science has become a decisive investigation tool in computational social science where systems are constituted by many agents that interact through complex structures [1], [2]. It has been recently shown that real-world networks, in particular social ones, manifest ubiquitous features: high asymmetry and, most importantly, strong non-normality, i.e., their adjacency matrix $\mathbf{A}$ is such that $\mathbf{AA}^T \neq \mathbf{A}^T\mathbf{A}$ [3]. Such topological feature has a crucial impact on dynamical behavior by amplifying the response to external perturbations, eventually allowing self-organized patterns to emerge [4]. Furthermore, strongly non-normality networks are characterized by the exclusive presence of single node strongly connected components (sink or source), named leaders, giving rise to a novel underlying hierarchical structure [5].

Our study analyzes and quantifies the effects of non-normality and, in particular, the role of leader nodes on social dynamical models to determine how non-normality drives and further shapes opinion polarization. We have considered the paradigmatic Susceptible-Infected-Susceptible (SIS) model with a nonlinear incidence term over a contact network where nodes represent agents (actor) of the social network. Each node has been assigned an attribute (opinion, merit, richness, etc.) identified with the probability of being infected. We follow a symmetry-breaking mechanism where the system initially settles in a consensus regime; namely, the actors share the same opinion. Using the eigenvector centrality, we prove that the leader nodes not simply drive the emergence of a new opinion but also force the polarization of the rest of the agents, which segregate around clusters of similar opinion (see Fig. 1). Our predictions correlate well with simulation in empirical data, and our analysis reveals hidden clusters in real social networks.

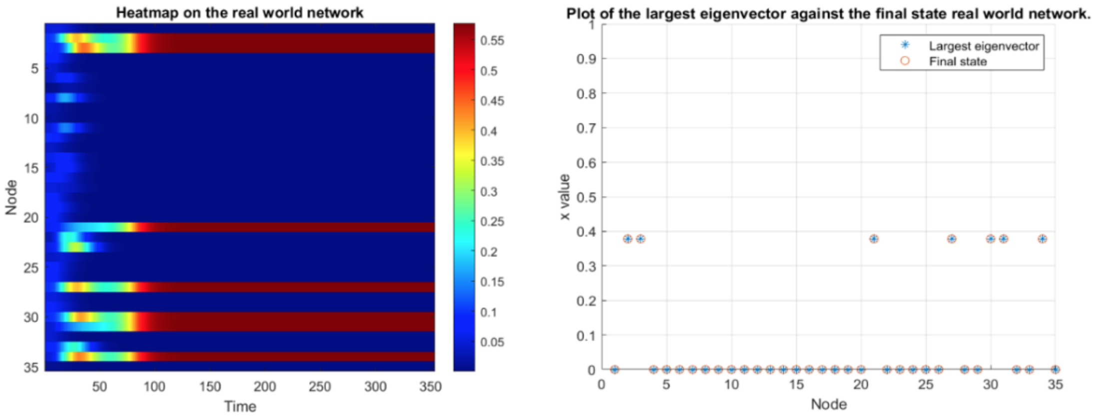


Figure 1: (Left) emergence of polarized opinion and (right) eigenvector prediction of patterns in social network of Dutch literary actors.

References

[1] Newman M.E.J., *Networks: An Introduction*. 2nd Edition, Oxford University Press (2018).
[2] Albert R., Barabási A.L., *Rev. Mod. Phys.* **74**, 47–97 (2002).
[3] Asllani M., Lambiotte R., Carletti T., *Sci. Adv.*, **4**: eaau9403, (2018).
[4] Muolo, R., Carletti, T., Gleeson, J. P., & Asllani, M., *Entropy* **23** (1), 36 (2021).
[5] O'Brien, J. D., Oliveira, K. A., Gleeson, J. P., & Asllani, M., *Phys. Rev. Res.* **3** (2), 023117 (2021).

Complementing Evolutionary Approaches in Public Policy with Complexity Theory: A Dissipative Structure Approach

T. Weber[1]*, J. Louçã*[1]*, L. Gerrits*[2]*. (1) Department of Information Science and Technology (ISTA), ISTAR–Instituto Universitário de Lisboa (ISCTE-IUL), Av. das Forças Armadas, 1649-026, Lisbon, Portugal, twrsh@iscte-iul.pt. (2) Institute for Housing and Urban Development Studies (IHS), Erasmus University Rotterdam, Burgemeester Oudlaan 50 Mandeville, Rotterdam, The Netherlands*

Public policy is a complex matter and thus several scholars have used complexity science to gain a better understanding of this field. Lately, there have been calls to update the policy subsystem concept as a classical component of public policy studies with complexity science to deal with cross-subsystem dynamics [1]. It is assumed that policy theories like the Punctuated Equilibrium Theory and the Advocacy Coalition Framework fail to adequately deal with these dynamics because they focus too much on internal subsystem processes.

We want to contribute to this debate by adding a focus on self-organization through communication among policy actors which we draw from our previous work. There, we have conducted a theory transfer of the theory of dissipative structures from thermodynamics to an application in social science [2]. In the original thermodynamic theory, inflows of matter and energy sustain structures of open systems. When we transfer this notion to policy subsystems, we think that policy subsystems are open systems consisting of collective belief structures (policy images) about a policy issue which are maintained by communications of coalitions of policy actors. Through the transfer we can analyze the evolution of these structures with the concept of ‘order through fluctuation’ from dissipative structure theory which fits well to models of punctuated equilibrium in public policy. This means that there are normal phases where fluctations from the prevailing policy image are dampened by the coalitions, and bifurcation periods where critical fluctuations can be amplified to change the qualitative nature of the policy image.

We illustrate these conceptual advances with a case study about the plans for a new runway at Munich airport. There, the prevailing policy image that the infrastructure project should be built was maintained for an extended time period and deviating opinions were dampened by the dominant coalition. But when the system approached a bifurcation period, these deviating opinions could eventually destabilize and subsequently change the policy image towards a moratorium of the whole project. We also create an agent-based model to test whether our theoretical assumptions can produce the observed outcomes in the empirical case study.

Overall, we want to innovate public policy studies and foster interdisciplinary research by showing how the conceptual framework of the theory of dissipative structures can help to grasp the empirical reality of policy subsystems in a sophisticated way. This can help to further refine evolutionary and complexity approaches for the study of public policy.

References

[1] Z. A. McGee, B. D. Jones, Reconceptualizing the Policy Subsystem: Integration with Complexity Theory and Social Network Analysis, *Policy Studies Journal*, 47 (2019), 138-158. https://doi.org/10.1111/psj.12319

[2] T. Weber, J. Louçã, L. Gerrits, Dissipative structures and the relation between individual and collective aspects of social behavior, *Systems Research and Behavioral Science,* (2021), 1–13. https://doi.org/10.1002/sres.2777

How to organize and adapt collective action in a complex social-ecological landscape where global change pressures the sustainability of multiple ecosystem services: case study in a French rurban area

__A. Brias__[1,2], J-B. Pichancourt[2], A. Bonis[1]. (1) Université Clermont Auvergne, 4 Rue Ledru, 63000, Clermont-Ferrand, France, antoine.brias@uca.fr. (2) INRAE, 9 Avenue Blaise Pascal, 63170, Aubière, France

In the context of fast and poorly predictable global changes, people from rurban areas face complex challenges for sustaining biodiversity and the provisioning of related ecosystem services essential for their livelihoods. A first layer of complexity results from the multiple ecological interactions between species, abiotic ressources and conditions and living organisms that are involved in ecosystem services provisions. A second layer of complexity comes from the network of social interactions, land-tenure system and land-use management preferences. Given this social-ecological complexity, there is a scientific challenge to formalize how environmental conditions and anthropic influences shape biodiversity and environmental patterns and thereafter ecosystems services. This will constitute an important tool to help people managing their actions to sustain ecosystem services, and whose timing of investment can be adapted on the long run in response to relevant global change warning signals.

In this paper, we focus on how actions around the adaptation of a common hedgerow network can be organized in response to global change in a rurban landscape (in central France near Clermont-Ferrand, widespread now in Europe). The hedgerow network is a collective semi-natural infrastructure that is key to sustain the functioning of a diversity of other semi-natural habitats that deliver a set of essential ecosystem services (e.g. food in adjacent fields, tourism, carbon sequestration in forests, habitat for biodiversity, protection against land-slide). To best organize investments, we propose a method that is based on the coupling of three major theories: the robustness framework of social-ecological systems developed by the 2009 Economic Nobel Price Laureate E. Ostrom's [1], the Aubin's viability theory [2] and the adaptation policy pathways framework [3].

For our case study, we show how this method can be used to organize data, construct mathematical models of social-ecological system dynamics, estimate viability kernels for a set of ecosystem services, estimate early warning signals that guide the triggering of collective action in response to various levels of global change, and finally organize these actions along an adaptation pathway plan.

Acknowledgements

The PACSEN project is co-financed by the European Union within the framework of the European Regional Development Fund (ERDF).

References

[1] Ostrom, E. (2009). *Social-Ecological Systems*. *325*(5939), 419–422.

[2] Aubin, J.-P., Bayen, A., & Saint-Pierre, P. (2011). *Viability Theory. New Directions*.

[3] Lavorel, S., Colloff, M. J., Locatelli, B., Gorddard, R., Prober, S. M., Gabillet, M., Devaux, C., Laforgue, D., & Peyrache-Gadeau, V. (2019). Mustering the power of ecosystems for adaptation to climate change. *Environmental Science and Policy*, *92*(July 2018), 87–97. https://doi.org/10.1016/j.envsci.2018.11.010

Zipf's law for cities: the end

***Vincent Verbavatz*[1,2], *Marc Barthelemy*[1,3], *(1) Université Paris-Saclay, CNRS, CEA, Institut de physique théorique, 91191, Gif-sur-Yvette, vincent.verbavatz@ipht.fr. (2) École des Ponts ParisTech, Champs-sur-Marne, France. (3)Centre d'Etude et de Mathématique Sociales, CNRS/EHESS, 54 Boulevard Raspail, 75006, Paris, France.*

Cities are among the most famous complex systems in social sciences: they concentrate a very large number of interacting people – inhabitants and workers - over short length scales. In this framework, modelling the population evolution of cities is at the core of all urban studies. Quantitatively, the most fundamental problem is to understand the population dynamics and the hierarchical organization of cities, usually thought to be described by a universal law due to Zipf [1]. Zipf's law states that the distribution of city populations in a country follows a power-law with exponent close to 1, a usual result in complex systems. In a seminal article [2], Gabaix proved how a stochastic growth process for cities could induce Zipf's law. However, the validity of Zipf's law has been challenged by recent empirical studies [3] while Gabaix's model is not able to reproduce the turbulent dynamics of cities through History [4]. Here we introduce a new kind of stochastic equation for modelling population growth in cities, which we construct from an empirical analysis of recent datasets (for Canada, France, UK and USA) that reveals how rare but large interurban migratory shocks dominate city growth [5]. This equation predicts a complex shape for the city distribution and shows that Zipf's law does not hold in general. It also predicts the existence of multiple temporal variations in the city hierarchy, in agreement with observations.

References

[1] Zipf, G. K. (1949). Human behavior and the principle of least effort: an introd. to human ecology.

[2] Gabaix, X. (1999). Zipf's law for cities: an explanation. *The Quarterly journal of economics*, *114*(3), 739-767.

[3] Arshad, S., Hu, S., & Ashraf, B. N. (2018). Zipf's law and city size distribution: A survey of the literature and future research agenda. *Physica A: Statistical Mechanics and its Applications*, *492*, 75-92.

[4] Batty, M. (2006). Rank clocks. *Nature*, *444*(7119), 592-596.

[5] Verbavatz, V., & Barthelemy, M. (2020). The growth equation of cities. *Nature*, *587*(7834), 397-401.

The opinions of the few: influencers at review platforms

***O. Mokryn**[1]. (1) Information Systems, University of Haifa, Israel. ossimo@gmail.com.*

Review platforms employ a voting mechanism, in which the crowd is invited to up-vote reviews that they find useful. Previous research identified several predictive characteristics, such as the reviews' sentiment, style of writing and information quality, peripheral cues, and specific emotions, such as anger, expressed in the review. However, several meta-analysis works showed that the literature contains conflicting findings.

This research investigates the attributes of over 1.2M reviews written by more than 327K users over three different review platforms. We find that reviews that are voted as helpful, on average, are authored by a small group of authors and that the best predictor for the success of a review is the author's history. We train an XGBoost classifier with textual features, the emotional effect of a review, measured by the amount of expressed emotions according to Plutchik's wheel of discrete emotions, and the reviewer's previous impact, over the SMOTE-balanced datasets. We find that the emotional effect is of a lesser explaining value than the reviewer's history but still performs better than textual features. The emotional effect is more significant for service goods such as restaurants and films (as opposed to products). Regression analysis for the exact score reveals a cross-platform predictive power of the author's impact and review length. Discrete negative emotions correlate positively with the usefulness of reviews, more so for experience goods. We define influential writers as reviewers who have a high review-platform h-index. A review platform h-index is defined as follows: Reviewer has index h if h of her Np reviews have at least h useful votes each, and the other $(Np - h)$ reviews have no more than h useful votes each.

A detailed analysis of the *path to glory* of the influential reviewers shows that reviewers with a high h-index are prolific writers, writing on average between 160 and 1000 reviews. As can be seen in Fig 1, the number of reviewers with a high h-index is very low. When we consider the imbalance in votes, where only a small percentage of the reviews are voted useful, we can conclude that a small group of reviewers authors most useful reviews. Their timeline shows that it takes them around 10 months to become influential, indicating there might be a community voting mechanism rather than the expected crowdvoting mechanism.

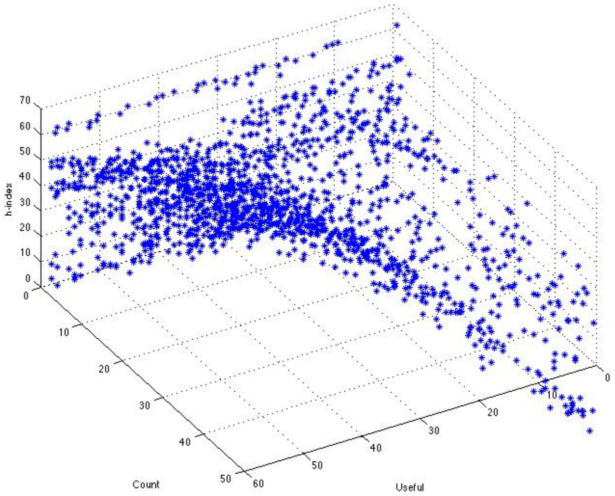


Figure 1: Influential authors' performance

What is your tendency to converge in a flock? Assessing the convergent behaviour of individuals collaborating as a complex system through self-reported measure

S. Magnanini[1]*, F. Guazzoni*[1]*, D.Trabucchi*[1]*, T.Buganza*[1]*, R.Verganti*[2] *(1) School fo Management, Politecnico di Milano, Via Lambruschini 4/C, Milano, Italy, silvia.magnanini@polimi.it. (2) House of Innovation, Stockholm School of Economics, Saltmätargatan 13, 113 59 Stockholm, Sweden*

There is a growing interest about how individuals in the organization can join and converge towards the same direction of change while articulating new strategic projects (Dell'Era et al., 2020; Pasmore et al., 2008; Verganti, 2017). Literature about Complex Adaptive Systems (Dooley, 1997; Palmberg, 2009) and social information processing (Fulmer & Ostroff, 2016; Pfeffer & Salancik, 1978; Weick, 1995) has interpreted convergence as a form of social construction where agents in the system collaboratively make sense of the changing environment by constructing shared knowledge (Canessa & Riolo, 2006, p. 1). From the sum of non-linear interactions, the system emerges and self-organizes itself (Ashmos et al., 2002; De Toni et al., 2012), exchanging information and ultimately building shared meaning (Stigliani & Ravasi, 2012; Weick, 1995). Besides, Flock models are becoming popular as agent-based simulation frameworks to conceptualize the complex emergent behavior of agents according to simple interaction rules (Goldstone & Gureckis, 2009; Reynolds, 1987; Uhl-Bien et al., 2007; Will, 2016).

Whereas Flock models can conceptualize human activities within a framework of emergence (Davis et al., 2007; J. R. Harrison et al., 2007; Reynolds, 1987), there is still a lack of empirical evidence on whether these virtual models can provide an exhaustive representation of individual behaviors within an actual collaborative setting (Will, 2016). Our study aims to answer to the following research question: *Can we assess the behavior of individuals converging within a CAS through self-reported measure?*

Through a laboratory experiment (G. W. Harrison & List, 2004), 60 students were involved to compare the tendency to converge manifested by individuals through self-reported measurement with their actual behavior expressed during a cooperative virtual game. First, the interaction rules governing the agent-based simulation of the Flock model (Will, 2016) were operationalized in a self-reported questionnaire build around key constructs from organization theory and social information processing. Constructs assess the properties that lead individuals to converge while collaborating. They are namely Cognitive Flexibility (Webster & Kruglanski, 1994), Uniqueness (Schumpe et al., 2016), Synthesis capabilities (Taggar, 2002) and Flexibility for change (Di Fabio & Gori, 2016). Secondly, a virtual cooperative game titled Build your City was purposefully designed to assess the questionnaire constructs so that - by tracking players' moves - was possible to capture the emergence of individual behaviours.

During data analysis, respondents' answers will be compared with data from ten game sessions of 6 players each. Information from each game session include participants' actions and created content, verbatim transcripts, supported by audio recordings.

We believe that expected findings can contribute to the extant literature in organization theory (Dooley, 1997; Fulmer & Ostroff, 2016; Palmberg, 2009) and social information processing (Pfeffer & Salancik, 1978), providing a method to measure behaviours in social systems. From a managerial perspective, the present study proposes a tool – conceived as a collaborative environment with gamified dynamics– that could support managers and people in the organization to collaboratively construct a shared understanding and ultimately foster convergence among individuals.

References

Ashmos, D. P., Duchon, D., McDaniel, Jr, R. R., & Huonker, J. W. (2002). What a Mess! Participation as a Simple Managerial Rule to 'Complexify' Organizations. *Journal of Management Studies*, *39*(2), 189–206. https://doi.org/10.1111/1467-6486.00288

Canessa, E., & Riolo, R. L. (2006). An Agent–Based Model of the Impact of Computer–Mediated Communication on Organizational Culture and Performance: An example of the Application of Complex Systems Analysis Tools to the Study of CIS. *Journal of Information Technology*, *21*(4), 272–283. https://doi.org/10.1057/palgrave.jit.2000078

Davis, J. P., Eisenhardt, K. M., & Bingham, C. B. (2007). Developing Theory Through Simulation Methods. *Academy of Management Review*, *32*(2), 480–499. https://doi.org/10.5465/amr.2007.24351453

De Toni, A. F., Biotto, G., & Battistella, C. (2012). Organizational design drivers to enable emergent creativity in web-based communities. *The Learning Organization*, *19*(4), 335–349. https://doi.org/10.1108/09696471211226699

Dell'Era, C., Magistretti, S., Cautela, C., Verganti, R., & Zurlo, F. (2020). Four kinds of design thinking: From ideating to making, engaging, and criticizing. *Creativity and Innovation Management*, *29*(2), 324–344. https://doi.org/10.1111/caim.12353

Di Fabio, A., & Gori, A. (2016). Developing a New Instrument for Assessing Acceptance of Change. *Frontiers in Psychology*, *7*(802), 10. https://doi.org/10.3389/fpsyg.2016.00802

Dooley, K. J. (1997). A Complex Adaptive Systems Model of Organization Change. *Human Sciences Press Inc.*, *1*(1), 29.

Fulmer, C. A., & Ostroff, C. (2016). Convergence and emergence in organizations: An integrative framework and review: CONVERGENCE AND EMERGENCE. *Journal of Organizational Behavior*, *37*, S122–S145. https://doi.org/10.1002/job.1987

Goldstone, R. L., & Gureckis, T. M. (2009). Collective Behavior. *Topics in Cognitive Science*, *1*(3), 412–438. https://doi.org/10.1111/j.1756-8765.2009.01038.x

Harrison, G. W., & List, J. A. (2004). Field Experiments. *Journal of Economic Literature*, *42*(4), 1009–1055. https://doi.org/10.1257/0022051043004577

Harrison, J. R., Lin, Z., Carroll, G. R., & Carley, K. M. (2007). Simulation modeling in organizational and management research. *Academy of Management Review*, *32*(4), 1229–1245. https://doi.org/10.5465/amr.2007.26586485

Palmberg, K. (2009). Complex adaptive systems as metaphors for organizational management. *The Learning Organization*, *16*(6), 483–498. https://doi.org/10.1108/09696470910993954

Pasmore, W. A., Stymne, B., Shani, A. B. (Rami), Mohrman, S. A., & Adler, N. (2008). The Promise of Collaborative Management Research. In A. B. (Rami) Shani, S. A. Mohrman, W. A. Pasmore, B. Stymne, & N. Adler, *Handbook of Collaborative Management Research* (pp. 6–32). SAGE Publications, Inc. https://doi.org/10.4135/9781412976671.n1

Pfeffer, J., & Salancik, G. R. (1978). *The External Control of Organizations: A Resource Dependence Perspective*. Social Science Research Network. https://ssrn.com/abstract=1496213

Reynolds, C. W. (1987). Flocks, herds and schools: A distributed behavioral model. *ACM SIGGRAPH Computer Graphics*, *21*(4), 25–34. https://doi.org/10.1145/37402.37406

Schumpe, B. M., Herzberg, P. Y., & Erb, H.-P. (2016). Assessing the Need for Uniqueness: Validation of the German NfU-G scale. *Personality and Individual Differences, 90*, 231–237. https://doi.org/10.1016/j.paid.2015.11.012

Stigliani, I., & Ravasi, D. (2012). Organizing Thoughts and Connecting Brains: Material Practices and the Transition From Individual to Group-Level Prospective Sensemaking. *The Academy of Management Journal, 55*, 1232–1259. https://doi.org/10.5465/amj.2010.0890

Taggar, S. (2002). Individual Creativity and Group Ability to Utilize Individual Creative Resources: A Multilevel Model. *Academy of Management Journal*, *45*, 315–330. https://doi.org/10.2307/3069349

Uhl-Bien, M., Marion, R., & McKelvey, B. (2007). Complexity Leadership Theory: Shifting leadership from the industrial age to the knowledge era. *The Leadership Quarterly*, *18*(4), 298–318. https://doi.org/10.1016/j.leaqua.2007.04.002

Verganti, R. (2017). *Overcrowded: Designing meaningful products in a world awash with ideas.* The MIT Press. https://mitpress.mit.edu/books/overcrowded

Webster, D. M., & Kruglanski, A. W. (1994). Individual Differences in Need for Cognitive Closure. *Journal of Personality and Social Psychology, 67(6)*, 1049–1062. https://doi.org/doi:10.1037/0022-3514.67.6.1049

Weick, K. E. (1995). *Sensemaking in Organizations*. SAGE Publications Inc. https://us.sagepub.com/en-us/nam/sensemaking-in-organizations/book4988

Will, T. E. (2016). Flock Leadership: Understanding and influencing emergent collective behavior. *The Leadership Quarterly*, *27*(2), 261–279. https://doi.org/10.1016/j.leaqua.2016.01.002

The Effects of Party Competition on Consensus Formation

G. Romero Moreno[1], J. Padilla[2], M. Brede[1]. (1) University of Southampton, Southampton SO17 1BJ, United Kingdom, Guillermo.RomeroMoreno@soton.ac.uk. (2) City University of New York, Address, New York City, NY 10017, United States of America

The fight over setting the political agenda is one of the basic mechanisms of party competition of every democracy. However, this political game may have side effects in other aspects of the public debate. One aspect of general interest is how it may alter consensus formation processes among citizens, which may result in states of consensus, polarisation, or opinion fragmentation in the population. In this paper, we study the interrelated dynamics of two processes affecting opinion dynamics when multiple issues are debated. First, we consider a noisy bounded-confidence model to describe the dynamics of citizens' opinion and consensus formation. Second, we build a model of party competition to gain votes. In political science, most research on party competition originates from spatial voting theory, where citizens vote for parties according to the distance of their positions to party positions. In this context, parties compete for vote shares by adjusting the saliency given to the different political issues (e.g. by campaigning) to promote aspects of the debate in which they hold favourable positions relative to the electorate. These changes in saliency affect in times the perception of opinion distance between citizens and their possibility of interaction, generating coupling between both processes. We use agent-based simulations for the study of our model. For simplicity, we focus on the case of two dimensions and three political parties, which already serves to illustrate possible model outcomes. We find that the effects of party competition on consensus formation are rich and non-trivially dependent on the configuration of party positions in the political space. We illustrate that –as one would intuitively expect– there are party configurations that foster polarisation for a wide range of model parameters. More surprisingly, we also show that other party configurations have the opposite effect, and can facilitate reaching a consensus state that could otherwise not have been achieved. We find that these differences in outcomes are generally robust to model parameters and are mainly dependent on the constellation of party positions. Our results serve to emphasise the importance that absolute and relative party positions have in the creation of polarisation and illustrate the richness of possible outcomes of interrelations between party competition and consensus formation.

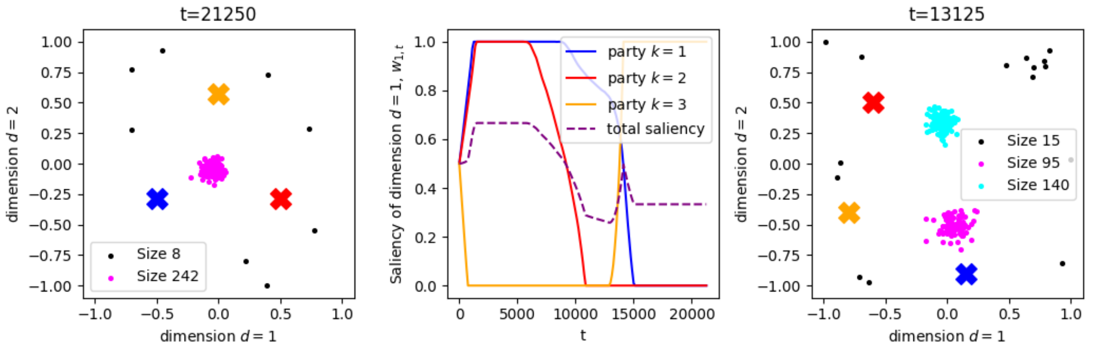


Figure 1: Left: Illustration a simulation where party competition fosters consensus. Dots are opinions of a population in the stationary state, and crosses represent fixed party positions. Center: Evolution of parties' support for dimension *d=1* of opinions for the scenario shown in the left panel. Right: Illustration of a simulation where party competition fosters polarisation.

Identifying Decarbonization Leverage Points in Supply Networks with Network Measures that Quantify Systemic Relevance

J. Stangl[1], *C. Diem*[1], *T. Reisch*[1,2], *S. Thurner*[1,2,3]. *(1) Complexity Science Hub, Josefstädter Str. 39, 1080 Vienna, Austria, stangl@csh.ac.at. (2) Medical University of Vienna, Spitalgasse 23, 1090 Vienna, Austria, (3) Santa Fe Institute, Santa Fe, 1399 Hyde Park Rd, NM 75791, USA*

Climate change and the decarbonization of our global economy are among the defining challenges of the 21st century. A large amount of greenhouse gas emissions arise from the production of economic goods and services [1]. Production processes depend on supply chains involving a multitude of individual companies that together form complex supply networks. Our goal is to find leverage points in these supply networks that minimize the system-wide greenhouse gas output, while keeping disruptions to production to a minimum. Here we show that these leverage points can be detected via combining the emissions of companies with the assessment of their socio-economic relevance. The latter we compute from network measures that have been used to compute systemic risk in economic networks [2]. Companies with high emissions and low socio-economic relevance are potential leverage points for decarbonization. We demonstrate this idea for the 20,000 companies in the Austrian pork supply network. Emissions of every company are calculated from from a life cycle assessment of Austrian pork [3]. Preliminary results are shown in Figure 1, where we identify companies with high emissions and low systemic risk (dots in red zone). Policymakers might target these companies with regulatory policies or tailored carbon taxation schemes.

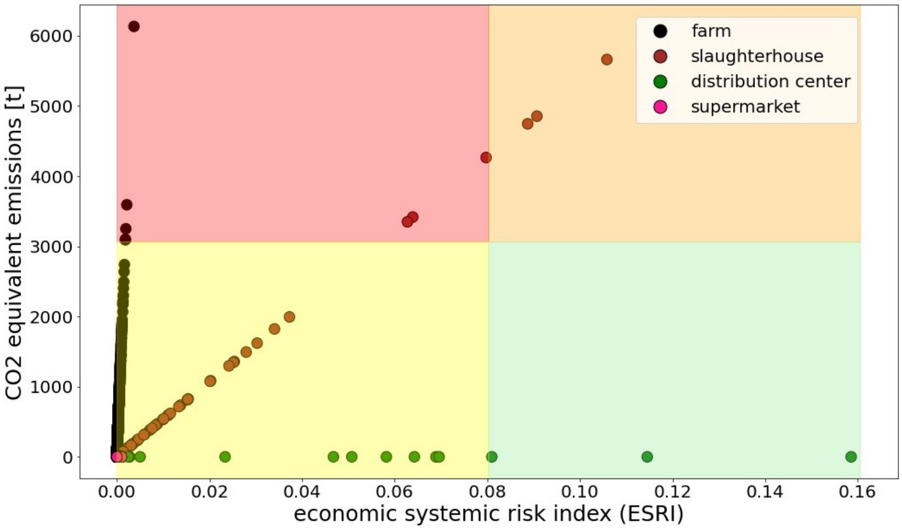


Figure 1: Annual CO2 emissions of companies in the Austrian pork supply network vs. their socio-economic relevance; Companies in the red zone are potential leverage points.

References

[1] IPCC, 2014: Climate Change 2014: Synthesis Report: Summary for Policymakers

[2] C. Diem, A. Borsos, T. Reisch, J. Kertesz, S. Thurner, *arXiv:2104.07260v1*

[3] T. Winkler, K. Schopf, R. Aschemann, W. Winiwarter, *J Cleaner Production*, 116 (2016) 80-89

An Ecological Approach to Structural Flexibility in Online Communication Systems.

M.J. Palazzi[1], C.A. Plata[2], V. Calleja-Solana[3], S. Meloni[3], A. Solé-Ribalta[1], S. Suweis[2], J. Borge-Holthoefer[1], (1) IN3, Universitat Oberta de Catalunya,Barcelona,Spain, mariapalazzi@uoc.edu. (2)IFISC (CSIC-UIB), Mallorca, Spain. (3)Dipartimento di Fisica e Astronomia G. Galilei, Università di Padova,Italy.

Human perceptual and cognitive abilities are limited resources. Today, in the age of cheap information -cheap to manipulate, to disseminate-, this cognitive bottleneck translates into hypercompetition for rewarding outcomes (acceptance, recognition) among actors. These incentives -summarised as visibility- pushes actors to mutualistically interact with specific memes, seeking the virality of their messages. Contents are driven by the evolving context, i.e. the chances to persist and reach widely are tightly subject to changes in the communication environment. Here, we show that the underlying architecture of the actors-memes interactions in information ecosystems, evolves towards emergent patterns, reminiscent of those found in natural ecosystems. By the analysis of Twitter data streams (**Fig. 1**(top-left)) we show that communication systems are structurally flexible, i.e. fluctuating from modular to nested, and back to modular, architectures as a response to environmental perturbations. We then propose an ecology-inspired modelling framework, bringing to light the precise mechanisms causing the observed topological reorganization **Fig. 1**(top-right). The model builds on the idea that the actor-meme network structure is effectively driven by an optimization process **[2]**, aiming at the maximization of visibility, and that each actor and meme specie has assigned a niche profile**[1]**, that represents their topical domain. Results from the numerical simulations predict -and the data confirm- that the users' struggle for visibility induces a re-equilibration of the network's mesoscale towards a transition from meso- to macroscale nested arrangements, top and bottom subpanels in **Fig. 1**(top-left) and (top-right), respectively. Finally, results at the microscale suggests that flexibility at the structural level is not mirrored at the dynamical one, where environmental shocks do leave a trace **Fig. 1**(bottom).

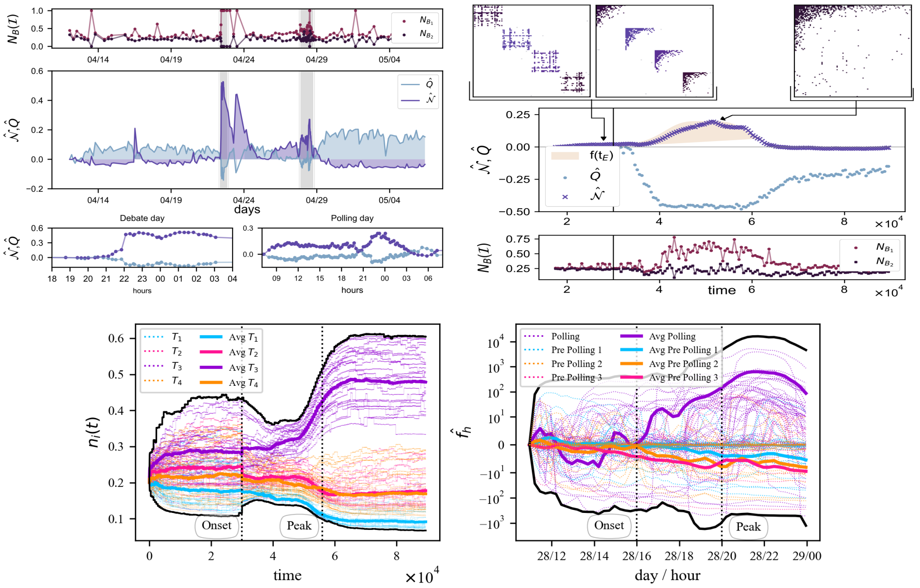


Figure 1: Top: Structural measures over time for a Twitter data stream from the Spanish general elections (2019) (left), and numerical simulations (right). Bottom: Evolution of abundances for the numerical simulations and the empirical data, respectively.

References

[1] Cai Weiran, et al., *Nat Comms*, 11, 5470, 2020. DOI: 10.1038/s41467-020-19154-5.

[2] S. Suweis, et al. *Nature*, 500, 449, 2013. DOI: 10.1038/nature12438.

[3] M. J. Palazzi, et al. *Nat Comms*, 12,1941, 2021. DOI: 10.1038/s41467-021-22184-2.

Shadowing and shielding: Effective heuristics for continuous influence maximisation in the voting dynamics

G. Romero Moreno[1]*, S. Chakraborty*[1]*, M. Brede*[1]*. (1) University of Southampton, Southampton SO17 1BJ, United Kingdom, Guillermo.RomeroMoreno@soton.ac.uk.*

Influence maximisation, or how to affect the intrinsic opinion dynamics of a social group, is relevant for many applications, such as information campaigns, political competition, or marketing. Previous literature on influence maximisation has mostly explored discrete allocations of influence, i.e. optimally choosing a finite fixed number of nodes to target. Here, we study the generalised problem of continuous influence maximisation where nodes can be targeted with flexible intensity. We focus on optimal influence allocations against a passive opponent and compare the structure of the solutions in the continuous and discrete regimes. We find that, whereas hub allocations play a central role in explaining optimal allocations in the discrete regime, their explanatory power is strongly reduced in the continuous regime. Instead, we find that optimal continuous strategies are very well described by two other patterns: (i) targeting the same nodes as the opponent (shadowing) and (ii) targeting direct neighbours of the opponent (shielding). Finally, we investigate the game-theoretic scenario of two active opponents and show that the unique pure Nash equilibrium is to target all nodes equally. These results expose fundamental differences in the solutions to discrete and continuous regimes and provide novel effective heuristics for continuous influence maximisation.

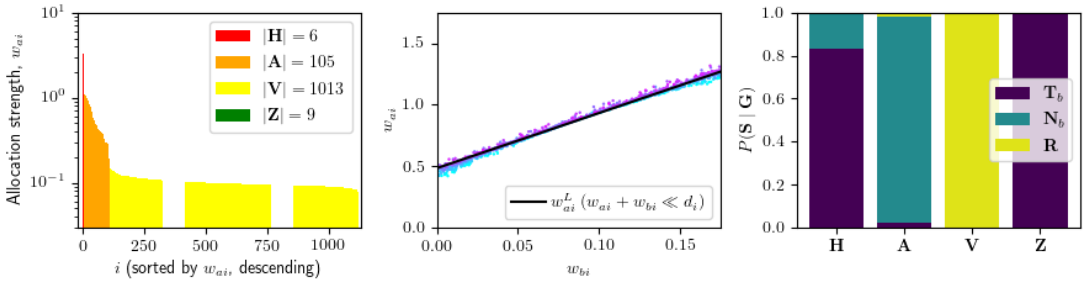


Figure 1: Left: Distribution of optimal influence allocations sorted in descending order when the passive controller discretely targets 16 randomly chosen nodes and both controllers hold the same budget. Center: Dependence of optimal allocations on the allocations of a passive controller when the value of external allocations are generally lower than node degrees. The active controller is in budget superiority and the passive controller targets continuously with weights randomly drawn from a uniform distribution. Numerical results are given by a cloud of points and analytical results are given by black curves. Right: Overlap between the allocation groups shown in the left panel and shielding groups (*Tb*: targeted by the opponent, *Nb*: neighbours of targeted by the opponent, *R*: rest) for a passive controller targeting 16 nodes in a discrete fashion.

References

[1] G. Romero Moreno, S. Chakraborty, M. Brede, *Shadowing and shielding: Effective heuristics for continuous influence maximisation in the voting dynamics*, PLOS ONE (2021), doi:10.1371/journal.pone.0252515.

Assessing disinformation through the dynamics of supply and demand in the news ecosystem

__P. Gravino__[1], G. Prevedello[1], M. Galletti[1], V. Loreto[1,2,3] (1) Sony Computer Science Laboratories, 75005 Paris, France, pietro.gravino@gmail.com (2) Sapienza University of Rome, Physics Department, 00185, Rome, Italy (3) Complexity Science Hub Vienna, A-1080 Vienna, Austria

Social dialogue, the foundation of our democracies, is currently threatened by disinformation and partisanship, with their disrupting role on individual and collective awareness and detrimental effects on decision-making processes. Despite a great deal of attention to the news sphere itself, little is known about the subtle interplay between the offer and the demand for information. Still, a broader perspective on the news ecosystem, including both the producers and the consumers of information, is needed to build new tools to assess the health of the infosphere. Here, we combine in the same framework news supply, as mirrored by a fairly complete Italian news database - partially annotated for fake news, and news demand, as captured through the Google Trends data for Italy. Our investigation focuses on the temporal and semantic interplay of news, fake news, and searches in several domains, including the virus SARS-CoV-2. Two main results emerge. First, disinformation is extremely reactive to people's interests and tends to thrive, especially when there is a mismatch between what people are interested in and what news outlets provide. Second, a suitably defined index can assess the level of disinformation only based on the available volumes of news and searches. Despite our results were mainly obtained on the *Coronavirus* keyword, we provide hints that the same findings can have more general applications. We contend these results can be a powerful asset to inform campaigns against disinformation and provide news outlets and institutions with potentially relevant strategies.

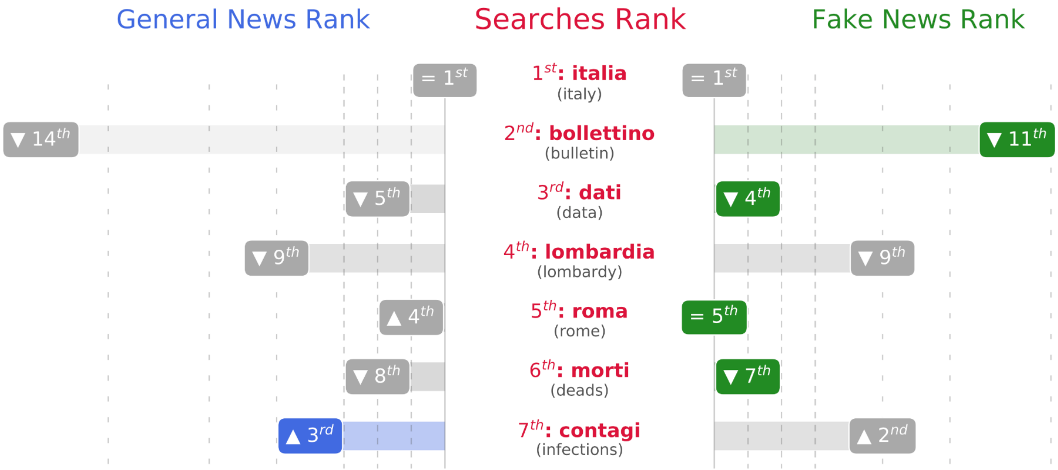


Figure 1: The most searched coronavirus-related keywords ranked by information demand (center, red). On the sides, tags indicate the rank in General News (left) and in Fake News (right). Tags are distanced from the center by the amount of rank mismatch with respect to Searches rank. Tags are colored to highlight the rank closest to the Searches rank: blue for General News and green for Fake News.

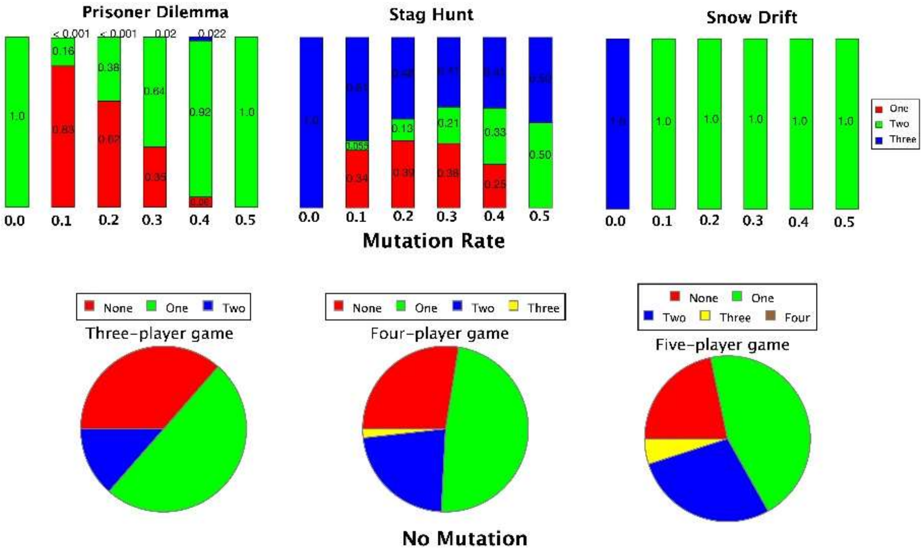
Acknowledgments

The Authors wish to warmly thank AGCOM for providing the database of Italian news. The database was shared in the framework of the Task Force on "Digital Platforms and Big Data - Covid-19 Emergency", established by the AGCOM to contribute, among other things, to the fight against online disinformation on issues related to the SARS-CoV-2 crisis.

Statistics of equilibria points in pairwise social dilemmas under replicator-mutator dynamics

M.H. Duong[1]*, T. A. Han*[2] *(1) University of Birmingham, Birmingham, UK, Email: h.duong@bham.ac.uk (2) Teesside University, Middlesbrough, UK.*

The replicator-mutator equation is a set of differential equations describing the evolution of frequencies of different strategies in a population that takes into account both selection and mutation mechanisms. It is a fundamental mathematical framework for the modelling, analysis and simulation of complex biological, economical and social systems and has been utilized in the study of, just to name a few, population genetics, autocatalytic reaction networks, language evolution and the evolution of cooperation, see e.g. [1]. In this talk we discuss our recent works [2,3] on the statistics of the equilibria of the replicator-mutator equation. For two-player social dilemmas (namely the Prisoner's Dilemma, Snow Drift, Stag Hunt and Harmony), we characterize (stable) equilibrium points and analytically calculate the probability of having a certain number of equilibria when the payoff entries are uniformly distributed. For multi-player random games whose pay-offs are independently distributed according to a normal distribution, by employing techniques from random polynomial theory, we compute the expected or average number of internal equilibria. In addition, we perform extensive simulations by sampling and averaging over a large number of possible payoff matrices to compare with and illustrate analytical results. Our results provide new insights into the behavioural diversity of dynamical systems, including biological, social and artificial life ones.



(Upper row) impact of mutation rate on the probability of having a certain number of internal equilibrium points for three popular social dilemmas (two-player two-strategy games). (Lower row) probability of having a certain number of internal equilibrium points for two-strategy games with three, four and five players without mutation.

References

[1] Komarova NL (2004) Replicator-mutator equation, universality property and population dynamics of learning. J Theor Biol 230(2):227–239.

[2] Duong MH, Han TA (2020) On Equilibrium Properties of the Replicator–Mutator Equation in Deterministic and Random Games. *Dyn Games Appl* 10**,** 641–663

[3] Duong MH, Han TA (2021) Statistics of equilibria of social-dilemmas with mutation, In preparation.

The dynamics of faculty hiring networks

Eun Lee, Daniel B. Larremore, and Aaron Clauset
The University of Colorado Boulder, Boulder, Colorado 80302, USA

Faculty hiring networks---who hires whose graduates as faculty---exhibit sharp inequalities [1]. These inequalities in faculty hiring reinforce other inequalities: from resource distributions [5, 6] to the productivity of individual careers [1]. As a result, the demographics of a few top PhD programs shape both the demographics of the entire fields [2] and the most popular research topics [3,4], slowing efforts to diversify academia more broadly. Understanding the mechanisms driving these patterns would inform new efforts to diversify the academy and scholarship.

Here, we investigate mechanisms that can and cannot describe observed structural inequality and network characteristics in current faculty hiring networks. To this end, we consider adaptive rewiring network models which reinforce institutional prestige via five potential mechanisms. In each time step, a random retirement creates an open position, and a new hiring follows one of the mechanisms to fill the position (Fig.1). Each mechanism determines the probability of the hiring from an institution based on a preference toward a type of institutional prestige which is inferred from the existing hiring network structure. This process of retirement and hiring is continued until the network reaches a statistically stable state to see how much the stabilized system is similar (dissimilar) to the current hiring system.

Among the five mechanisms, we found that the hiring mechanism which favors toward a department placing its graduates to many places best captured a number of characteristics of real faculty hiring networks, including Gini coefficient, steepness of linear hierarchy, and centrality distributions. On the other hand, the hiring mechanism which favors previously hired faculty's institutions outperforms in describing the perception bias on the representation of institutions in a field. Here, the perception bias denotes a possible error between a local and network-level visibility of an institution based on the social sampling theory [7]. This competing result suggests that the underlying mechanism reinforcing strong hierarchies in the faculty hiring may be a mixture of global and local preference in the institutional prestige. Furthermore, we observed that the less influential institutions can suffer from the `Ranking Trap' – for them, improving their ranking in a system preferring the institutional prestige could be almost impossible. We hope our study can be a steppingstone that stimulates a systematic analysis on the reinforcing mechanism of hierarchical prestige in academics on the way of improvement of diversity in the system.

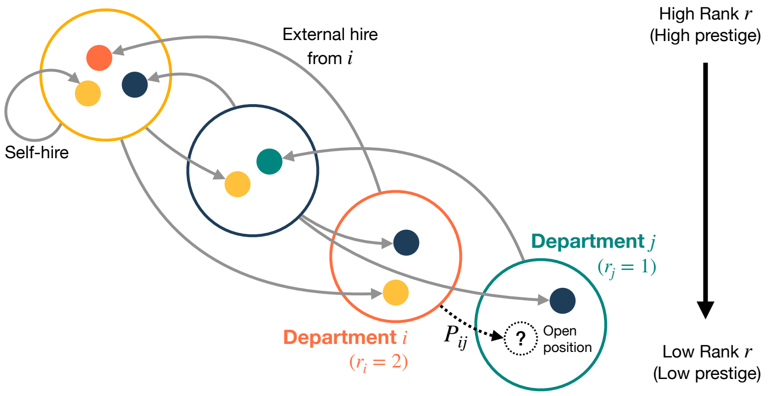


Fig. 1. Schematic figure for a faculty hiring model. Each color-bounded circle represents a department, and color-filled circles represent Ph.D. graduates from the institution having same color- bounded circle. Grey directed arrows denote exchanges of graduates between departments. When a faculty retired, it makes an open position at an institution j. The institution j needs to hire a faculty based on the hiring model preferring institutional prestige.

[1] A. Clauset, S. Arbesman, and D. B. Larremore, Systematic inequality and hierarchy in faculty hiringnetworks, Science Advances1, e1400005 (2015).
[2] S. F. Way, D. B. Larremore, and A. Clauset, Gender, Productivity, and Prestige in Computer ScienceFaculty Hiring Networks, , 1169 (2016).
[3] B. Hofstra, V. V. Kulkarni, S. M.-N. Galvez, B. He, D. Jurafsky, and D. A. McFarland, The Diver-sity–Innovation Paradox in Science, Proceedings of the National Academy of Sciences , 201915378(2020).
[4] A. C. Morgan, D. J. Economou, S. F. Way, and A. Clauset, Prestige drives epistemic inequality in thediffusion of scientific ideas, EPJ Data Science7, 40 (2018), 1805.09966.
[5] B. Pusser and R. Geiger, Knowledge and money: Research universities and the paradox of the mar-ketplace, Academe90, 115 (2004).
[6] S. Cole and J. R. Cole, Visibility and the Structural Bases of Awareness of Scientific Research,American Sociological Review33, 397 (1968).
[7] K. Lerman, X. Yan, and X.-Z. Wu, The "Majority Illusion" in Social Networks, PLOS ONE11,e0147617 (2016).

Evolutionary game model of risk propensity in group decision making

A. Civilini[1], *N. Anbarci*[2], *V. Latora*[1,3,4,5]. *(1) Queen Mary University of London, London E1 4NS, United Kingdom, a.civilini@qmul.ac.uk (2) Durham University, Durham DH1 3LB, United Kingdom, (3) The Alan Turing Institute, London NW1 2DB, United Kingdom, (4) Complexity Science Hub Vienna, A-1080 Vienna, Austria, (5) Università di Catania and INFN, Catania I-95123, Italy*

Choice dilemmas describe the most general situations in which decision makers are faced with an alternative between two possibilities: a risky strategy, that either brings a high reward with a probability w_p, or a low one with probability $1 - w_p$, and a safe strategy with an intermediate reward. Expected Utility Theory and its strategic version Game Theory assume that rational decision makers always act to maximize their private utilities/payoffs. However, two anomalous behaviours emerge when decision makers take decision not as single entities but in social groups. The first anomaly is the herd behaviour leading to irrational outcomes, observed for example during financial bubbles, where decision makers can follow the opinion of others, apparently disregarding their own interests. The second anomaly is group polarization (also known as group choice shift), where the average opinion of individual decision makers is exacerbated when they act in group. In our work we propose to use evolutionary game theory to model group choice dilemmas in a population of interacting decision makers represented as the nodes of a hypergraph, whose hyperedges describe the interactions in groups of two or more agents. The numerical simulations on structured population have been compared to an analytically treatable mean-field version of the model dynamics. We show that group choice shifts emerge naturally in our model, and we are able to explain how they depend on the mechanisms of preference aggregation and on the structure of the interaction network (Fig. 1a). In particular, when implemented on heterogeneous structured populations, our model predicts the spontaneous emergence of irrational herd behaviours towards the riskier strategy (Fig. 1b), where the adoption of a strategy is followed by a collapse of the average utility (Fig. 1c). By unveiling the role played by the structure of the hypergraph on the emergence of irrational and risky behaviours, our model opens new research avenues to explain the appearance of extreme collective behaviours in social systems.

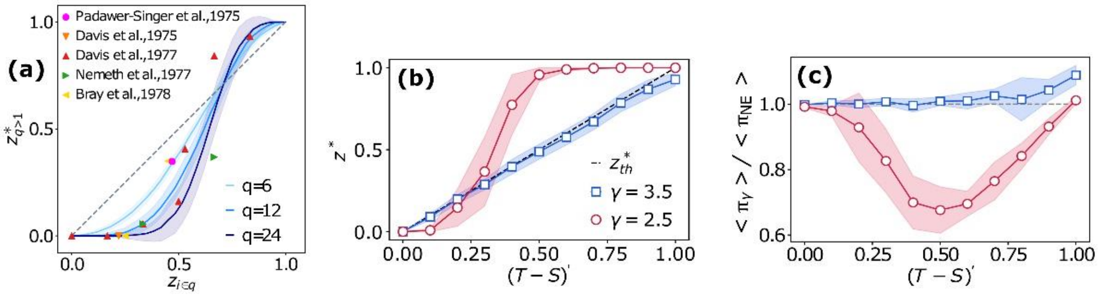


Figure 1: Model predictions for empirically observed extreme collective behaviours.

(a) Group polarization (GP) from numerical simulations compared to empirical data from mock trial juries (coloured dots). Counter-intuitively, our model predicts that GP increases with q, the group size. **(b)** For populations with scale-free co-membership degree distribution (red circles), the computed quasi-stationary distribution deviates from z^*_{th}, the Nash Equilibrium derived analytically in mean-field approximation. **(c)** This emergent herd behaviour is irrational from an Expected Utility Theory perspective, since the collective adoption of the risky strategy is followed by a collapse of the average utility/payoff.

Opioid addiction recovery on Reddit: behavioral shift and social support

D. Balsamo[1], *C. Monti*[2], *G. De Francisci Morales*[2], *P. Bajardi*[2], *R. Schifanella*[1,2]. *(1) University of Turin, Turin, Italy, duilio.balsamo@unito.it. (2) ISI Foundation, Turin, Italy*

Individuals affected by *Substance Use Disorders* often participate to mutual help groups like Alcoholics Anonymous and Narcotics Anonymous to receive the support they need to proceed in a radical change of life, detoxification. Nowadays, these individuals often call on online communities to search for advice and peer support, like in the case of Reddit. In this work, we focus on Reddit users who share their opioid detoxification process on the community *r/OpiatesRecovery*, with the goals of uncovering the behavioral changes that undergo users' recovery paths on the platform, and of understanding the group dynamics that help such users in the process, such as social feedback and social support provided by the community. We estimated the starting date of detoxification t_0 for a large number of users in the community by training a machine learning model to classify submissions containing self-reports of their temporal progress in detoxification. In order to measure the presence of behavioral shift, we computed several day-by-day metrics of behavior based on activity and content, aggregating the results by aligning users' posting timelines on t_0, and assessing statistical differences before and after the start of detox. We considered both engagement features, such as activity and participation to specific communities, as well as social feedback features, like the amount of support, trust, and similarity [1] expressed by the users and received by the community. The presence of shift in online behavior was statistically assessed both by performing *interrupted time series analysis* on the data, observing the sign and the statistical significance of the level change, and by comparing the population average of the difference between each feature before and after t_0, with a bootstrapped data. Our results show that the users who face detoxification abruptly lower their participation in opioid-consumption-related subreddits, such as *r/opiates,* and concentrate on recovery, while experiencing a significant increase in support, trust, and status received by the community (Figure 1).

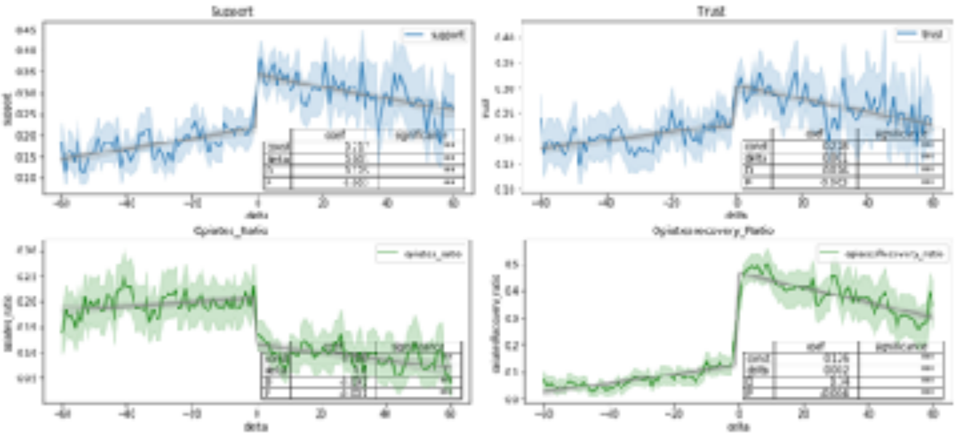


Figure 1: Interrupted time series of Support and Trust received (blue) and of participation to r/opiates and r/opiatesRecovery.

References

[1] M.Choi, et al., *Proceedings of The Web Conference 2020*, pages 1514–1525, 2020.

Modelling network dynamics with Markov chains to understand structures in online communities: cross comparison of climate change and vaccine

V. Chomel[1,2]*, M. Panahi*[2]*, D. Chavalarias*[1,2]*.* victor.chomel@iscpif.fr

(1) Institut des Systèmes Complexes Paris Ile-de-France (ISC-PIF), CNRS, Paris, France

(2) Ecole des Hautes Etudes en Sciences Sociales (EHESS), Paris, France

Communities are major keys to understanding the spread of information on social networks. Unveiling the network structure helps to highlight the role within the communities of some activist accounts. Through the application of a new method based on Markov chains, metrics of community dynamic such as turnover, activism or social mobility inside the community can be computed. Based on these indicators, information dynamics are compared on two radically different debates: climate change and vaccines. This study is based on Twitter data collected since 2019 for the climate change debate (160 million tweets and retweets) and on the French one on vaccine since 2017 (more than one million tweets).

For the first issue, the debate seems to be settled with two well-defined communities and gatekeepers setting an information agenda. For the second one, even though vaccine is an old debate, Covid puts a new light on the topic, changing the boundaries of the debate. Unlike climate change, communities are quickly evolving leading to the appearance of new communities with more extreme opinions on the topic. We will compare the impact on narratives of these two kinds of structures and how it can be linked to misinformation patterns. This analysis enables us to suggest possible solutions to fight against misinformation online.

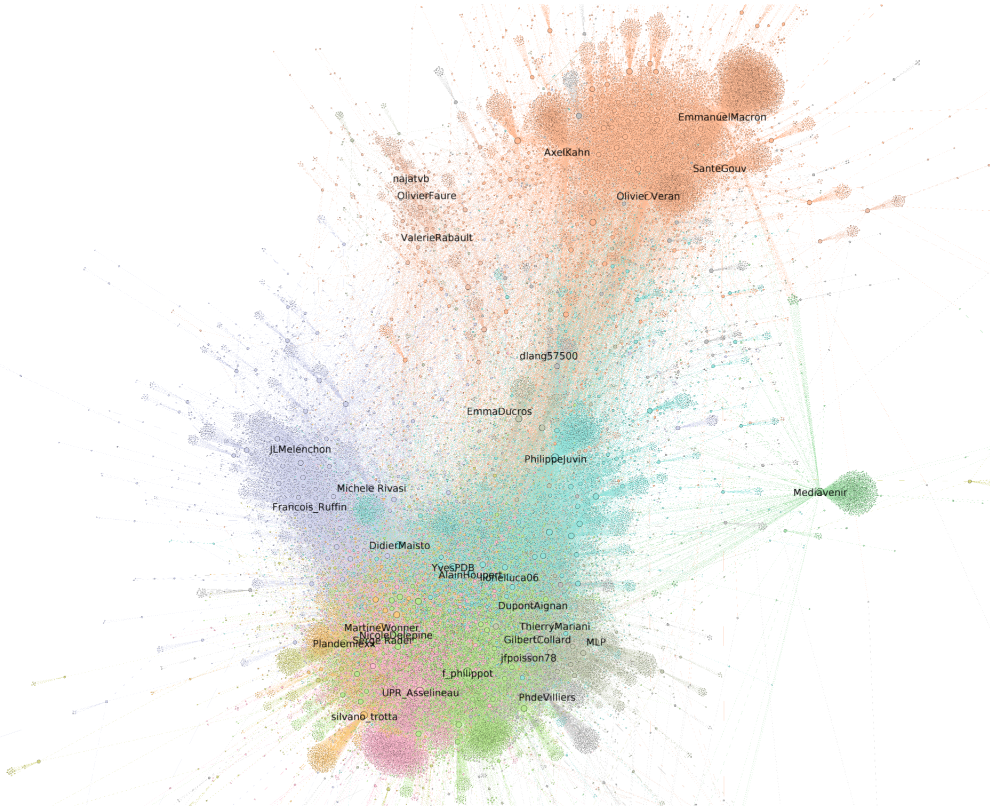
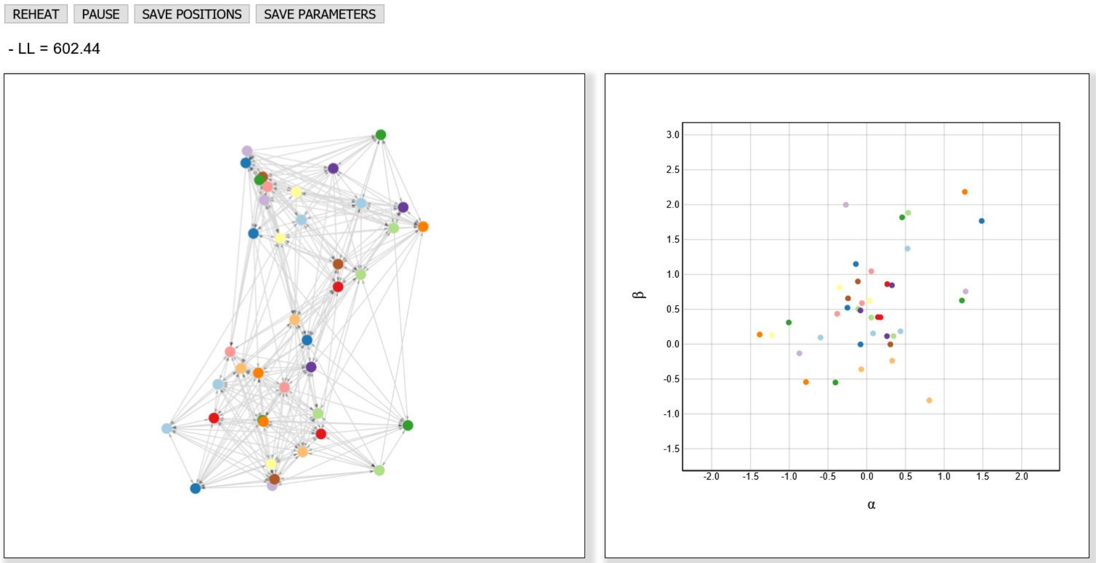


Figure 1: Retweet network of the French debate about vaccine in 2020

Understanding force directed layouts through latent space models

F. Gaisbauer*, A.Pournaki, S.Banisch,E. Olbrich Max Planck Institute for Mathematics in the Sciences, Leipzig, Germany, felix.gaisbauer@mis.mpg.de*

This contribution brings together two strands of research: Latent space approaches to network analysis and force-directed layout algorithms. The former can be considered as extensions of spatial random graph models for social networks, which have the goal of embedding a graph/network in an underlying social space [1]. They have recently been employed successfully in the estimation of (one-dimensional) ideological positions from social media data via Bayesian ideal point estimation [2]. The latter are used ubiquitously for data exploration, illustration, and analysis. Nevertheless, an interpretation of the outcomes of graph drawings with force-directed algorithms is not straightforward [3]. We argue and show that interpretability can be provided by latent space approaches where distances between nodes can be interpreted as proximity/separation in this space. To this end, we derive a force-directed layout algorithm on the basis of explicit latent space models of network interactions. The positions of the actors in an assumed latent space influence the probability of ties between them - the closer the positions of the actors, the more probable it is that they form a tie. We show that force-directed layout algorithms can be considered as maximum likelihood estimators of the model. We also present ready-to-use implementation of the layout algorithm and show its application to Twitter retweet networks.



- LL = 602.44

Figure 1: Screenshot of the interactive implementation of the force-directed layout. Left: inferred latent positions of the nodes. Right: additional parameters α and β for each node, which represent activity and popularity of the nodes.

References

[1] P. D. Hoff, A. E. Raftery, and M. S. Handcock (2002). Latent space approaches to social network analysis. Journal of the American Statistical association, 97(460), 1090-1098.
[2] P. Barberá (2015). Birds of the same feather tweet together: Bayesian ideal point estimation using Twitter data. Political analysis, 23(1), 76-91.
[3] T. Venturini, M. Jacomy, P. Jensen (2021). What do we see when we look at networks. Visual network analysis, relational ambiguity, and force-directed layouts. Big Data & Society 8(1), 20539517211018488.

Dynamic communities and engagement in public debate about Covid-19 in the German Twittersphere

Yasaman Asgari [1], Armin Pournaki [2], Eckehard Olbrich [2], Felix Gaisbauer [2]

(1) Sharif University of Technology, P.O. Box 11165-9161, Tehran, Iran, y.asgari@student.sharif.edu

(2) Max Planck Institute for Mathematics in the Sciences, Inselstrasse 22, Leipzig

Twitter as a popular social medium is not immune to Covid-19. This contribution investigates engagement in the German debate about the recent Covid-19 pandemic and state measures on Twitter over time from February to August 2020. The approach used is a dynamical extension of a method that has been established for cross-sectional data sets [1] based on a user-centered tweet collection around prominent German scientists, news outlets, and politicians. We find that the number of users interacting with each other is increased by the rise of the pandemic, and public health experts (Health ministers, virologists, and medical doctors) became influencers, while before the disease, they were important only within their scientific communities. Combining force-directed layouts and modularity clustering of 10-day time slices of retweet networks, we find that opposition to the state measures in Germany quickly becomes visible on Twitter. Over time, the opposition on the platform grows in size and becomes diversified in the sense that a weak separation within the cluster is visible, while the cluster of users who are less fundamentally critical of the state remains structurally diverse throughout. Furthermore, origin-destination flows between the communities detected are studied to determine the users in the evolving opposition. Then we investigate their engagement in reply sections of the seed users. Moreover, we go beyond the referenced method by also analyzing tweet content via topic modeling; we investigate the difference in topic usage of the different modularity clusters in the retweet network.

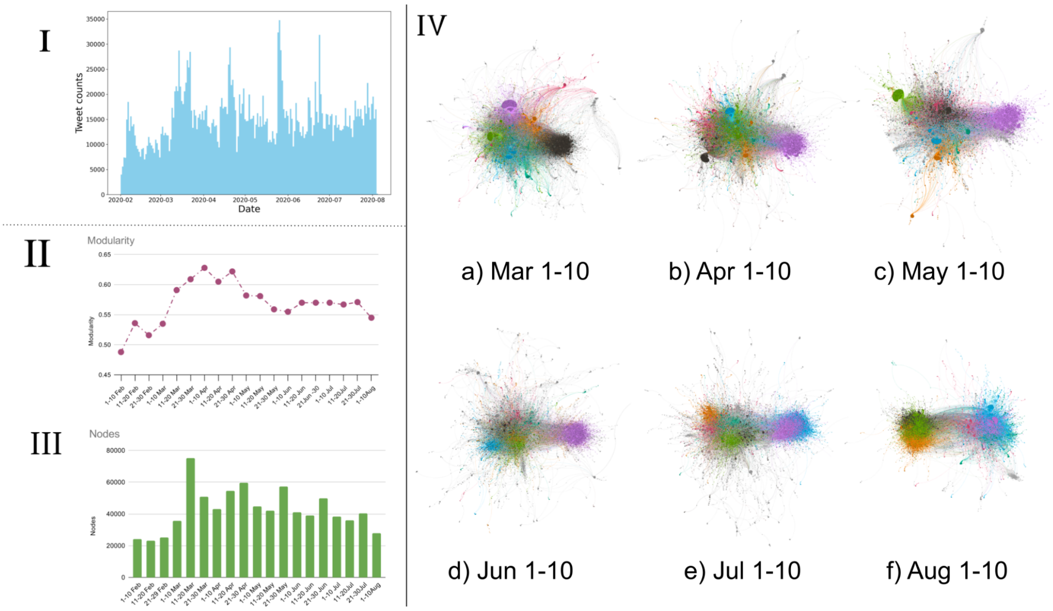


Figure 1: Activity statistics and dynamics of retweet network properties :Number of tweets over time(I), Modularity of 10-days retweet network (II), Number of nodes of 10-days retweet network (III), Force-directed layout and modularity clustering of retweet networks over time; opposition pole is always on the right (IV).

References

[1] F. Gaisbauer, A. Pournaki, S. Banisch, and E.Olbrich, *PLOS ONE 16, no. 3*

Social Network Reorganization during Emergency Lockdown

L. Napoli[1]*, V. Sekara*[2]*, M. Garcia-Herranz*[3]*, M. Karsai*[1] *(1) Department of Network and Data Science, Central European University, Quellenstraße 51, 1100 Vienna, Austria,* Napoli_Ludovico@phd.ceu.edu *(2) Department of Computer Science, IT University of Copenhagen, Rued Langgaards Vej 7, 2300 Copenhagen, Denmark (3) UNICEF Office of Innovation, 3 United Nations Plaza, New York, NY 10017, USA*

Socioeconomic status is a crucial driver during the structural emergence of social networks, as it has been observed with traditional tools for many decades. However, the role of socioeconomic status in the evolution of large-scale social networks represents an open challenge for modern network science, which can be tackled with relevant large digital data available recently. In particular, the impact of exogenous shocks, like the COVID-19 outbreak and related response policies, on the socioeconomic structure of a large-scale social network still lacks profound comprehension. Here, we use a mobile phone dataset from Sierra Leone, referring to the COVID-19 crisis period, and couple it with a fine-grained socioeconomic map, to obtain a large-scale socioeconomic network.

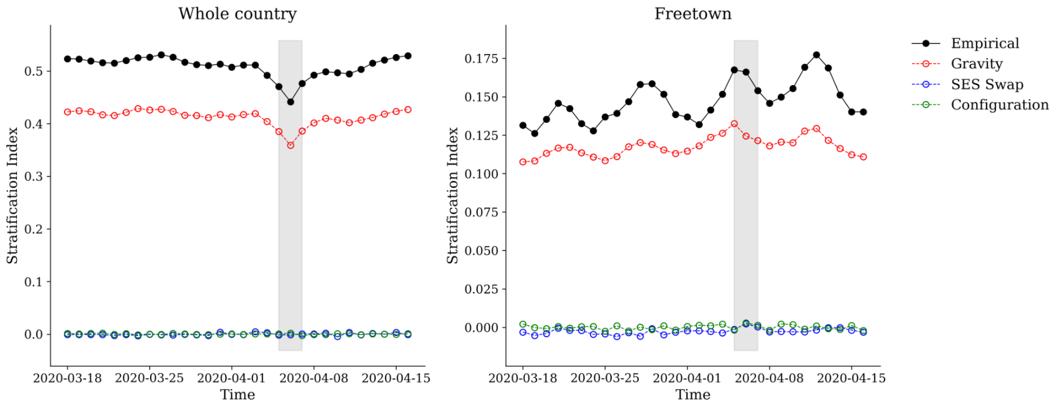


Figure 1: The Stratification Index [2] of the communication network from Sierra Leone measured day by day, with a symmetric three-day rolling time window, during the four-week period of observation. Analysis on the whole country (left) and within the capital Freetown (right). The different curves refer to the empirical measurement (black) and to the values obtained from three reference models: a gravity model (red), a multiple random swap of individual socioeconomic status (blue), and a configuration model (green). The grey area indicates the three days of full lockdown.

With this combined dataset, we analyze the impact of a three-day full lockdown, implemented by the government in response to the pandemic outbreak. We dynamically observe the social-economic network structure where individuals are assigned with a relative wealth indicator and are connected if they call each other on a given day. Via matrix measures [1] we observe how socioeconomic stratification changes locally (within the capital Freetown) and globally (over the whole country). By measuring the diagonality of these matrices with the Stratification Index [2] we observe significant patterns of stratification. Further, we find that while social segregation increases locally during the lockdown period, it decreases globally, as evidenced in Fig. 1, where we compare its dynamics with reference measures generated by the Gravity model or random reference network models. We show that the main reason behind the global decrease is the relative increase of communications between the richest (Freetown) and the poorest rural areas of the country.

References

[1] Y. Leo et al., *Journal of The Royal Society Interface,* 13 (2016) 125

[2] X. Dong et al., *EPJ Data Science,* 9 (2020) 20

Aging in a Schelling Segregation Model

David Abella[1], *Maxi San Miguel*[1], *Jose J. Ramasco*[1]. *(1) Instituto de Física Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), Campus Universitat de les Illes Balears, 07122 Palma de Mallorca, Spain. david@ifisc.uib-csic.es.*

There are several variants of the Schelling segregation model. Here we consider the effect of aging in a noisy, constrained, long range model in a 2d-lattice [1]. Aging [2] is understood as a property of the agents such that they are less prone to change their location as they get older in a satisfying location. The results are counterintuitive: aging favors segregation and increases the final number of satisfied agents. The original model exhibits 3 phases when increasing the tolerance: frozen, segregated and mixed. The segregated phase is modified by aging: i) The power law characterizing coarsening shows an exponent lower than for the original Schelling Model, implying much slower dynamics, ii) The final configuration has a larger interface density. More important, in the presence of aging, the mixed phase is replaced by a partially segregated phase. This new phase is characterized by a rather large number of clusters in the final configuration as compared to the two dominant clusters of the segregated phase, by a rougher interface and also shows a smaller coarsening exponent (see Fig. 1). The results of this work will be available soon at [3].

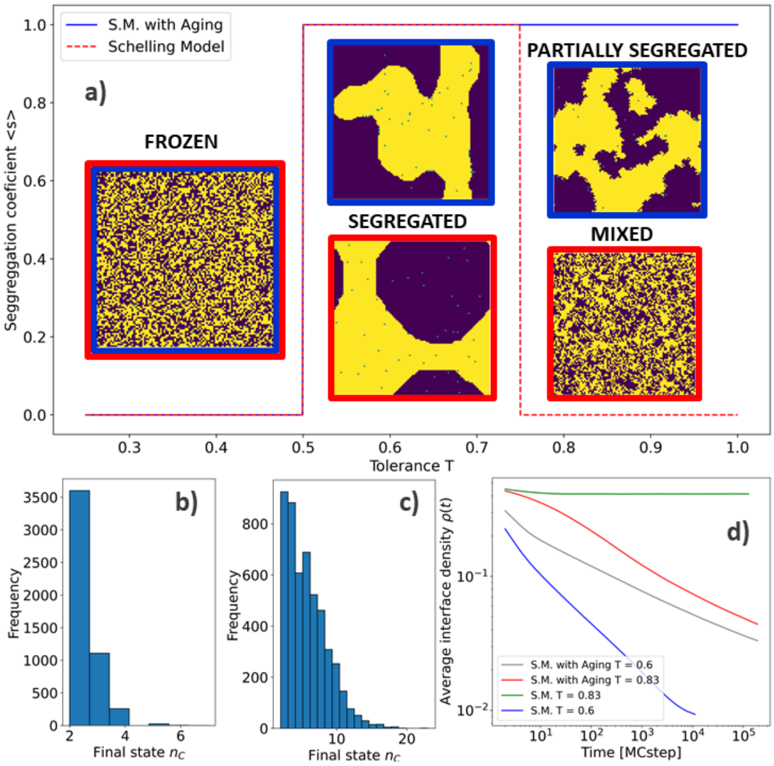


Figure 1: (a) Phase diagram of the Schelling Model (red line) and with aging (purple line). Segregation coefficient $s = 0$ indicates disorder and $s = 1$, segregation. Insets show snapshots of asymptotic states for the different phases for the original model (framed in red) and with aging (blue). **(b)** Cluster size distribution with aging for T = 0.6 (segregated phase) and **(c)** T = 0.83 (partially segregated) **(d)** Coarsening power laws for the average interface density.

References

[1] Gauvin, Laetitia et al. (2009). *The European Physical Journal B*, 70(2), 293-304.

[2] Artime, Oriol, et al. (2018). *Physical Review E* 98.3: 032104.

[3] Abella, David, et al. (2021), Aging in a Schelling Segregation Model, in preparation.

Collective Memory with Endogenous Recall Effects in Wikipedia

N. Igarashi[1], *Y. Okada*[2], *Y. Sano*[3], *(1)(2)(3) University of Tsukuba, 1-1-1 Tennodai, Tsukuba, Japan,*
naokigarashi@gmail.com

Memories such as wars and earthquakes in society are called “collective memory” [1]. Some studies have focused on decay in collective memory and modeled the temporal decay of collective memory as a bi-exponential function based on communicative and cultural memory by using citation, patents, music, and Wikipedia data [2].

Since temporal decay of collective memory is not only caused by the one-way forgetting but also by the recall effect, however, this recall effect has not been received much attention. Kanhabua et al. showed that when a terrorist attack or other event occurs, similar events in the past also will be recalled, and Wikipedia views increase [3]. In this case, Kanhabua et al. analyzed the cases where it is recalled by external news.

Here we analyzed collective memory decay with endogenous recall effect by using Wikipedia page view. Figure 1 shows the monthly change in Wikipedia page views over time for global celebrities who died in 2016, starting four months before they died. Obviously, there are large peaks every 12 months due to the yearly anniversary, probably by numerous mass media reports. Moreover, there are also small peaks one month before every large peak that are non-trivial. We consider this small spontaneous increase to be the endogenous recall effect in collective memory.

In this study, by using the time series data of Wikipedia page views on events with spontaneous recall effects (e.g., memory decay due to obituaries), we propose an extended model of temporal decay of collective memory by adding a “recall term“ to the temporal decay model. Finally, We anticipate our results of the investigation using a mathematical approach with social media to lead to a greater affinity between sociology and science regarding collective memory.

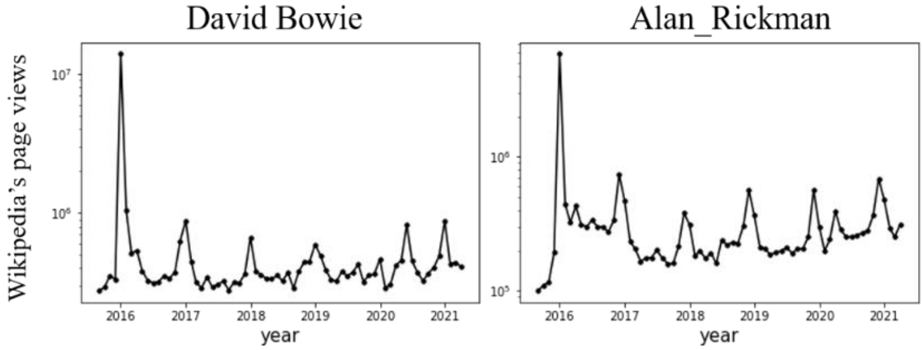


Figure 1: Monthly Wikipedia page views for global celebrities who died in 2016

[1] Halbwachs, M. (2020). *On collective memory*. University of Chicago Press.

[2] Candia, C., Jara-Figueroa, C., Rodriguez-Sickert, C., Barabási, A. L., & Hidalgo, C. A. (2019). The universal decay of collective memory and attention. *Nature Human Behaviour*, *3*(1), 82-91.

[3] Kanhabua, N., Nguyen, T. N., & Niederée, C. (2014, September). What triggers human remembering of events? A large-scale analysis of catalysts for collective memory in Wikipedia. In *IEEE/ACM Joint Conference on Digital Libraries* (pp. 341-350). IEEE.

Detecting Transition Points in Adaptive Social Networks: Large-Scale Numerical Simulations and Neural Network-Based Visualization

H. Sayama[1,2] *(1) Center for Collective Dynamics of Complex Systems, Binghamton University, Binghamton, NY, USA; (2) Waseda Innovation Lab, Waseda University, Tokyo, Japan; sayama@binghamton.edu.*

Understanding social fragmentation transition, i.e., transition of social states between many disconnected communities with distinct ideas and a well-connected single network with homogeneous ideas, is a timely research topic with high relevance to various current societal issues. We had previously studied this problem using numerical simulations of adaptive social network models [1] and found that two individual behavioral traits, *homophily* and *attention to novelty*, among others, had the most significant impact on the outcomes of social network evolution. However, our previous study was limited in terms of the range of parameter values examined, and possible interactions between multiple behavioral traits were largely ignored.

In the present study, we have conducted a substantially larger-scale numerical experiment with an expanded parameter sweep range by an order of magnitude in each parameter dimension, resulting in a total of 116,640 simulation runs. To capture nontrivial interactions among behavioral parameters, we modeled and visualized the outcome dependence on parameters using neural networks. Results (Figure 1) have shown that, while the competition between homophily (h) and attention to novelty (a) is the primary determinant of social fragmentation in a low-conformity (c) regime (left), another transition plane emerges at an intermediate homophily level in a high-conformity regime (right), which was not previously known. This implies that social homogenization may be caused by more than one behavioral mechanism.

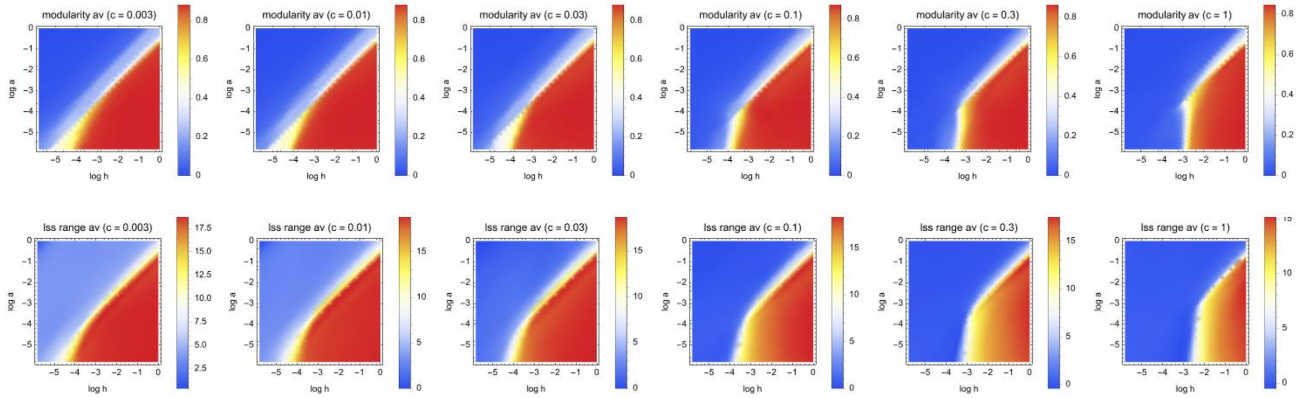


Figure 1: Phase diagrams of adaptive social network evolution. Results of large-scale parameter-sweep numerical simulations were modeled and visualized using neural networks. Each plot shows outcome dependence on homophily (h, horizontal axis), attention to novelty (a, vertical axis) and conformity (c, varied from left to right). **Top row:** How final network topology (modularity) depends on h, a and c. **Bottom row:** How final idea diversity (range of average community states) depends on h, a and c. Red and blue regions correspond to fragmented and homogenized network states, respectively. Number of nodes $n = 300$.

Acknowledgments

This work was supported in part by JSPS KAKENHI Grant Number 19K21571.

Reference

[1] Sayama, H. (2020). *Artificial Life Conference Proceedings* (pp. 113-120), MIT Press. https://doi.org/10.1162/isal_a_00349 (open access)

Complex systems approach to social sciences empirical research – a pilot for future data acquisition and analysis

F. Peter[1,2,*], *R. Cárdenas*[3], *J. Perelló*[1,2,*]. *(1) OpenSystems Research Group, Departament Física de la Matèria Condensada, Universitat de Barcelona, Martí i Franquès, 1, 08028 Barcelona, Spain, fpeter@ub.edu. (2) Universitat de Barcelona Institute of Complex Systems . (3) Universidad Yachay Tech, Hacienda San José s/n y Proyecto Yachay, San Miguel de Urcuquí, Ecuador.*

The exploration of biological systems by means of mathematical modeling not only elicited many important concepts in chaos theory and complex systems theory, but it also thoroughly revolutionized our understanding of living matter. Right now, the social systems and more particularly social interactions are being revolutionized in this same fashion. Complex systems science already achieves considerable insights by exploring large sets of meta-data, often prerecorded for other purposes, such as social media data, phone calls or urban structures. We propose an experimental setup that allows long time recording of large data sets with the ethical agreement of the human subjects participating and with large freedom in the design of the experiment and therefore, the research question. We present a pilot experiment in which we measure opinion heterogeneity and clusters of opinion profiles in a three months recording of two hundred participants. The participants answered binary opinion questions on social, political, and philosophical matters that were created in a citizen science framework. We present statistical physics and complex systems analysis methods that enrich the long-standing tradition of plain surveys in sociology and psychology. First results from a one-year follow-up experiment with several hundreds of participants and a narrowed scope of questions are presented.

Acknowledgements

This work was partially supported by MINEICO (Spain), Agencia Estatal de Investigación (AEI) and Fondo Europeo de Desarrollo Regional (FEDER) through grant PID2019-106811GB-C33 (FP and JP); by Generalitat de Catalunya (Spain) through Complexity Lab Barcelona (contract no. 2017 SGR 608; FP and JP); by the European Union’s Horizon 2020 research and innovation programme (CoAct, under grant agreement no. 873048; FP and JP).

On the resilience of extremism: the heterogeneous Hegselmann-Krause model in networks

R. Perrier, H. Schawe, L. Hernández. Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS, CY Cergy Paris Université, France.

The role of heterogeneous confidences in the Hegselmann-Krause (HK) bounded confidence model has been extensively studied recently in a mixed population setting revealing a complex phase diagram with a reentrant phase of consensus for confidence values where the homogeneous model leads to fragmentation [1]. On the other hand, the homogeneous HK model in networks has also been studied [2] and a very recent work shows how networks foster consensus [3]. However, the outcomes of the HK dynamics combining confidence and neighbourhood heterogeneities are still unclear.

In this work we present a detailed study of HK bounded confidence opinion dynamics with heterogeneous confidence ε_i drawn from uniform distributions in different intervals $[\varepsilon_l, \varepsilon_u]$ on different topologies, from fully connected to sparse networks.

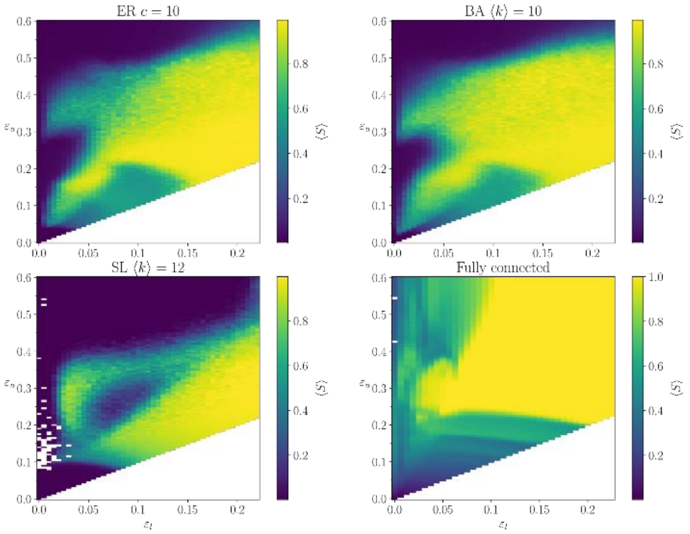


Figure 1. Phase diagram: the influence of different levels of heterogeneity on the consensus in a society for various topologies. $\langle S \rangle$: fraction of agents which have the majoritarian opinion, i.e., $\langle S \rangle \approx 1$ signals consensus in the society, while smaller values signal a fragmented opinion landscape. 4455 data points, each point gives the interval of the heterogeneities of the considered sample, results are averaged over 100 samples. N = 16384 agents. ER = Erdős-Rényi; BA = Barabási-Albert; SL = Square Lattice

The exploration of the whole parameter space, along with an extensive finite-size analysis in the region where very close minded agents are present, $\varepsilon_l = 0.05$, reveal rich, highly complex and non-monotonous behaviors. Specifically, we uncover a phenomenon of phase coexistence induced by the topology, where both the size of the largest opinion cluster and the actual value of the majoritarian opinion have to be taken into account. We show that consensus is lost faster when increasing ε_u in the phase associated to the majoritarian opinion being *mild* than in the phase associated to the majoritarian opinion being *extreme*.

References:

[1] H. Schawe, L. Hernández, Scientific Reports volume 10, Article number: 8273 (2020)
[2] S. Fortunato, International Journal of Modern Physics C16, 259 (2005)
[3] H. Schawe, S. Fontaine and L. Hernández, Phys. Rev. Research to be published (2021) arXiv:2102.10910v2

A little Knowledge Is A Dangerous Thing: Excess Confidence Explains Negative Attitudes Towards Science

S.Lackner[1]*, F. Francisco*[2]*, J.Gonçalves-Sá*[1,3]*. (1) LIP – Laboratório de Instrumentação e Física de Partículas, Lisboa, Portugal, simone.lackner@lip.pt (2) Faculdade de Ciências da Universidade do Porto, Porto, Portugal, (3) Physics Department, Instituto Superior Técnico, Lisboa, Portugal*

Scientific knowledge is accepted as the main driver of development, allowing for longer, healthier, and more comfortable lives. However, an interesting paradox can be observed: the rise in science communication and outreach efforts seems to correlate with the rise of openly anti-science groups. This has serious social consequences, from the anti-vaccination to "post-truth" movements. Such fears are particularly relevant now, as trust in science is fundamental to fight the COVID19 pandemic. It has been argued that overconfidence (broadly defined as a tendency to have a subjective assessment of one’s own skills, intellect, or talent, that is reliably greater than the objective accuracy of such skills, intellect, or talent), might play an important role in these movements. If confidence in one’s knowledge grows faster than knowledge itself (Dunning-Kruger effect), this could create a false sense of understanding. Here, we present a new, non-self-reported metric of confidence and ask how knowledge and confidence play a role in determining public attitudes towards science. We propose a testable theoretical model and discuss how our results can inform science communication. The rationale is as follows: given any True/False/Don’t Know knowledge questionnaire for which there are consensual answers, an individual who knows the answer should answer it correctly and an individual who does not know the answer should answer Don’t Know. Therefore, we argue that every incorrect answer can correspond to a situation in which the individuals overestimate their own knowledge and this is a possible measure of overconfidence. We collected three large scale questionnaires (Science and Technology Eurobarometer, 1998 to 2005, in 34 European countries; American General Social Survey, 2006-2018; Pew Research Center, 2019) and focused on a set of variables gauging, knowledge and attitudes towards science. We found that our confidence metric varies non-linearly with knowledge, with the highest confidence appearing at the intermediate knowledge levels. This corresponds to the majority of the population, across countries, years, education levels, and age. We further report that overconfident individuals also tend to display the most negative attitudes towards science, with important implications for science communication and policy. This is contrary to current views, that either place knowledge as secondary (following the dismissal of the Deficit Model) and do not replicate the findings of the Dunning-Kruger effect, as both confidence and negative attitudes peak at intermediate levels of knowledge.

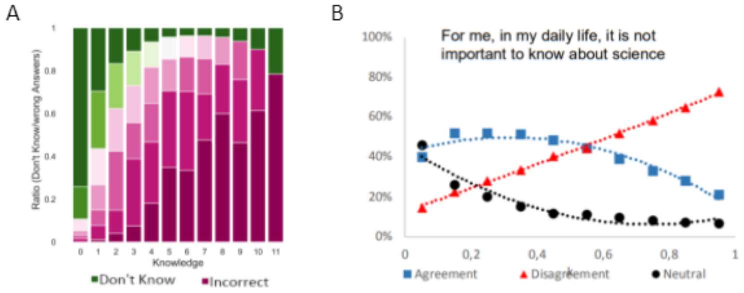


Figure 1: (A) Stacked bar chart shows fractions of respondents according to knowledge. Subgroups of each bar show fraction of respondents according to ratio of “Don’t know” to wrong answers. (B) Relative frequencies of agreement, disagreement and neutral stance for each knowledge category for one example of attitude items.

The Impact of Homophily on Digital Proximity Tracing

G. Burgio[1]*, B. Steinegger*[1]*, G. Rapisardi*[1,2]*, A. Arenas*[1]*. (1) Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, Tarragona, Spain, giulio.burgio@urv.cat. (2) Barcelona Supercomputing Center (BSC).*

Preventing disease outbreaks is one of the greatest challenges humanity faced in its history. The advent of digital proximity tracing (DPT) apps, and their large-scale implementation to prevent the spread of SARS-COV-2, sparked numerous studies that analyzed how this novel technology impacts disease propagation. A question which has not been addressed so far is how homophily affects the efficacy of DPT. Indeed, empirical studies suggests that the voluntary adoption of DPT apps strongly varies among the population, correlating with age, income and nationality [1,2]. Accordingly, adoption is much more probable among the contacts of an app adopter compared to a random individual, a phenomenon referred to as 'homophily' of social contacts. Through a simple model we reveal, analytically and numerically, a nonlinear phenomenology in terms of the reproduction number and the attack rate. The reproduction number has a minimum at a certain level of the mixing rate α between adopters and non adopters (see Figure 1A). Looking instead at the attack rate, we identify three different dynamical regimes (see Figure 1B): (i) critical, in which mixing is beneficial; (ii) saturated, in which it is detrimental; and (iii) intermediate, in which it is beneficial or detrimental, depending on the particular level of mixing. The key driver behind this rich phenomenology is the necessity that both infector and infectee are adopters in order for DPT to reduce onward transmission. We corroborate our findings with Monte Carlo simulations on a set of real contact networks. In summary, our results indicate that depending on infectivity and app coverage, homophilic adoption can be beneficial as well as detrimental for disease control.

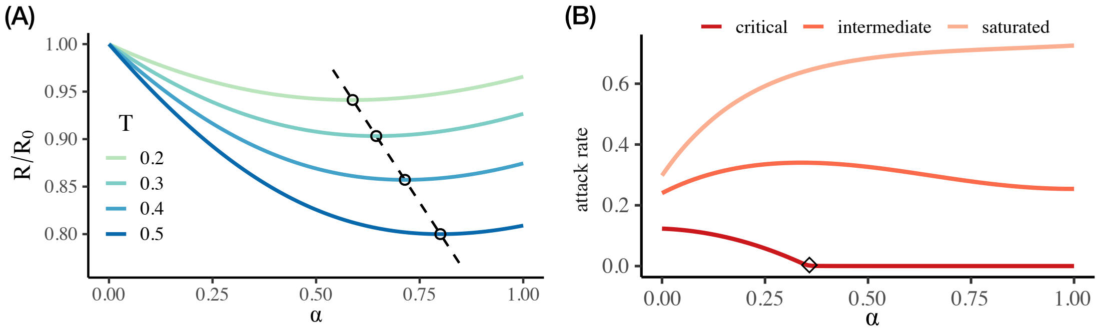


Figure 1: (**A**) Reproduction number R, normalized with respect to the basic reproduction number R_0, versus the mixing rate α, for different adoption coverages T. Dots indicate the minimum at α^*, while the dashed line shows its continuous variation with T. (**B**) The three dynamical regimes for the attack rate. The black diamond indicates the value of α for which $R = 1$.

References

[1] J. A. Moreno López, et al., *Science Advances 7(15), eabd8750* (2021).

[2] S. Munzert, et al., *Nature Human Behaviour 5(2), 247–255* (2021).

Analysing the validity and robustness of the iconic *World 3* global model: what can sensitivity and feedback loop analysis say?

M. Jochaud du Plessix[1], *S. Fenet*[2,1], *P.Y. Longaretti*[1]. *(1) Université Grenoble Alpes, CNRS, Inria, Grenoble INP, LJK, Grenoble, France, mathilde.jochaud-du-plessix@inria.fr. (2) Université Lyon 1, LIRIS, Lyon, France. (3)*

World3 is a flagship complex model. It was the first to widely explore potential future global socio-environmental trends, and to point out in a focused and sophisticated manner the possible dire consequences of forced growth in a finite and constrained environment. Since the inception of its first version in 1971 [1], it has attracted a lot of attention, comments, praises and criticisms. Many were however based on shallow arguments, and of the few in-depth analysis of the model that have been performed, most were carried out in an early period and lead to some incorrect conclusions due to their highly generic level of assessment of the model input parameters (see in particular [2]). The interest in the model has been revived in the last decade or so, in light of the growing concerns about planetary limits and the impact of human activity on natural systems, and following an *ex post* analysis of Turner [3]. Along with the availability of recent data, this revival of interest fueled the need to re-evaluate both the robustness of the model and the validity of its conclusions. This task presents however a delicate epistemological dimension: how can we evaluate the validity of a model with only one historical instantiation, and whose focal point of projection lies in the future? We will present the first two steps in the answer to this question.

First, we have performed an extensive analysis of the parameterization choices made by the authors of the model, based on a comprehensive sensitivity analysis of the whole model, made possible by modern computing power. This relies both a one-step-at-a-time screening method and a variance decomposition method. This first analysis allowed us to identify the most influential parameters in an objective way. Second, we performed an analysis of the inner dynamics of the model based on the relative importance of its feedback loops, based on a complete loop structural dominance analysis [4]. In order to deal with the large number of feedback loops of the model, we have used an enumeration of the smallest independent loop sets.

In conclusion, we will briefly outline the next step in this study, bearing on the role of modelling choices, and discuss in light of our results the relevance of the model to the real world and emerging Anthropocene era, some fifty years after its initial development.

References

[1] D.L. Meadows, W. W. Behrens, D. H. Meadows, R. F. Naill, J. Randers, E. Zahn. Dynamics of growth in a finite world. Cambridge, MA: Wright-Allen Press, 1974.

[2] P. Vermeulen, D. C. de Jongh. Parameter sensitivity of the ‘Limits to Growth’world model. *Applied Mathematical Modelling*, *1*(1), 29-32, 1976.

[3] G. M. Turner. A comparison of The Limits to Growth with 30 years of reality. *Global environmental change*, 18(3), 397-411, 2008.

[4] C. E. Kampmann, R. Oliva. Structural dominance analysis and theory building in system dynamics. *Systems Research and Behavioral Science: The Official Journal of the International Federation for Systems Research*, 25(4), 505-519, 2008.

Relationship between repost network evolution of trending hashtags and their ranking dynamics in microblogging complex system

***H. Cui*[1]**, *J.Kertész[1]. (1) Department of Network and Data Science, Central European University, Quellenstrasse 51, A-1100, Vienna, Austria, cui_hao@phd.ceu.edu.*

Online social networks are complex systems where people generate and propagate information. Some hashtags become popular by being reposted by a large number of users. Studies have been carried out on the repost network evolution characteristics of trending hashtags, as well as their ranking dynamics [1] in microblogging complex systems. However, few studies have explored the relationship between the repost network characteristics in the pre-trending period of a popular hashtag and the ranking patterns after it appears on the trending list. The largest Chinese microblogging website Sina Weibo is one such online complex system which is suitable for this study. Trending popular hashtags appear on Weibo's Hot Search List (HSL) and are ranked by their level of popularity indicated by search volume indices. As old popular hashtags vanish and new popular hashtags emerge, the Hot Search List updates in real-time, displaying only the most popular 50 hashtags. The popularity of the hashtags is associated with time-dependent ranks. For hashtags of different topical categories, the time series rank trajectories may have different patterns. We would first extract several rank trajectories categories which contain similar patterns, then explore the network characteristics of these trending hashtags during the pre-trending period, finally link the two aspects together to understand the relationship between the network characteristics in pre-trending period and the ranking dynamics after appearing on the Hot Search List. The Hot Search List in the microblogging system is a reflection of the collective attention in the whole system. Understanding the characteristics of the repost network evolution of trending hashtags during the pre-trending period is crucial to study the emergence and evolution of the collective attention in the microblogging complex system.

References

[1] Cui, H., Kertész, J. Attention dynamics on the Chinese social media Sina Weibo during the COVID-19 pandemic. *EPJ Data Sci.* 10, 8 (2021).

Tracking the political opinion landscape in Twitter during electoral periods

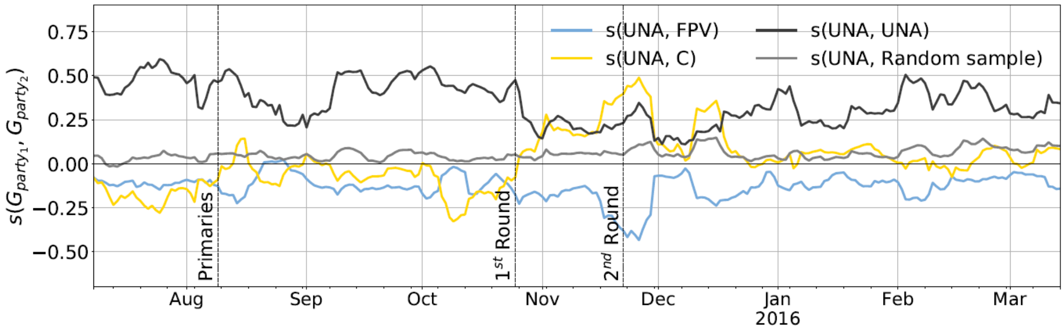
Tomás Mussi Reyero[1], Mariano G. Beiró[1,2], J. Ignacio Alvarez-Hamelin[1,2], Laura Hernández[3] and Dimitris Kotzinos[4]
1.Universidad de Buenos Aires. Facultad de Ingeniería, Paseo Colón 850, C1063ACV Buenos Aires, Argentina.
2. CONICET Universidad de Buenos Aires. INTECIN, Paseo Colón 850, C1063ACV Buenos Aires, Argentina.
3.Laboratoire de Physique Théorique et Modélisation, UMR-8089 CNRS, CY Cergy Paris Université, 2 Av. Adolphe Chauvin, 95302 Cedex, France.
4. ETIS UMR-8051 CY Cergy Paris Université, ENSEA, CNRS, 2 Av. Adolphe Chauvin, 95302 Cedex, France.

An interesting way to study the structure of opinions in Twitter exploits the hashtags chosen by the users, assuming that this choice reveals a concept that the user wishes to address. In a recent work [1], topics are defined by determining the community structure in a weighted network of hashtags, where two hashtags are connected if they appear together in the same tweet. Assuming that the coexistence of hashtags is semantically meaningful, the community structure of such network can reveal the general topics under discussion. In this way, the users may be characterized by a topic vector, with a dimension equal to the number of communities detected and where each component informs about whether the user is interested more or less than the average over the population, on the different topics.

In this work we extend these ideas to a dynamical study of the rapidly evolving opinion landscape that takes place in a society during an electoral campaign and with the case study of the two last presidential elections in Argentina (2015 and 2019). Specifically we understand by evolution of opinion, the evolution of the preferences of the social actors concerning the topics discussed in the platform, and we are particularly interested in detecting whether specific groups or users synchronize around some specific topics at some given times.

We compute similarity metrics among supporters of a given political party and across members of different parties as a function of time, and we are able to detect in a completely agnostic way (without choosing topics to follow) if the groups synchronize at a given time, meaning that they are all talking about the same topics, or on the contrary when they enter in complete opposition.

Our results show that the method is able to capture the reshaping of the political opinion landscape which has led to the inversion of the results between the two rounds of 2015 Argentinian election.



Cross-group similarities between one minority party (UNA) and the two major ones in 2015 : C (in yellow) and FPV (in light blue). The self-similarity of the minority party (black) along with the similarity of a randomized sample (grey) are plotted for comparison. The change of the sign of the crossed similarity with C (the winner party) is clearly observable between the two rounds.

References:

[1] Cardoso, F.M., Meloni, S., Santanche, A., Moreno, Y.: *Topical alignment in online social systems*. Frontiers in Physics 7, 58 (2019)

Dynamics of User Eccentricity on GAB Social Media

***Sriniwas Pandey*[1] *and H. Sayama*[1,2]** *(1) Center for Collective Dynamics of Complex Systems, Binghamton University, Binghamton, NY, USA; (2) Waseda Innovation Lab, Waseda University, Tokyo, Japan; spandey4@binghamton.edu.*

This study aims to investigate how ideas exchanged on social media drift away from the norm and how users become more *eccentric* (meaning "off-central") over time. Specifically, we collected posts and connection information for approximately three thousand users of GAB, a far-right oriented social media site, from August 2016 till January 2021, gathering a total of 147,000 gabs (posts). Each gab was converted into a numerical vector using Doc2Vec and principal component analysis. For each user, we reconstructed their "knowledge base" at every time point, which is a set of gab-vectors posted by the user or by their neighbors during the past five days. Using this knowledge base as a frame of reference, we computed the *eccentricity* of each new gab, which was defined as the distance of the new gab-vector from the "center of gravity" in the knowledge base of the user. This captures how off-central the new idea was relative to general discussion going on around the user. We then calculated the f-score (weighted average of absolute change in eccentricity over time) and the g-score (change in eccentricity with direction of change included) of each user's gabs' eccentricities. Based on these metrics, users were classified into four categories: (i) High f-score and g-score: 'Marginalized', (ii) High f-score and low g-score: 'Fickle minded', (iii) Low f-score and g-score: 'Stable', and (iv) High f-score and a very low g-score: 'Mainstreamed'. We also classified users based on the degree of deviation from their own previous posts (called *self-eccentricity*).

Results are summarized in Figure 1, showing the distributions of users in the f-score vs. g-score space for eccentricity (Fig. 1a) and self-eccentricity (Fig. 1b). Figure 1a shows that the users were nearly equally distributed in both positive and negative sides in their g-score, suggesting that it was equally likely for them to move towards or away from general discussion around them. However, Fig. 1b reveals that the users were generally drifting away from their older posts. A plausible interpretation of these two observations is that the GAB users tended to explore idea space outward (increasing both eccentricity and self-eccentricity) but they also assimilated to each other locally to form clusters of opinionated groups, which was captured as "Mainsteamed" behavior in Fig. 1a.

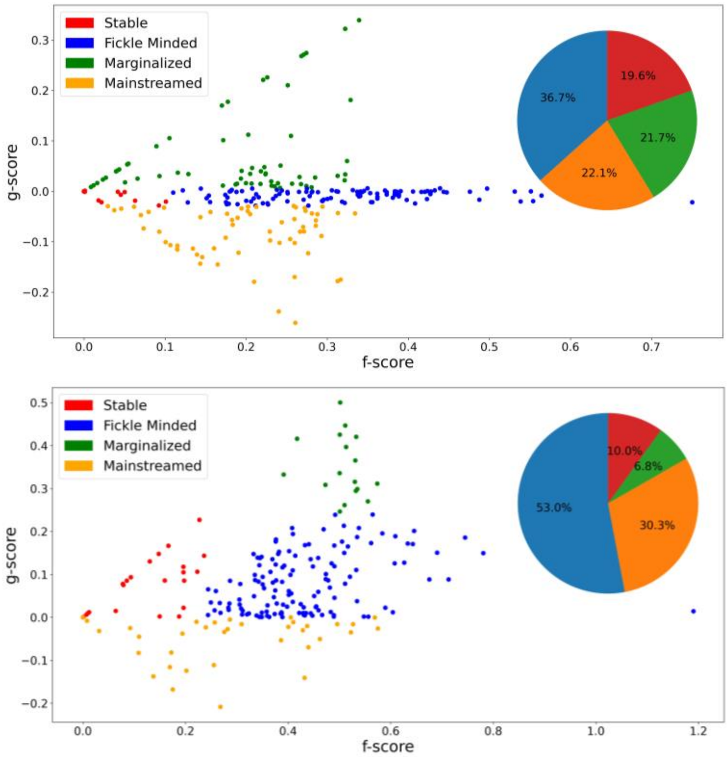


Figure 1: GAB users' behaviors plotted in a f-score (average absolute change in eccentricity) vs. g-score (average direction of eccentricity change) space. **Top:** Based on their gabs' eccentricity. **Bottom:** Based on their self-eccentricity (i.e., how different their new ideas were from their own older ones).

Catalyzing Collaborations: A Model for the Dynamics of Team Formation at Conferences

E. Zajdela[1]*, D. Abrams*[1]*, A. Feig*[2]*, R. Wiener*[2]*. (1) Northwestern University, 633 Clark St, Evanston, IL, USA, emmazajdela@u.northwestern.edu. (2) Research Corporation for Science Advancement, 4703 East Camp Lowell Drive, Suite 201, Tucson, AZ, USA*

The COVID-19 pandemic has brought to the fore the importance of collaboration among scientists to address global challenges. One of the main ways that collaborations are catalyzed is by gathering scientists together at conferences, but across the globe, conferences amount to billions of dollars per year in terms of travel expenses, organizing costs, and loss of research time. The goal of this research is to understand the mechanisms of team formation at conferences. In this work, we present a mathematical model for predicting the formation of scientific collaborations at conferences, inspired by the chemical process of catalysis. Specifically, the model uses differential equations to estimate the probability that pairs of participants will form a collaboration at any given time. It takes as inputs the participants' prior knowledge of each other (before the conference) and their level of interaction throughout the conference. Model predictions are tested using data from four multi-year series of conferences known as "Scialogs," organized by the Research Corporation for Science Advancement over the period 2015-2020. We analyze data from 12 Scialog conferences corresponding to over 13,000 pairs of participants and find that scientists who interact more intensely during a conference have a higher likelihood of forming a collaboration. Furthermore, we find that the likelihood of collaborating remains at a higher level even after the interaction between participants has ceased. These findings, along with the predictive model we develop, may have an impact on stakeholders from public, private, and nonprofit sectors who wish to optimize future conferences to promote new collaborations and may further generalize to help with promotion of team formation in a wide range of contexts.

Acknowledgements

Material based upon work supported by National Science Foundation Graduate Research Fellowship Grant No. DGE-1842165, the United States Department of Agriculture and the Northwestern Buffett Institute for Global Affairs.

7 Complexity in Physics and Chemistry

Contents

Inertia-driven Turing patterns on directed network topologies

R. Muolo[1], *M. Asllani*[2],*T. Carletti*[1] *(1) Department of Mathematics and naXys, Université de Namur, Rue Grafé 2, Namur, Belgium, riccardo.muolo@unamur.be (2) School of Mathematics and Statistics, University College Dublin, Belfield, Dublin 4, Ireland*

Nature offers a variety of systems where basic entities interact through nonlinear terms to eventually display order in space and/or time, i.e., patterns. Examples vary from biology (e.g., the patches observed in animals coats) to engineering (e.g., working regime of power grids). One of the most popular theories accounting for this self-organizing behavior was developed by Alan Turing, who proposed a mechanism where a diffusion-driven instability occurs after perturbing a homogeneous stable state [1]. Recently, such framework has been extended to complex networks [2], where the diffusion operator is modeled by the (symmetric) Laplacian matrix, and successively to directed networks [3], observing that the directed topology enhances the arising of Turing patterns. Despite the wide range of applications of the Turing ideas and the many obtained results, we would like to observe that they rely on a diffusion mechanism governed by the Fick law, hence they assume an infinite fast propagation of signals and matter. To overcome such unphysical assumption, the theory has recently been generalized on continuum [4] and on symmetric networks [5]. In the latter scenario, authors have proved the emergence of a new class of instabilities called *inertia-driven*, as they are not caused by diffusion, but by the sole inertia, i.e., a quantity related to the finite propagation velocity of the involved species [5]. In this work, we have made a step further by studying these newly found inertia-driven patterns on directed networks. Our numerical results show that an asymmetric topology enhances the possibilities of instability to occur (Figure 1), in analogy to what happens for diffusion-driven instability [3].

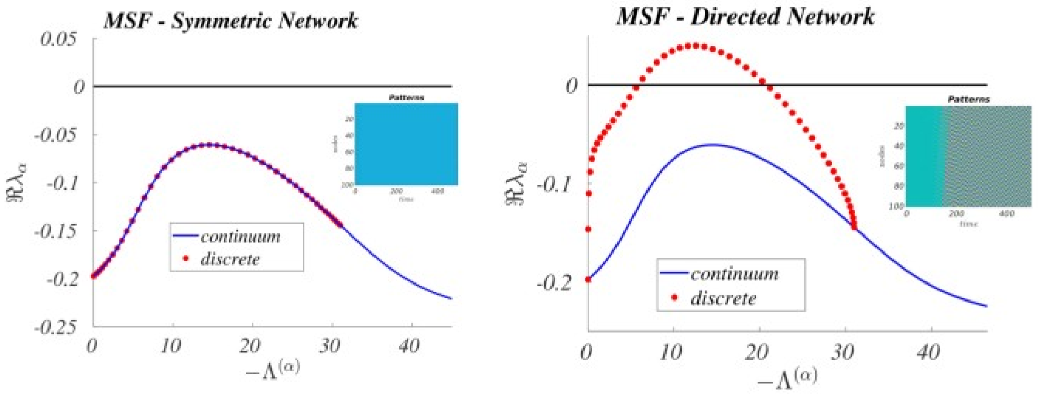


Figure 1: Master Stability Function (MSF) for the Brusselator model: the continuous MSF is always negative (blue line), but when the network is asymmetric (right panel) the discrete MSF (red dots) becomes positive and turns the system unstable.

References

[1] Turing AM, Phil. Trans. R. Soc. Lond. B 237 (1952) 37.

[2] Nakao H, Mikhailov AS, Nature Physics 6 (2010) 544.

[3] Asllani M, et al., Nature Communications 5 (2014) 4517.

[4] Zemskov EP, Horsthemke W., Physical Review E 93 (2016) 032211.

[5] Carletti T, Muolo R, preprint https://arxiv.org/abs/2104.04319 (2021).

Eigenvalue ratio spacing statistics of complex networks: Disorder vs Randomness

Ankit Mishra[1], *Sarika Jalan*[1,2]. *(1) Complex Systems Lab, Discipline of Physics, Indian Institute of Technology Indore, Khandwa Road, Simrol, Indore-453552, India Country, ankitphy0592@gmail.com (2) Department of Physics, Birla Institute of Technology and Science, Pilani 333031, India.*

The distribution of ratio of consecutive eigenvalue spacing of random matrices has emerged as important tool to study spectral properties of much body system as it doesn't require unfolding procedure as compared with the traditional level spacing distribution. We explore the ratio distribution of adjacency matrix of various model networks namely small-world network, scale-free network, dis(assortative) scale-free networks, ER random networks and dis(assortative) ER random networks. We also incorporate diagonal disorder from box distribution in the adjacency matrix in our analysis. We find that for all model networks the ratio distribution depicts GOE statistics. When diagonal disordered is added in the adjacency matrix, increasing the width of the diagonal distribution lead to the gradual transition from GOE to the poisson statistics. Moreover, the critical disorder (w_{c}) required to procure poisson statistics increases with the increase in the rewiring probability (randomness in connections) in the case of small-world networks. We further find that the assortative get to poisson statistics earlier compared to corresponding dissortative network. Moreover, we relate our results with the dynamic of maximal entropy random walker.

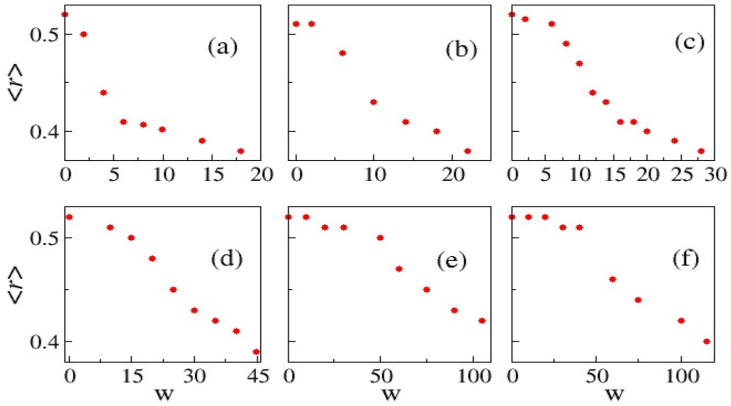


Figure 1: Plot of <r> as a function strength of diagonal disorder (w) for various rewiring probabilities of small-world newtworks. Here, $N = 2000$ and <k> = 20 with 20 network realizations and 40 disorder realizations. (a) p_{r} = 0.001 (b) p_{r} = 0.005 (c) p_{r} = 0.01 (d) p_{r} = 0.05 (e) 0.5 (f) 1

References

[1] Anderson, P. W., *Phys. Rev*, 109 (1958) 1492.

[2] Watts, Duncan J and Strogatz, Steven H., Nature , 393 (1998).

[3] Wegner, F., Z. Phys. B, 209 (1980)

Signatures of quantum synchronization in tomograms

Soumyabrata Paul *[1], S. Lakshmibala[1], V. Balakrishnan[1], S. Ramanan[1].*

(1) Department of Physics, Indian Institute of Technology Madras, Chennai-600 036, India, soumyabrata@physics.iitm.ac.in.

Quantum synchronization has been examined in recent years in diverse systems ranging from two coupled oscillators subject to an external driving field, to hybrid systems of atoms interacting with radiation fields. In contrast to classical synchronization, several measures of quantum synchronization have been proposed in the literature (see for instance, [1]). It has been suggested that the degree of quantum synchronicity (for example, [2]) is reflected in the manner in which entanglement between subsystems evolves temporally. Commonly used measures of entanglement in this context are concurrence and the subsystem linear entropy. Both are defined in terms of the reduced density matrices corresponding to the entangled subsystems. The density matrix however, is not directly obtained from experiments, but is reconstructed from tomograms, which are histograms of measured observables.

The reconstruction procedure involves maximum likelihood estimates, which are inherently error-prone, leading to non-optimal state reconstruction, particularly in the case of systems with large Hilbert spaces. Thus, in tracking entanglement dynamics, proposals to extract entanglement indicators without using density matrices, would be relevant, particularly because experiments on quantum synchronization are now being performed, and hence tomograms are readily available. We estimate entanglement directly from appropriate tomograms, thus providing a ready-reckoner of synchronization in the quantum context.

We have examined the dynamics of continuous variable synchronization in a hybrid tripartite system comprising a three-level Lambda atom interacting with two radiation fields, where transitions between the two lower energy levels is dipole forbidden and transitions between either of these levels to the excited state is mediated by a field. From optical tomograms, different entanglement indicators [3] based on tomographic entropy, mutual information, distances between probability distributions etc., are proposed. We examine the roles played by nonlinearities in the fields, the atom-field coupling strength, the initial state considered etc., in relating entanglement dynamics on the one hand to the manner in which the degree of synchronization changes in time, on the other hand.

Acknowledgements

The authors acknowledge partial support from Project no. SB20210807PHMHRD008128, of the Main Project No. SP20210777DRMHRDDIRIIT, sanctioned by the Ministry of Human Resource Development, under the IoE-CoE scheme, to IIT Madras.

References

[1] Tian-tian Huan, Ri-gui Zhou, Hou lan, *Scientific Reports*, 10, (2020), 12975.

[2] A. Mari, A. Farace, N. Didier, V. Giovannetti, R. Fazio, *Phys. Rev. Lett.*, 111, (2013) 103605.

[3] B. Sharmila, S. Lakhsmibala, V. Balakrishnan, *J. Phys. B: Atomic, Molecular and Optical Physics*, 53, (2020), 245502.

The Role of Particle Statistics in Quantum Reservoir Computing

***C. Charalambous**[1], G. L. Giorgii[1], R. Zambrini[1]. (1) IFISC, Instituto de Fisica interdisciplinar y Sistemas Complejos (CSIC-UIB), Campus Universitat de les Illes Balears, 07122, Palma de Mallorca, Spain, christos@ifisc.uib-csic.es.*

Reservoir computing (RC) is a neuro-inspired machine learning approach to time series processing. As such, it forms an example of a natural unconventional analog computer designed to perform a given computational task. Its power in solving nonlinear and temporal tasks depends on the reservoir possessing a high dimensional state space and the ability to retain memory of information for sufficiently long time. Quantum systems, with their large number of degrees of freedom and their complex real time dynamics satisfy both requirements, and for this reason are good candidates to serve as substrates for RC [1]. In addition, quantum effects such as superposition could lead to improvement in the performance of a RC. Recently the role of dynamical phases of closed quantum systems on the performance of quantum reservoir computers was studied [2]. In general, an open question in quantum reservoir computing (QRC) is the following: what properties a physical system must exhibit in order to be a good QRC? The aim of this particular work is to establish the relation between particle statistics of the particles comprising the complex computing system and its performance as a QRC.

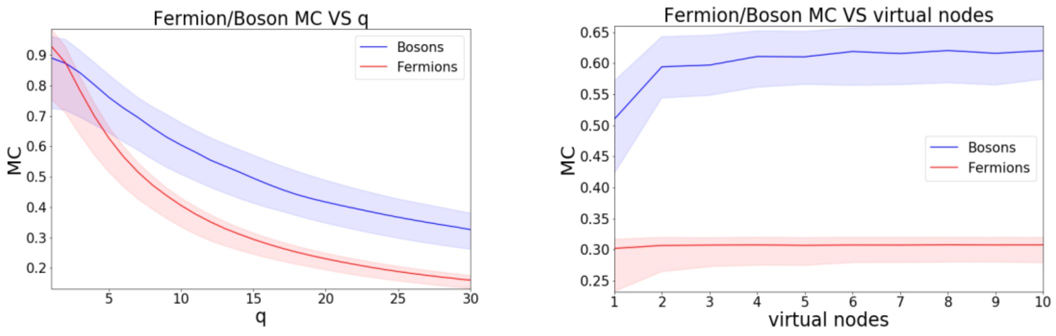


Figure 1: We use as a measure of the memory capacity of the reservoir the Pearson correlation. We find that for more nonlinear tasks, bosonic reservoirs perform better. Furthermore, time multiplexing enhances the performance of the bosonic reservoir, while it leaves the fermionic one unaffected.

Acknowledgements

This study was financially supported by the Spanish State Research Agency, through the Severo Ochoa and Marıa de Maeztu Program for Centers and Units of Excellence in R&D (MDM-2017-0711), CSIC extension to QuaResC (PID2019-109094GB), and CSIC Quantum Technologies PTI-001.

References

[1] R. Martinez-Peña, J. Nokkala, G. L. Giorgi, R. Zambrini, M. C. Soriano, *Cognitive Computation*, (2020) 1-12.

[2] R. Martinez-Peña, G. L. Giorgi, J. Nokkala, M. C. Soriano, R. Zambrini, *arxiv:2103:05348* (2021).

Rigidity percolation in disordered 3D rod systems

S. Heroy[1]*, D. Taylor*[2]*, F. Shi*[3]*, M.G. Forest*[3]*, and P.J. Mucha*[3]*. (1) University College London, 90 Tottenham Court Road, London W1T 4TJ, London, UK s.heroy@ucl.ac.uk (2) University at Buffalo, State University of New York, Buffalo, NY 14260, USA(3) University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA*

In composite materials composed of soft polymer matrix and rigid high aspect ratio particles, the host material undergoes a transition in mechanical strength when the incorporated particle phase surpasses a critical density. This phenomenon, termed rheological percolation, is well-known to occur in many polymer-rod and polymer-platelet composites at a critical density that exceeds the conductivity percolation threshold (which occurs when the conducting particles form a large connected component that spans the composite, see A-B below). Contact percolation in rod-like composites has been routinely exploited for conductive properties, and the characterization of contact percolation is well established. Mechanical or rheological percolation, however, has evaded a complete theoretical explanation and predictive description. A natural hypothesis is that mechanical percolation is due to a rigidity phenomenon, whereby a large *rigid* component of rods (or platelets) spans the composite (C-D of figure).

In brief, a rigid component is a necessarily connected component wherein inter-particle contacts eliminate all of the component’s nontrivial degrees of freedom. Rigidity percolation is frequently studied in biological/physical applications but we lack an understanding of rigidity percolation in systems of three-dimensional rods.

Here we build an algorithm to detect the rigidity percolation threshold in rod-polymer composites. We model the composites as systems of randomly distributed, soft-core (intersecting at contact) rods and study the emergence of a giant (spanning) rigid component. Building on our previous results for two-dimensional composites, we develop an approximate algorithm that identifies spanning rigid components through hierarchically identifying and compressing provably rigid motifs---equivalently, decomposing a giant rigid component into rigid interactions of smaller rigid components. We apply this algorithm to random monodisperse systems generated in Monte Carlo simulations to estimate a rigidity percolation transition point (critical density) and explore the dependence on rod aspect ratio. We show that this transition point scales inversely with the excluded volume of a rod. Interestingly, this scaling - unlike the contact percolation scaling derived by Balberg - is valid at relatively low aspect ratios without a deviation parameter. Moreover, this implies that the critical contact number (the number of contacts per rod at the rigidity percolation threshold) is independent of rod aspect ratio and lower than the prediction from Maxwell's isostatic condition.

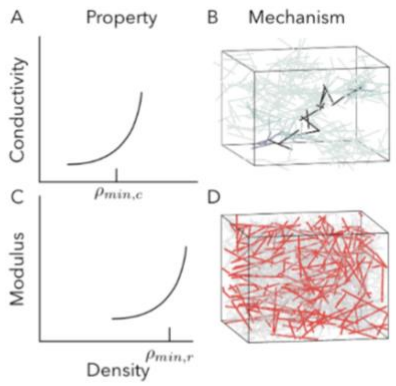


Figure 1: Conductivity percolation is driven by contact percolation, while rheological percolation is driven by rigidity percolation.

References:

[1] S. Heroy, D. Taylor, F. Shi, M.G. Forest and P.J. Mucha. Rigidity percolation in disordered 3d rod systems *Arxiv2103.12612 (2021).* https://arxiv.org/abs/2103.12612

Recurrence Network Analysis of the Dynamics of Quantum Observables: Comparison with Classical Dynamics

*Pradip Laha[1], **S. Lakshmibala[2]**, V. Balakrishnan[2].*

(1) Department of Optics, Palacky University,17. Listopadu 1192/12. 77146, Czech Republic

(2) Department of Physics, Indian Institute of Technology Madras, Chennai-600 036, India
slbala@physics.iitm.ac.in.

We present results obtained through time-series and recurrence network analysis (see for instance [1]) of the dynamics of quantum observables. This attempt is in the spirit of the Ehrenfest connection which suggests that quantum expectation values are the counterparts of classical dynamical variables. Novel features emerge, pertaining to the entanglement dynamics of relevant observables, and also in the identification of network quantifiers that mirror the results obtained through time-series analysis. For illustrative purposes, we examine two generic entangled tripartite systems with intrinsic nonlinearities and intensity-dependent couplings. One system comprises radiation fields interacting with an atom, and the other is an optomechanical set-up.

The mean photon number is identified as the appropriate observable that mimics the entanglement dynamics. We carry out a detailed time-series analysis of this observable and identify the manner in which the maximal Lyapunov exponent changes with variations in the intensity-dependent coupling strength. We analyze an interesting bifurcation cascade that appears over shorter time scales as the coupling strength varies, and examine the recurrence time distributions of the mean photon number. We then examine network quantifiers such as the clustering coefficient, assortativity etc., and identify promising quantifiers [2] which capture features obtained from the time-series analysis. Further, a similar investigation has been carried out on interacting classical coupled Duffing oscillator [3] which is an approximate counterpart of the atom-field interaction model mentioned above.

This program sheds light on the differences that arise in the dynamics of classical variables and quantum observables, as showcased through the time-series and network analysis, and also spotlights on the efficacy and limitations of network quantifiers in this context.

Acknowledgments

This work was supported in part by a seed grant from IIT Madras to the Centre for Quantum Information Theory.

References

[1] Y. Zou, R.V. Donner, N. Marwan, J. F. Donges, and J. Kurths, *Phys. Rep.*, 787, (2019) 1.

[2] Pradip Laha, S. Lakshmibala, and V. Balakrishnan, *Europhys. Letts.*, 125, (2019) 60005.

[3] Pradip Laha, S. Lakshmibala, and V. Balakrishnan, *Int. Jour. Theor. Phys.*, 59, (2020), 3476.

Structure of Wavefunction for Many-Boson Systems in Mean-Field with Random k-Body Interactions

*P. Rao**, *N. D. Chavda. Department of Applied Physics, Faculty of Technology and Engineering, The Maharaja Sayajirao University of Baroda, Vadodara-390001, India, *Email: pnrao-apphy@msubaroda.ac.in.*

With the recent experimental developments on ultra-cold quantum gases and production of Bose-Einstein (BE) condensates, there is a renewed interest in theoretical investigations of finite interacting many-particle complex quantum systems. These investigations in turn are useful in addressing various open problems of quantum statistical physics such as BE condensation, quantum many-body chaos and thermalization. In order to address these problems, it is extremely important to analyze wavefunction properties of these systems. The statistical properties related to the structure of the wavefunctions can reveal different layers of chaos and hence give profound understanding of various problems in the field of quantum many-body chaos and thermalization, in isolated finite interacting particle systems such as atomic nuclei, quantum dots, interacting spin systems modeling quantum computing core, ultracold atoms and quantum blackholes with SYK model and so on [1,2]. To analyze the wavefunction properties, it is very crucial to examine the so-called strength functions (also known as local density of states) in detail as they give information about how a particular basis state spreads onto the eigenstates. The chaos measures like number of principal components (NPC), information entropy, fidelity decay etc. can also be determined by examining the general features of the strength functions [3]. Very recently, q-Hermite polynomials have been employed to study spectral densities of SYK model [2] and it is shown that the Gaussian to semi-circle transition in the spectral density, can be well represented by the so-called q-normal distribution (f_{qN}) [4]. Recent analytical work on strength functions of fermion systems shows that the conditional q-normal density (f_{CqN}) can be used to represent strength functions in the thermalization region [5]. We analyze the wavefunction structure for interacting bosons in the presence of a mean-field using embedded random matrix ensembles of k-body interactions. We consider the Hamiltonian H, which is a sum of one-body $h(1)$ and an embedded Gaussian Orthogonal Ensemble (GOE) of k-body interaction $V(k)$ with interaction strength λ. It is denoted by BEGOE (1+k) (where 'B' represents bosons) [3]. A complete analytical description of the variance of the strength function as a function of λ and k is derived and the marker λ_t defining thermalization region is obtained. In the thermalization region ($\lambda > \lambda_t$), f_{CqN} describes Gaussian to semi-circle transition in strength functions as body rank k of the interaction increases. This interpolating form of the strength function is utilized to obtain the smooth form for the NPC, which is a measure of chaos in finite interacting many-particle systems [6].

References

[1] F. Borgonovi, F. M. Izrailev, L. F. Santos, V. G. Zelevinsky, *Phys. Rep.*, 626 (2016) 1.

[2] Y. Jia, J. J. M. Verbaarschot, *J. High Energy Phys.*, 07 (2020) 193.

[3] V. K. B. Kota, *Embedded Random Matrix Ensembles in Quantum Physics,* Springer-Verlag, Heidelberg, 2014.

[4] M. Vyas, V. K. B. Kota, *J. Stat. Mech. Theory Exp.*, 10 (2019) 103103.

[5] V. K. B. Kota, M. Vyas, arXiv:2011.05799v1.

[6] P. Rao, N. D. Chavda, *Phys. Lett. A*, 399 (2021) 127302.

Persistent individual bias in a voter model with quenched disorder

J. W. Baron *(1) Instituto de Fısica Interdisciplinar y Sistemas Complejos IFISC (CSIC-UIB), 07122 Palma de Mallorca, Spain*

Many theoretical studies of the voter model (or variations thereupon) involve order parameters that are population-averaged. While enlightening, such quantities may obscure important statistical features that are only apparent on the level of the individual. In this talk, I discuss which factors contribute to a single voter maintaining a long-term statistical bias for one opinion over the other in the face of social influence. To this end, I introduce a modified version of the network voter model that also incorporates quenched disorder in the interaction strengths between individuals and the possibility of antagonistic relationships. I demonstrate that a sparse interaction network and heterogeneity in interaction strengths give rise to the possibility of arbitrarily long-lived individual biases, even when there is no population-averaged bias for one opinion over the other. This is done by calculating the eigenvalue spectrum of the weighted network Laplacian using the theory of sparse random matrices.

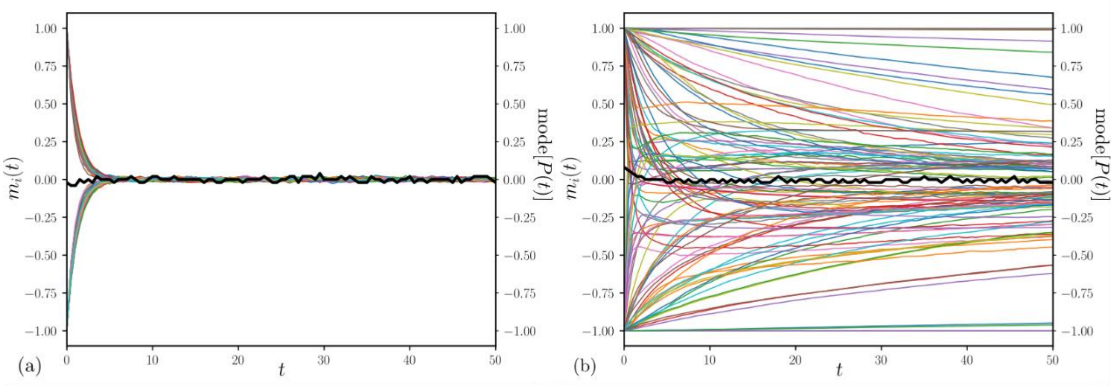


Figure 1: Panel (a): In the regular network voter model, individual biases quickly decay to zero. Panel (b): In the network voter model with quenched interactions, individual biases can persist. In both panels, the population averaged opinion (thick black line) fluctuates about zero.

Acknowledgements

Partial financial support has been received from the Agencia Estatal de Investigación (AEI, MCI, Spain) and Fondo Europeo de Desarrollo Regional (FEDER, UE), under Project PACSS (Grant No. RTI2018-093732-B-C21) and the Maria de Maeztu Program for units of Excellence in R&D (Grant No. MDM-2017-0711).

References

[1] J. W. Baron, *Phys Rev E,* 103 (2021) 052309.

8 COVID-19

Contents

Digital proximity tracing on empirical contact networks for pandemic control

G. Cencetti[1], *G. Santin*[1], *A. Longa*[1,2], *E. Pigani*[1,3], *A. Barrat*[4,5], *C. Cattuto*[6,7], *S. Lehmann*[8], *M. Salathé*[9], *B. Lepri*[1].

(1) Fondazione Bruno Kessler, Trento, Italy, gcencetti@fbk.eu. (2) University of Trento, Trento, Italy. (3) University of Padua, Padua, Italy. (4) Aix Marseille Univ, Université de Toulon, CNRS, CPT, Turing Center for Living Systems, Marseille, France. (5) Tokyo Tech World Research Hub Initiative (WRHI), Tokyo Institute of Technology, Tokyo, Japan. (6) University of Turin, Turin, Italy. (7) ISI Foundation, Turin, Italy. (8) Technical University of Denmark, Copenhagen, Denmark. (9) École Polytechnique Fédérale de Lausanne (EPFL), Lausanne, Switzerland.

Digital contact tracing is a relevant tool to control infectious disease outbreaks, including the COVID-19 epidemic. Early work evaluating digital contact tracing omitted important features and heterogeneities of real-world contact patterns influencing contagion dynamics. We fill this gap with a modeling framework informed by empirical high-resolution contact data to analyze the impact of digital contact tracing in the COVID-19 pandemic. We investigate how well contact tracing apps, coupled with the quarantine of identified contacts, can mitigate the spread in real environments. We find that restrictive policies are more effective in containing the epidemic but come at the cost of unnecessary large-scale quarantines. Policy evaluation through their efficiency and cost results in optimized solutions which only consider contacts longer than 15–20 minutes and closer than 2–3 meters to be at risk. Our results show that isolation and tracing can help control re-emerging outbreaks when some conditions are met: (i) a reduction of the reproductive number through masks and physical distance; (ii) a low- delay isolation of infected individuals; (iii) a high compliance. Finally, we observe the inefficacy of a less privacy-preserving tracing involving second order contacts. Our results may inform digital contact tracing efforts currently being implemented across several countries worldwide.

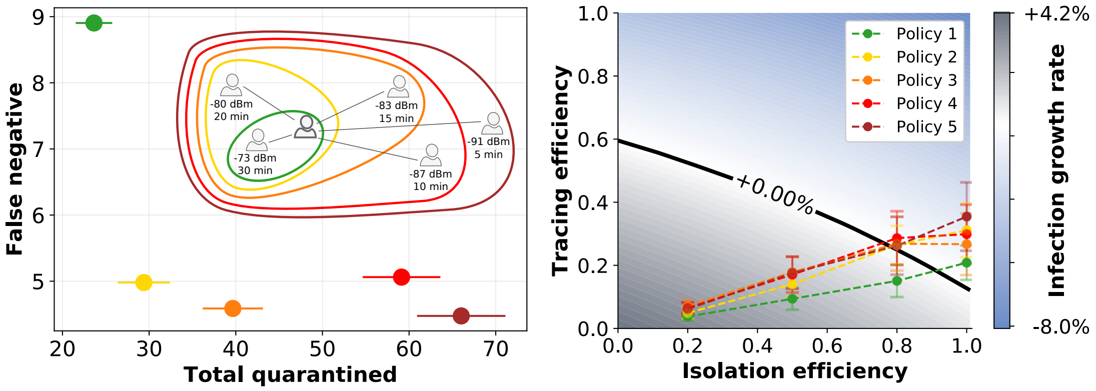


Figure 1: Tracing policies evaluation in terms of benefit and cost.

How to Mitigate the COVID-19 Outbreak While Maintaining Economic Activities?

T. Kano[1], *K. Yasui*[1,2], *T. Mikami*[1,3], *M. Asally*[4,5,6], *A. Ishiguro*[1]. *(1) Research Institute of Electrical Communication, Tohoku University, 2-1-1 Katahira, Aoba-ward, Sendai 980-8577, Japan, tkano@riec.tohoku.ac.jp. (2) Frontier Research Institute for Interdisciplinary Sciences, Tohoku University, Sendai 980-8578, Japan. (3) Graduate School of Engineering, Tohoku University, Sendai, 980-8579, Japan. (4) School of Life Sciences, University of Warwick, Coventry, CV4 7AL, UK. (5) Warwick Integrative Synthetic Biology Centre, University of Warwick, Coventry, CV4 7AL, UK. (6) Bio-Electrical Engineering Innovation Hub, University of Warwick, Coventry, CV4 7AL, UK*

While the minimization of human contact is effective in managing COVID-19 outbreaks caused by SARS-CoV2, it results in severe economic damages to several occupational categories such as restaurants and travel industries. Strategies are required to address this dilemma by considering the interrelation between the spread of the virus and economic activities. To address this problem, we had proposed an abstract agent-based model of the COVID-19 outbreak that accounts for economic activities [1].

In this paper, we discuss how to solve the aforementioned problem by extending our earlier study. Specifically, we modified the model such that wealthy people do not tend to undergo voluntary restraint but tend to use their money even during an outbreak. We compared the simulation results between the modified and the original model, and observed that the number of deaths caused by economic loss were smaller for the modified model than for the original model (Fig. 1). This result suggested that urging wealthy people to use their money might aid in mitigating outbreaks while minimizing economic losses.

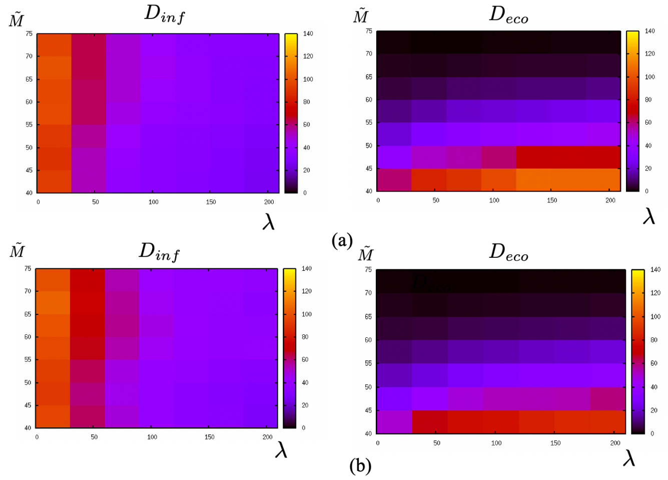


Figure 1: Simulation results. (a) Original model. (b) Modified model. The left and right color maps indicate the number of deaths caused by infection and economic loss, respectively. The horizontal and vertical axes in each color map represent the degree of voluntary restraint and the average amount of money in the community at the initial condition, respectively.

Acknowledgements

This work was supported by the JSPS KAKENHI Fund for the Promotion of Joint International Research (Fostering Joint International Research (B)) 19KK0103.

References

[1] T. Kano, K. Yasui, T. Mikami, M. Asally, and A. Ishiguro, *Proceedings of the Royal Society A*, 477 (2021) 20200604.

Epidemic Spreading and Digital Contact Tracing: Effects of Heterogeneous Mixing and Quarantine Failures

Abbas K. Rizi[1,*], Ali Faqeeh[1,2,3], Arash Badie-Modiri[1], and Mikko Kivelä[1]

[1]Department of Computer Science, School of Science, Aalto University, FI-0007, Finland
[2]Mathematics Applications Consortium for Science & Industry, University of Limerick, Ireland
[3]CNetS, School of Informatics, Computing, & Engineering, Indiana University, Bloomington, IN, USA
*abbas.karimirizi@aalto.fi

Contact tracing via digital tracking applications installed on mobile phones is an important tool for controlling epidemic spreading. Its effectivity can be quantified by modifying the standard methodology for analyzing percolation and connectivity of contact networks. We apply this framework for networks with varying degree distribution, the number of application users and the probability of quarantine failure. Further, we include structured populations with homophily and heterophily and the possibility of degree-targeted application distribution. Our results are based on a combination of explicit simulations and mean-field analysis. They indicate that there can be major differences in the epidemic size and epidemic probabilities which are equivalent in the normal SIR processes. Further, degree heterogeneity is seen to be especially important for the epidemic threshold but not as much for the epidemic size. The probability that tracing leads to quarantines is not as important as the application adaption rate. Finally, both strong homophily and especially heterophily with regards to application adoption can be detrimental. Overall, epidemics are very sensitive to all of the parameter values we tested out, which makes the problem of estimating the effect of digital contact tracing an inherently multidimensional problem [1].

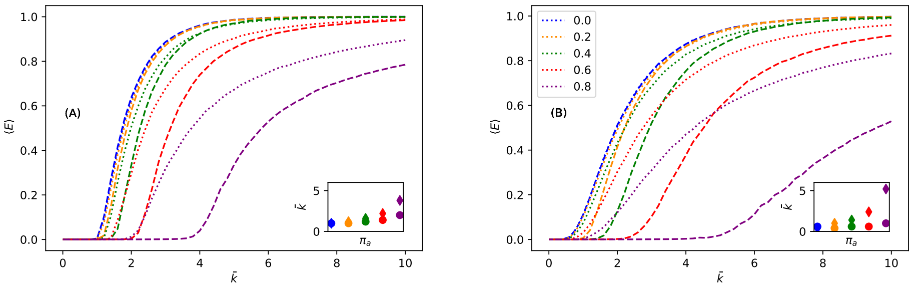


Figure 1: Effects of contact tracing on the expected epidemic size $\langle E \rangle$ and the epidemic threshold (insets) using two different strategies, random app adoption (small dots) and high-degree targeting (dotted lines) in (a) homogeneous networks with Poisson degree distribution and (b) heterogeneous networks with a power-law degree distribution $P(k) \propto k^{-3}$. Results are shown for $\pi_a \in [0, 0.2, 0.4, 0.6, 0.8]$ where π_a is the fraction of app users in the network.

References

[1] K. Rizi, Abbas, et al. ”Epidemic Spreading and Digital Contact Tracing: Effects of Heterogeneous Mixing and Quarantine Failures.” arXiv:2103.12634 (2021).

High-resolution stochastic dynamics underlying SARS-CoV-2 infection transmission and evolution

Bosiljka Tadic[1*], and Roderick Melnik[2]

[1] Department of Theoretical Physics, Jozef Stefan Institute, Jamova 39, Ljubljana, Slovenia

[2] Department of Mathematics, Wilfrid Laurier University, 75 University Avenue West, Waterloo, ON, Canada
bosiljka.tadic@ijs.si

Abstract

Several stochastic processes linked at different scales may contribute to the development of SARS-CoV-2 epidemics. Their essential features steam from the complex interplay of various biological and social factors, which emphasize the epidemics spreading as a collective phenomenon that remains elusive to strictly medical or purely social control measures. Therefore, a deeper understanding of active mechanisms can be sought within multi-scale and high-resolution dynamics modelling. In this context, the agent-based approaches that consider individual actors' properties offer valuable insights complementing traditional mean-field descriptions, augmented by compartmental models for interacting social groups.

Recently, we have developed an agent-based modelling framework [1] considering different biological and social factors entangled at the microscopic interaction scale between the viruses and hosts with high temporal resolution. The social dynamics that drives the epidemic spreading concerns the hourly resolution of day-to-day social activity under circulating active viruses. Meanwhile, every infection event depends on the agent's susceptibility to the virus in question, the virus survival time, traceability and mutations.

Here, we present the main features of the dynamical model and its capabilities to address several outstanding topics in the current epidemics. More precisely, we show how the model predicts the contribution of asymptomatic virus carriers to the overall infection growth and highlight the factors that critically affect it [2]. Furthermore, assuming a concrete level of social activity, we demonstrate the competing roles between massive immunization and the increased transmissibility due to virus mutations.

[1] B. Tadic, R. Melnik: *Modeling latent infection transmissions through biosocial stochastic dynamics*, PLoSone 15(10): e0241163 (2020)

[2] B. Tadic, R. Melnik: *Microscopic dynamics modeling unravels the role of asymptomatic virus carriers in SARS-CoV-2 epidemics at the interplay between biological and social factors*, Computers in Biology and Medicine, 133:104422 (2021)

Keywords:
epidemics, SARS-CoV-2, agent-based modeling, asymptomatic, immunization

Optimizing testing strategies for COVID-19 pandemic mitigation in school settings

E. Colosi[1], *G. Bassignana*[1], *C. Poirier*[1] *D. Contreras*[2], *A. Barrat*[2,3], *V. Colizza*[1,3]. *(1) INSERM, 27 rue de Chaligny, 75012 Paris, France, giulia.bassignana@iplesp.upmc.fr. (2) Aix Marseille Université, Université de Toulon, CNRS, CPT, Turing Center for Living Systems, 13009 Marseille, France. (3) Tokyo Tech World Research Hub Initiative (WRHI), Tokyo Institute of Technology, Tokyo, Japan*

Indoor environments such as educational settings are known to favor the transmission of infectious diseases such as COVID-19 because contacts between individuals are prolonged and in close proximity, while air circulation is limited. Therefore, different governments around the world proposed school closures as a first measure to stop viral spreading, especially during the first wave. This study evaluates screening strategies in educational settings to allow the quick detection and isolation of COVID-19 cases, with the final aim to control epidemic activity while avoiding unnecessary closures. Using temporal high-resolution data of face-to-face interactions between students and teachers in a French primary school [1], we developed an agent-based model to simulate the epidemic activity. Then we compared the outcomes of different testing protocols. In all protocols, symptomatic individuals are detected with a certain probability, tested, and isolated for 7 days if positive. We took this scenario as a baseline and compared the effect of symptomatic testing (ST) alone and coupled with (i) preventive isolation (PI) of classes where symptomatic cases are detected or (ii) regular screening (RT). We explored various testing frequencies and adherence levels (i.e., how many people accept to be tested). Preliminary results showed that preventive isolation presents similar final size and number of tests to symptomatic testing alone, but results in a larger number of person-day lost. Instead, regular screening was the most effective protocol, presenting the lowest values of both final size and person-day lost, when adherence is equal or bigger than 75%. However, regular screening has a high cost in terms of total number of tests required (100 times more than that of the symptomatic testing). These results can be used to inform public authorities providing a range of alternatives to closure. The option implemented can vary depending on both the epidemic activity and the final objectives (reduce the epidemic size, minimize the number of person-day lost, limit the number of tests).

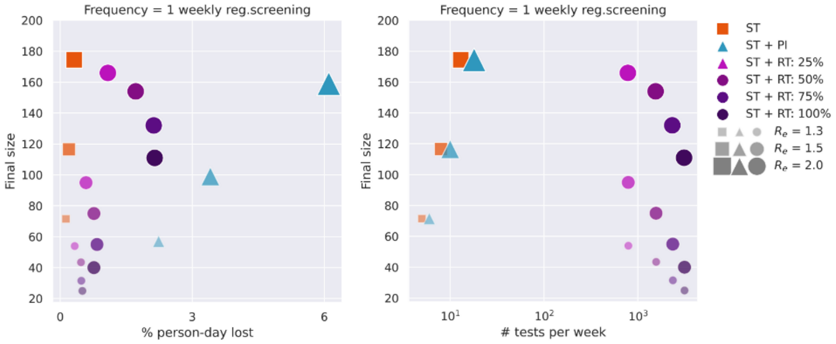


Figure 1: Median epidemic size for different values of effective reproductive number Re, for 500 simulations of 90 days of school. The initial seeding is set to three exposed individuals within the same class. Test sensitivity is θ = 80%. Colors and symbols stand for different testing protocols. Regular screening frequency is one test per week. Adherence ranges from 25% to 100%. Symbol size and color saturation increase with Re. Panel a) Epidemic size as a function of percentage of person-day lost. Panel b) Epidemic size as a function of number of tests.

References

[1] The SocioPatterns collaboration. http://www.sociopatterns.org (Accessed May. 13, 2021).

Impact of January 2021 curfew measures on SARSCoV-2 B.1.1.7 circulation in France

Laura Di Domenico[1,*], *Chiara E Sabbatini*[1], *Giulia Pullano*[1,2], *Daniel Lévy-Bruhl*[3], *Vittoria Colizza*[1,4]

(1) INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health, Paris, France
(2) Orange Labs, Sociology and Economics of Networks and Services (SENSE), Chatillon, France
(3) Santé publique France, Saint-Maurice, France
(4) Tokyo Tech World Research Hub Initiative, Institute of Innovative Research, Tokyo Institute of Technology, Tokyo, Japan

**laura.di-domenico@inserm.fr*

In response to the increasing spread of the SARS-CoV-2 B.1.1.7 variant, social distancing was strengthened in France in January 2021. We used a stochastic two-strain epidemic model, calibrated on genomic and hospitalization data, to assess the differential impact of implemented measures on the variant and the historical strain. The model accounts for four age classes; changes in contact patterns over time are informed from data on social contacts, mobility and school attendance, together with evidence from behavioral surveys quantifying loss of adhesion to social distancing in the general population and an increased risk aversion in the older age class. We showed that the progressive strengthening of social distancing measures implemented in January 2021, including an anticipation of curfew, allowed hospitalizations to plateau, bringing the effective reproductive number of the historical strain below 1, but were not sufficient to offset the spread of the B.1.1.7 variant. Winter school holidays further slowed down the epidemic in February. However, despite the foreseen increase in vaccination rhythm, a rapid resurgence in hospitalizations was predicted in March in absence of more stringent measures, as later confirmed by the observed epidemic progression.

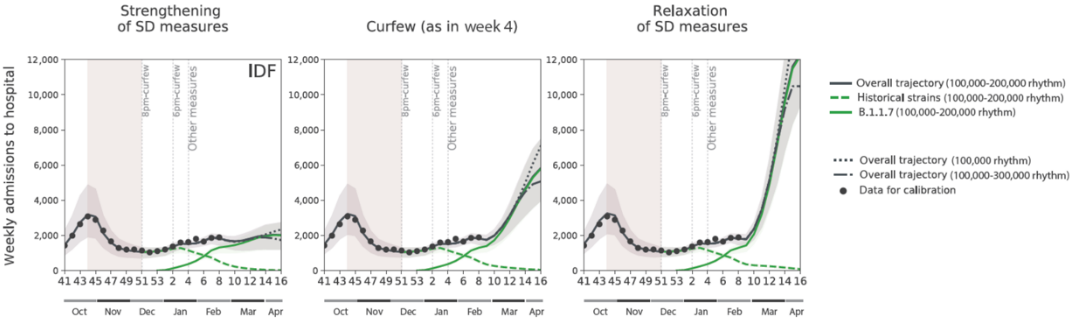


Figure 1: **Projected weekly hospital admissions due to SARS-CoV-2 historical strains and B.1.1.7 variant in Île-de-France.** From left to right, different scenarios after winter school holidays: strengthening of social distancing (SD) measures scenario; curfew scenario, estimated in week 4 and assuming no additional changes; relaxation of SD measures scenario. The solid grey curve refers to the median overall trajectory, due to the concurrent circulation of the historical strains (dashed green curve) and the B.1.1.7 variant (solid green curve). A slower (dotted curve) and an optimistic (dot-dashed curve) vaccination rhythm are also shown. The model is fitted to daily hospital admissions; the figure shows weekly data (black dots) to simplify the visualization.

References

[1] Di Domenico, L., Sabbatini, C. E., Pullano, G., Lévy-Bruhl, D. & Colizza, V. Impact of January 2021 curfew measures on SARS-CoV-2 B.1.1.7 circulation in France. Euro Surveill. 2021;26(15):pii=2100272. https://doi.org/10.2807/1560-7917.ES.2021.26.15.2100272

Adherence and sustainability of interventions informing optimal control against COVID-19 pandemic

Laura Di Domenico[1°], Chiara E. Sabbatini[1°], *Pierre-Yves Boëlle[1], Chiara Poletto[1], Pascal Crépey[2], Juliette Paireau[3,4], Simon Cauchemez[3], François Beck[4], Harold Noel[4], Daniel Lévy-Bruhl[4], Vittoria Colizza[1,5]*

(1) INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health, Paris, France, chiare.sabbatini@gmail.com (2) Univ Rennes, EHESP, REPERES « Recherche en Pharmaco-Epidémiologie et Recours aux Soins » – EA 7449, 35043, 6 Rennes, France. (3) Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, UMR2000, CNRS, Paris, France (4) Santé Publique France, French National Public Health Agency, Saint-Maurice, France (5) Tokyo Tech World Research Hub Initiative, Institute of Innovative Research, Tokyo Institute of Technology, Tokyo,Japan. · °co-first authors

Despite the maintenance of long lasting social distancing interventions to suppress the COVID-19 epidemic, some European countries still experience sustained viral circulation. This puts a significant burden on the health care system as well as on the population's quality of life. As the prospect of phasing out this phase approaches, it is all the more essential to balance the effectiveness and discomfort of the adopted measures, as the adherence of the population rapidly decreases with time. Using an age stratified transmission model and building on the analysis of the current situation in Île-de-France, we compared the efficacy of different interventions. The model integrates mobility data to parameterize social contacts over time and accounts for the co-circulation of the historical strains and B.1.1.7 variant. We fitted the model to daily hospital admissions, up to the strengthening of the measures in the region, and we compared predicted trends based on unchanged conditions with trajectories resulting from the application of stricter interventions. We show that high intensity lockdowns would be able to rapidly curb viral circulation, allowing a significant reduction of the burden of hospitalization despite loss of adhesion. Milder lockdowns require a much longer time to achieve the same result. These situations can be compared through a data-driven "distress" index summarizing the intensity and duration of social distancing and providing a measure of the constraints felt on average by an individual due to implemented policies (**Figure 1**). For intermediate values of the distress index, shorter strict lockdowns are largely more performant than longer moderate lockdowns. If moderate lockdown restrictions were less respected over time, epidemiological and healthcare indicators would considerably worsen, for relatively small gains in lowering the policy-induced distress.

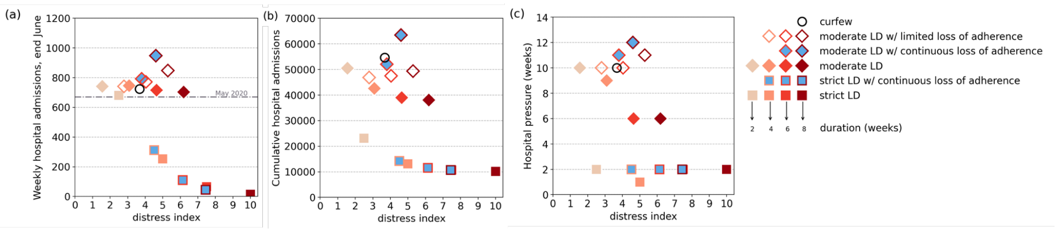


Figure 1: Weekly hospital admissions at the end of June (a), cumulative hospital admissions (b), hospital pressure, defined as the number of weeks in which hospital admissions remain above the peak level achieved during the second wave (c) as functions of the distress index. Symbol types refer to the stringency of intervention, color shades of the symbol contour refer to the duration of the lockdown intervention. Adherence to moderate and strict lockdowns is coded with the fill color.

References: Di Domenico L, Sabbatini CE, Boëlle P-Y, Poletto C, Crépey P, Paireau J, et al. *Adherence and sustainability of interventions informing optimal control against COVID-19 pandemic*. medRxiv https://www.medrxiv.org/content/10.1101/2021.05.13.21257088v1

COVID-19 Policy Analysis: Labor structure dictates lockdown mobility behavior

***S. Heroy*[1]*,I. Loaiza Saa*[2]*, A.S. Pentland*[2]*, and N. O'Clery*[3]*. (1) University College London, 90 Tottenham Court Road, London W1T 4TJ, London, UK* *s.heroy@ucl.ac.uk* *(2) MIT Media Lab,*

Countries and cities around the world have resorted to unprecedented mobility restrictions to combat Covid-19 transmission. Here we exploit a natural experiment whereby Colombian cities implemented varied lockdown policies based on ID number and gender to analyse the impact of these policies on urban mobility. Using mobile phone data, we find that the severity of local lock-down quotas, measured in the number of days citizens are allowed to go out, does not correlate with mobility reduction. Instead, wefind that larger, wealthier cities with a more formalized and com-plex industrial structure experienced greater reductions in mobility. Commuters are more likely to stay home when their work is located in wealthy or commercially/industrially formalized neighbourhoods. Hence, our results indicate that cities' employment characteristics and work-from-home capabilities are the primary determinants of mobility reduction. This finding underscores the need for mitigations aimed at lower income/informal workers, and sheds light on critical dependencies between socioeconomic classes in Latin America.

Experimental setup: We examine mobility reductions (extracted from call detail records) in 22 municipalities of interest that imposed various restrictions wherein residents were allowed out on days of the week corresponding to their gender/ID number. Do more severe quotas result in lower mobility, or are socioeconomic factors more important?

Larger cities had more dramatic responses, while policy had hardly any measurable effect: We measure change in trip frequency, relative change in daily distance traveled for a set of sufficiently tele-active users across the studied cities and find no statistically significant effect of policy severity, but strong evidence that residents of larger cities (which have higher GDP, labor formality, economic complexity) have more pronounced mobility disruptions. These results are additionally supported by GPS (Google Mobility) data at an aggregate level.

Trickle-down social distancing: Across the largest municipalities, we find that commuters who worked in wealthier and/or formalized commercial areas were much more likely to have their commute disrupted in lockdown. This finding is consistent with the above section, but the effect for low income commuters is especially interesting. Low income workers in higher income areas were apparently more likely to be shut off from labor, as their work caters to a wealthier clientele (e.g. low skill labourers in or near high income firms) than low income peers worker in lower income/informal areas. As a consequence, larger cities (with more income, formalization) have more commute disruptions.

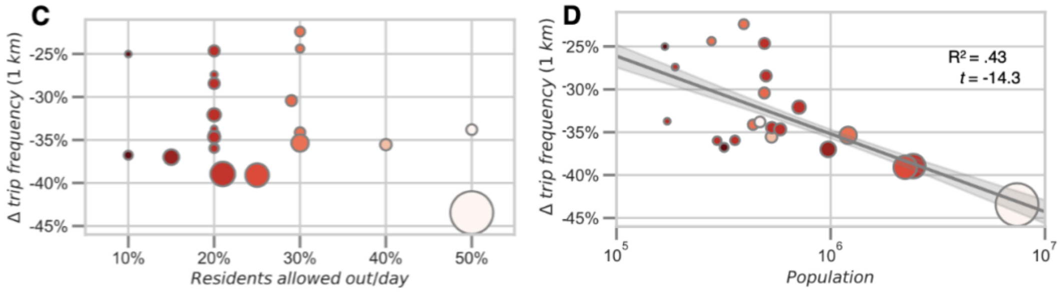


Figure 1: Increasing population (but not mobility quota) has a strong reducing effect on municipality-level median reduction in trip frequency.

References:

[1] Heroy, S., Loaiza, I., Pentland, A., & O'Clery, N. (2021). COVID-19 policy analysis: labour structure dictates lockdown mobility behaviour. *Journal of the Royal Society Interface*, *18*(176), 20201035.

Exploring the impact of mobility restrictions on the COVID-19 spreading through an agent-based approach

Martina Fazio[1]*,* ***Alessandro Pluchino***[1,2]*, Giuseppe Inturri*[3]*, Michela Le Pira*[4]*, Nadia Giuffrida*[5]*, Matteo Ignaccolo*[4]

(1) Department of Physics and Astronomy, University of Catania, Catania, Italy, alessandro.pluchino@ct.infn.it
(2) INFN Section of Catania, Catania, Italy
(3) Department of Electrical, Electronic and Computer Engineering, University of Catania, Catania, Italy
(4) Department of Civil Engineering and Architecture, University of Catania, Catania, Italy
(5) Spatial Dynamics Lab, University College Dublin, UCD Richview Campus, D04 V1W8, Belfield, Dublin, Ireland

Mobility restriction is considered one of the main policies to contain COVID-10 spreading. However, there are multiple ways to reduce mobility via differentiated restrictions, and it is not easy to predict the actual impact on virus spreading. This is a limitation for policy-makers who need to implement effective and timely measures. Notwithstanding the big role of data analysis to understand this phenomenon, it is also important to have more general models capable of predicting the impact of different scenarios. Besides, they should be able to simulate scenarios in a disaggregated way, in order to understand the possible impact of targeted strategies, e.g. on a geographical scale or in relation to other variables associated with the potential risk of infection. This paper presents an agent-based model (ABM) able to dynamically simulate the COVID-19 spreading under different mobility restriction scenarios. The model uses the Italian case study with its 20 administrative regions and considers parameters that can be attributed to the diffusion and lethality of the virus (based on an a-priori risk model [1]) and to population mobility patterns. The model is calibrated with real data and reproduces the impact that different mobility restrictions can have on the pandemic diffusion based on a combination of static and dynamic parameters. Results suggest that virus spreading would have been similar if differentiated mobility restriction strategies based on a-priori risk parameters instead of a national lockdown would have been put in place in Italy during the first wave of the pandemic. Further analysis on the second wave's data shows that the adopted strategy with differentiated restrictions did not change the a-priori risk ranking of the Italian regions. The proposed model could give useful suggestions for decision-makers to tackle pandemics and virus spreading at a strategic level.

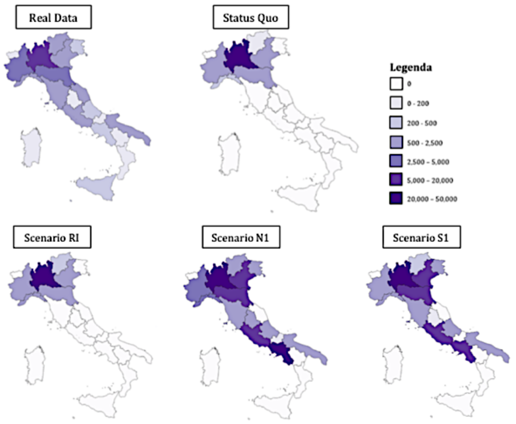


Figure 1: Distribution of the number of deaths in different scenarios

References

[1] Pluchino, A., Biondo, A. E., Giuffrida, N., Inturri, G., Latora, V., Le Moli, R., Rapisarda, A., Russo, G., Zappalà, C. *A Novel Methodology for Epidemic Risk Assessment: the case of COVID-19 outbreak in Italy*. Scientific Reports, 11 (2021) 5304. (https://www.nature.com/articles/s41598-021-82310-4)

Secondary Bacterial Infection Risk Index

P. Churchyard[1]*, R. Peruvemba*[1]*, A. K. Gupta*[1]*, C. Woloshyn*[1]*. (1) HSR.Health, 15200 Shady Grove Rd, Rockville, MD, United States, impact@healthsolutionsresearch.org.*

Bacterial co/secondary infection (BSI) can significantly increase the morbidity and mortality of viral infections. For example, up to 75% of those infected with influenza that go on to acquire pneumonia, are confirmed to have bacterial co-infection [1]. Further, the treatment for such infection is typically broad-spectrum antibiotics whose overuse contributes to long-term antimicrobial resistance [2]. Hence, the high prevalence of BSI warrants deeper investigation. In a pandemic or epidemic setting, BSI can have devastating consequences, particularly in high-risk groups such as the immunocompromised and the immunosuppressed. Immunosuppression is associated with more severe morbidity and a significantly higher risk of mortality from BSI [3]. During the 2009 Swine influenza pandemic, there was an increase in hospital pneumonia cases because of secondary bacterial pneumonia, which was identified in 29–55% of mortalities [4, 5]. Recently, the COVID-19 pandemic has highlighted the need for early recognition and appropriate treatment of those with BSI to minimize the associated costs both in lives lost and increased healthcare expenditures. We have developed a model for tracking and forecasting the risk of BSI befalling patients hospitalized with COVID-19. This Secondary Bacterial Infection (SBI) Index incorporates disparate data sources including geocoded health, social and environmental data, and historical clinical data. We believe our results substantiate this approach as a tool that enables early recognition of potential SBI, probable causative organism, and antibiotic susceptibility patterns. Future directions of this work include extrapolating this approach to the general hospitalized patient population. A critical component of any robust pandemic early warning system is an SBI Index; it represents a sophisticated solution for identifying the risk of BSI among hospitalized patients (e.g., COVID-19), and provides insights allowing for targeted treatment, hastened recovery, and reduced antibiotic over-utilization.

References

[1] M. C. Zambon, “The pathogenesis of influenza in humans,” *Reviews in medical virology*, vol. 11, no. 4, pp. 227–241, 2001.

[2] C. Llor and L. Bjerrum, “Antimicrobial resistance: risk associated with antibiotic overuse and initiatives to reduce the problem,” *Therapeutic advances in drug safety*, vol. 5, no. 6, pp. 229–241, 2014.

[3] T. W. Rice, L. Rubinson, T. M. Uyeki, F. L. Vaughn, B. B. John, R. R. Miller III, E. Higgs, A. G. Randolph, B. E. Smoot, B. T. Thompsonet al., “Critical illness from 2009 pandemic influenza a (h1n1) virus and bacterial co-infection in the united states,” *Critical care medicine*, vol. 40, no. 5, p. 1487, 2012.

[4] J. Louie, C. Jean, T. Chen, S. Park, R. Ueki, T. Harper, E. Chmara, J. Myers, R. Stoppacher, C. Cataneseet al., “Bacterial coinfections in lung tissue specimensfrom fatal cases of 2009 pandemic influenza a (h1n1)-united states, may-august 2009.” *Morbidity and Mortality Weekly Report*, vol. 58, no. 38, pp. 1071–1074, 2009.

[5] J. R. Gill, Z.-M. Sheng, S. F. Ely, D. G. Guinee Jr, M. B. Beasley, J. Suh, C. Deshpande, D. J. Mollura, D. M. Morens, M. Brayet al., “Pulmonary pathologic find-ings of fatal 2009 pandemic influenza a/h1n1 viral infections,” *Archives of pathology & laboratory medicine*, vol. 134, no. 2, pp. 235–243, 2010.

Developing a Business Risk Index for COVID-19 Transmission based on Mobility Networks

L. Meena[1]*, C. S. Kwan*[1]*, A. J. Gupta*[1]*, C. J. Woloshyn. (1) HSR.Health, 15200 Shady Grove Rd, Rockville, MD, United States, impact@healthsolutionsresearch.org.*

The COVID-19 pandemic has led to a significant number of local and state level business closures, affecting the worldwide economy on numerous fronts. In addition, state level blanket lockdowns have shed light on the vulnerability of local businesses [1], and the burden placed on employers' and employees' mental health [2]. Hence, these events demand a more strategic approach in order to mitigate the personal and professional disruptions caused by potential future pandemics. Early attempts to reopen the economy without effective measures in place could result in an uptick of COVID-19 cases, leading to as much as 18 percent more deaths than previously projected [3].

By analyzing big data from both HSR.*health*'s US state-level COVID-19 GeoHealth Platform, and SafeGraph's mobility and foot traffic data for over 3 million businesses, we capture disease dynamics by using a regression-based machine learning approach along with HSR.*health*'s proprietary Transmission Risk Index (TRI) to quantify the risks pertaining to COVID-19 transmission at places of interest (POI) such as restaurants, hotels, grocery stores, malls, etc. Moreover, we conduct a geospatial and time series analysis that forecasts COVID-19 cases at the aforementioned POIs. We present a network-based approach from this analysis to build a Business Risk Index (BRI), leveraging the mobility of the population, foot traffic, the physical dimensions of the building to determine occupant density within a business, and the TRI. The resulting BRI ranks POIs by their risk of disease transmission, enabling public health officials to target specific hotspots when developing response strategies and allowing for a controlled reopening of the economy.

Based on the reproduction number (R) value of potential future pandemics and our TRI, we can estimate a projected number of infections in a country, state, or community and, in conjunction with the BRI for that pandemic, can work with public health officials on targeted lockdowns, business capacity restrictions, and social distancing guidelines. This gives us the framework necessary to respond and deploy effective actions early.

References

[1] A. W. Bartik, M. Bertrand, Z. Cullen, E. L. Glaeser, M. Luca, and C. Stanton, "The impact of covid-19 on small business outcomes and expectations," *Proceedings of the National Academy of Sciences*, vol. 117, no. 30, pp. 17 656–17 666, 2020.

[2] G. Giorgi, L. I. Lecca, F. Alessio, G. L. Finstad, G. Bondanini, L. G. Lulli, G. Ar-cangeli, and N. Mucci, "Covid-19-related mental health effects in the workplace: a narrative review," *International journal of environmental research and public health*, vol. 17, no. 21, p. 7857, 2020.

[3] M. Berger and J. Chesak, "The consequences of reopening are here: Covid-19 cases on the rise." 2020. [Online]. Available: https://www.healthline.com/health-news/covid-19-reopening-consequences-cases-rising

Anatomy of digital contact tracing: Role of age, transmission setting, adoption and case detection

***B. Arregui García**[1,2] beatriz@ifisc.uib-csic.es, J. A Moreno López[1,2], P. Bentkowski[1], L. Bioglio[3], F. Pinotti[1], P-Y Boëlle[1], A. Barrat[4,5], V. Colizza[1] and C. Poletto[1]. (1) INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health, Paris, France (2) Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), CSIC-UIB, Palma de Mallorca, Spain, (3) Department of Computer Science, University of Turin, Turin, Italy, (4) Aix-Marseille Univ, Université de Toulon, CNRS, CPT, Turing Center for Living Systems, Marseille, France, (5)Tokyo Tech World Research Hub Initiative (WRHI), Tokyo Institute of Technology, Tokyo, Japan*

The efficacy of digital contact tracing against coronavirus disease 2019 (COVID-19) epidemic is debated: Smartphone penetration is limited in many countries, with low coverage among the elderly, the most vulnerable to COVID-19. We developed an agent-based model to precise the impact of digital contact tracing and household isolation on COVID-19 transmission [1]. The model, calibrated on French population, integrates demographic, contact and epidemiological information to describe exposure and transmission of COVID-19. We explored realistic levels of case detection, app adoption, population immunity, and transmissibility. Assuming a reproductive ratio R = 2.6 and 50% detection of clinical cases, a ~20% app adoption reduces peak incidence by ~35%. With R = 1.7, >30% app adoption lowers the epidemic to manageable levels. Higher coverage among adults, playing a central role in COVID-19 transmission, yields an indirect benefit for the elderly. These results may inform the inclusion of digital contact tracing within a response plan against COVID-19 and future emerging disease epidemics.

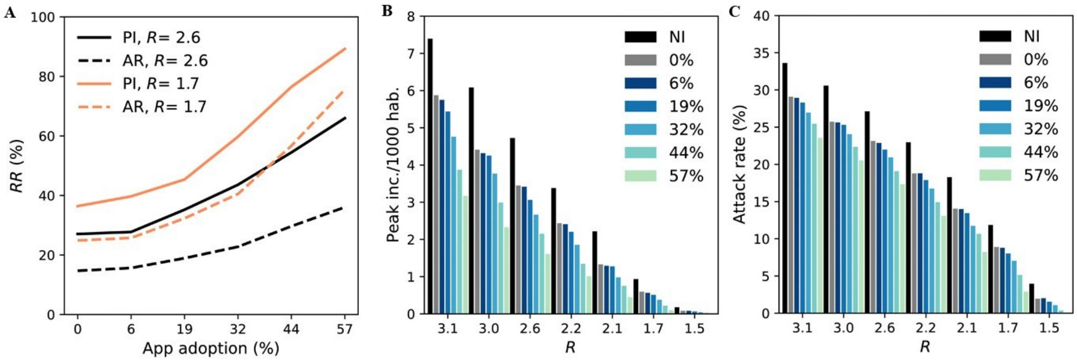


Figure 1: **Impact of digital contact tracing and household isolation on the epidemic: A** - Relative reduction (*RR*) versus no intervention (*NI*) scenario in attack rate (*AR*) and peak incidence (*PI*) as a function of the app adoption. The attack rate is computed as cumulative incidence discounting initial immunity (10%). Two reproductive ratio (R=1.7 and 2.6 are considered). **B**,**C** – Peak incidence and attack rate according to reproduction ratio R and app adoption. Parameter values were as follows: initial immunity, 10%; clinical case detection, 50%; subclinical case detection, 5%; compliance to isolation of contacts notified by the app, 90%; and compliance to isolation of household contacts, 90%.

References

[1] J. A. Moreno López, B. Arregui García, P. Bentkowski et al. Anatomy of digital contact tracing: role of age, transmission setting, adoption and case detection. *Sciences Advances* | 09 Apr 2021 Vol. 7, no. 15, DOI: 10.1126/sciadv.abd8750

Dark Web Marketplaces and COVID-19

***M. Nadini**[1,2], A. Bracci[1], M. Aliapoulios[3], D. McCoy[3], I. Gray[4], A. Teytelboym[5,6], A. Gallo[7], A. Baronchelli[1,2,8]. (1) City, University of London, Department of Mathematics, London, UK, Matthieu.Nadini@city.ac.uk (2) The Alan Turing Institute, British Library, London, UK (3) Center for Cybersecurity, New York Univ. Tandon School of Engineering, Brooklyn, NY, USA (4) Global Intelligence Team, Flashpoint. New York, NY, USA (5) Institute for New Economic Thinking, Oxford Martin School, University of Oxford, Oxford, UK (6) Department of Economics, University of Oxford, Oxford, UK (7) Department of Finance, The Bayes Finance Faculty, London, UK (8) UCL Centre for Blockchain Technologies, University College London, UK*

The COVID-19 pandemic has brought a worldwide public health emergency, economic distress, and disinformation-driven panic, whose combination has reshaped our online socio-economic behaviour. In particular, we analysed 23 millions of listings on 164 dark web marketplaces (DWMs) appearing between January 2020 and April 2021 and detected that shortages of COVID-19 goods (e.g., masks, COVID-19 tests and vaccines) pushed customers, and vendors, towards illicit online trade [1,2]. Figure 1 shows that the complex behaviour of DWMs is correlated with major COVID-19 events and the activity of social networks, such as, Twitter and Wikipedia [1]. Figure 2 illustrates that COVID-19 vaccines were offered on DWMs since March 2020, while officially approved vaccines appeared much later, close to their rollout date. In late February 2021, the first fabricated proofs of vaccination (like COVID-19 passports) started to were offered on DWMs. Illicit trade of COVID-19 products poses a concrete threat to public health and risks to undermine the efficacy of the ongoing vaccination campaign.

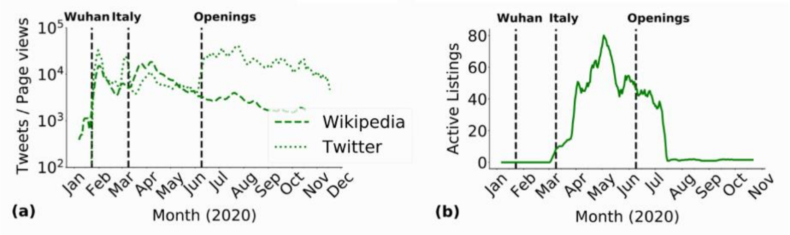


Figure 1: (a) Evolution of Twitter posts and Wikipedia page visits between Jan 2020-Nov 2020. (b) Evolution of Personal Protective Equipment on DWMs on the same period [1].

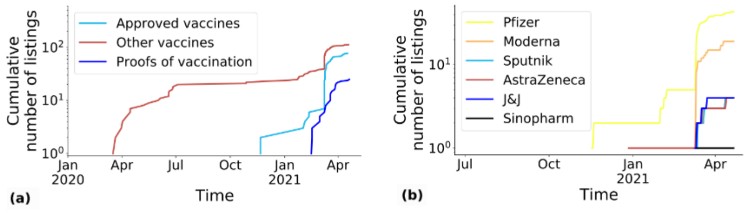


Figure 2: Evolution of COVID-19 vaccines on DWMs between Jan 2020-Apr 2021 [2].

References

[1] A. Bracci, M. Nadini, M. Aliapoulios, D. McCoy, I. Gray, A. Teytelboym, A. Gallo, A. Baronchelli. EPJ Data Science 10.1 (2021): 1-26

[2] A. Bracci(+), M. Nadini, M(+). Aliapoulios, D. McCoy, I. Gray, A. Teytelboym, A. Gallo, A. Baronchelli. *Preprint arXiv:2102.05470 (2021)* (+) Contributed equally

Critical Interventions: Mobility-Based Thresholds for Epidemic Containment.

Jesús A. Moreno López[1], *Alberto Hernando*[2], *Sandro Meloni*[1], *José J. Ramasco*[1]. *(1) Institute for Cross-Disciplinary Physics and Complex Systems (IFISC), CSIC-UIB, Palma de Mallorca, Spain. jeslop@ifisc.uib-csic.es (2) Kido Dynamics SA, Avenue de Sevelin 46, 1004 Lausanne, Switzerland*

There is an expected mutual influence between mobility, contacts and the evolution of an epidemic. As observed in many countries, mobility has been a crucial target for containment measures against SARS-CoV-2 spreading such as confinements and other mobility restrictions. In this work we introduce a simple agent based model with a direct relationship between mobility and contacts. We explore how threshold interventions based on a 14 day cumulative incidence (CI), known as "COVID-19 traffic light system" amongst many countries, lead to a sharp oscillatory behaviour in incidence curves (Fig. 1A). Recent works have suggested the existence of a critical mobility threshold on the radius of gyration *Rg*, which would dictate the growth/decline of an epidemic depending on the mobility's level with respect to the threshold. This can inspire a new kind of preemptive intervention rather than the observed reactive interventions, controlling the epidemic before epidemic spreading accumulates an increasing inertia (Fig. 1C). Estimating and tracking the evolution of the critical mobility threshold *Rg,c* as time and immunity advance allows us to define an intervention based on remaining under the threshold. In our model, this method allows easier control over the epidemic, improving greatly upon mobility restrictions (Fig. 1B and 1D) and preventing very high incidence rates associated to sharp epidemic waves, which could possibly saturate health care resources. Our results will be included in an upcoming paper [1].

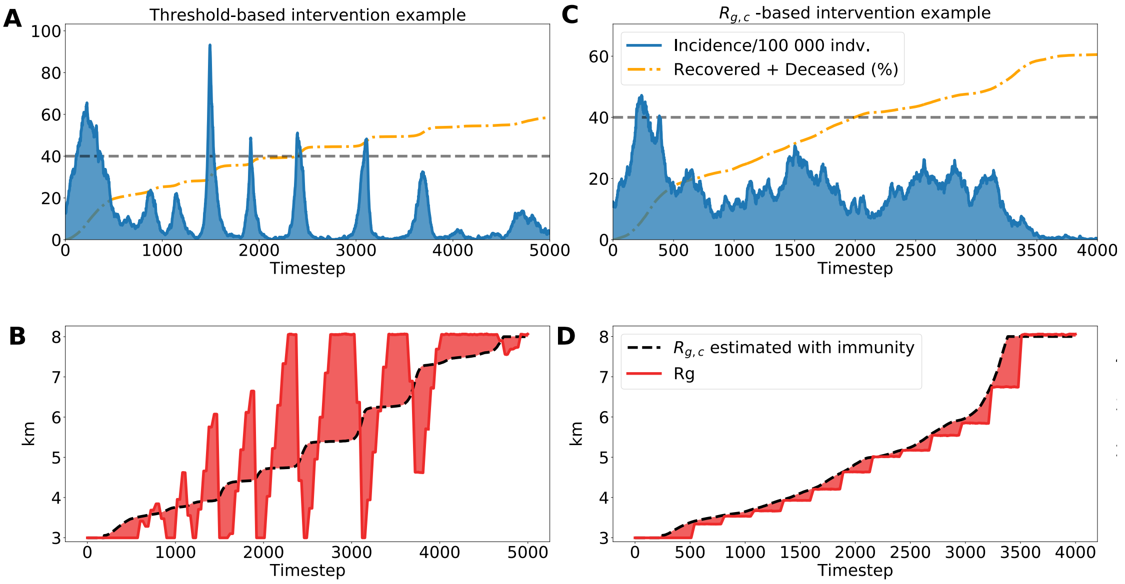


Figure 1: (**A**, **B**) Examples of incidence, recovered+deceased, *Rg* and our estimation of *Rg,c* in a given realisation of threshold interventions where the 14 day CI threshold = 150/100 000 and measures are revised every 30 timesteps. (**C**, **D**) Examples of incidence, recovered+deceased, *Rg* and our estimation of *Rg,c* in a given realisation of *Rg,c*-based interventions where mobility is re-adjusted to the estimation of Rg,c every 270 timesteps.

References:

[1] *J.A.M. López, A. Hernando, S. Meloni, J.J. Ramasco*. Critical interventions: Mobility-based thresholds for epidemic containment. *In progress.*

Infection rate analysis at each daily activity under COVID-19 pandemic using human mobility data

Jun'ichi Ozaki[1], *Yohei Shida*[2], *Hideki Takayasu*[1,3], *Misako Takayasu*[1,2]. *(1) Institute of Innovative Research, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan,* *ozaki.j.ac@m.titech.ac.jp* *(2) Department of Mathematical and Computing Science, School of Computing, Tokyo Institute of Technology, 4259 Nagatsuta-cho, Midori-ku, Yokohama 226-8503, Japan (3) Sony Computer Science Laboratories, 3-14-13 Higashi-Gotanda, Shinagawa-ku, Tokyo, Japan*

The COVID-19 pandemic has been one of the new challenges imposed on human beings. The pandemic has forced us to face the simultaneous problem of epidemic prevention and economic activity. Governments have struggled to solve that problem by restricting specific activities considered to be high-risk for infection while permitting other activities (e.g. [1]). However, their policy has not always been determined by fully scientific processes because of the limitation of time, data, or human resources. The back actions of the governments' policy, such as increasing stay-at-home rates [2], should be examined to check its effectiveness, to make better choices. Also, the trade-off between the infection and the economy has much to be investigated in detail.

In this paper, we analyze human mobility derived from mobile phone GPS tracking data in 2020, during the COVID-19 pandemic. The GPS dataset provided by Agoop, a Japanese private company, consists of individual locations for about 1,000,000 mobile phone users in Japan with mostly less 10 minutes intervals. The user trajectory can be tracked in a day with anonymized user IDs. Then we categorize daily activity by the GPS trajectory of each person into the following four types: *home*, *work*, *move*, *stay*. We calculate the population density in each spatial mesh at each activity, which is renormalized to the total population in each region for the infection modeling.

We propose an infection model based on the SIR model, and we discover that the infection intensity is strongly dependent on daily activity type. We discretize the time series and discard the changes in the Susceptible and Recovered variables because in the Japanese case, the infection is very low compared to the total population. We model that the infection rate, or the effective reproduction rate, is a function of the population density at each mesh and activity: sum of the contact infection intensity times infection coefficients of each activity as follows.

$$\alpha(t) = \sum_{a,m,\tau} C_a \rho_{am\tau}^{\gamma_a}(t),$$

where $\alpha(t)$ is the effective reproduction number in week t, and a, m, τ denotes activity type, spatial mesh, and date-time in each week, and C_a is the coefficient, $\rho_{am\tau}$ is the population density, γ_a is the contact exponential factor of each activity. The parameters are estimated to fit the real effective reproduction rate in Japan. Finally, the seasonality and countermeasure to COVID-19 such as mask effects are also considered, which are included through the time-dependent factor $C_a(t)$. Thus the different infection behaviors at the similar human mobility situation are explained.

References

[1] Japanese Ministry of Health, Labour and Welfare, https://www.mhlw.go.jp/stf/seisakunitsuite/bunya/0000164708_00079.html. (on 05/19/2021)

[2] Yabe, T., et al., *Scientific Reports* 10(1), 1–9 (2020).

Mobility induces a network interpretation of the Reproductive Number that explains COVID-19 spatial spread and informs surveillance

M. Mazzoli[1], *E. Valdano*[1], *Z. Susswein*[2], *S. Bansal*[2], *V. Colizza*[1]

(1) INSERM, Sorbonne Université, Institut Pierre Louis d'Epidémiologie et de Santé Publique, IPLESP, Paris, France

mattia.mazzoli@inserm.fr *(2) Department of Biology*, Georgetown University, Washington, DC, USA

The COVID-19 pandemic is still active, with worldwide circulation of viral variants of concern. Vaccines are largely effective, but their rollout is still insufficient to stop the epidemic, and surveillance and non-pharmaceutical interventions remain necessary in many contexts. This is the case of France, currently phasing out the third lockdown, while accelerating vaccination (30% of the population with at least 1 dose by mid-May) and strengthening surveillance of the variants of concern on the territory. Surveillance systems, however, mainly rely on local indicators, like the effective reproductive number R (the average number of secondary cases generated by each case), neglecting the impact of the spatial dynamics in sustaining the epidemic spread. To account for the spatial coupling of local epidemics due to mobility, we introduce here a theoretical framework based on a metapopulation scheme that defines the reproductive number as a spatial network among the 94 departments of mainland France. By integrating co-location records from Facebook at high spatial resolution together with official epidemic data and estimates, we define metrics for the risk of case importation (R_{imp}) and exportation (R_{exp}) at each location. Focusing on the summer-fall 2020 period, we show that spatial importation of cases at specific locations at the start of the summer (e.g. Loire in Figure 1, with a peak in R_{imp}) triggered a local epidemic (estimated number of local cases I_{loc} in the plot) opening the path to the second wave in the fall, with a large risk of exportation to other locations (peak in R_{exp}). Our indicator is able to identify early signals highlighting the potential for seeding events that should be the focus of interventions (at the source, to prevent exportations, and at destination, to prevent importation), and that are not identifiable through local surveillance. This becomes particularly important for the surveillance of B.1.351 and P1 variants currently expanding in the region of Paris.

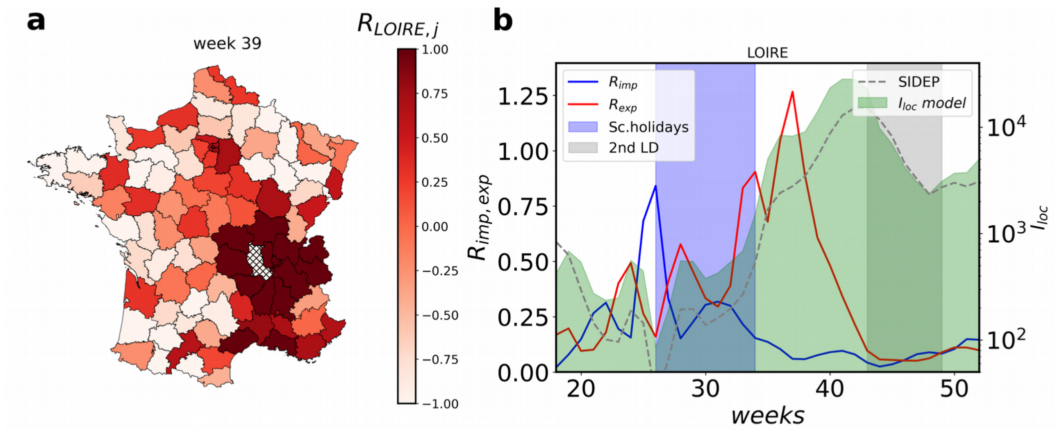


Figure 1: **(a)** Epidemic risk R_{ij} of spreading the historic variant from Loire to the rest of Metropolitan France in week 39 of 2020, colors in log scale. **(b)** The evolution in time of the risk of importation R_{imp}, in blue, exportation R_{exp}, in red, the number of locally developed cases I_{loc} by the model in Loire and the officially reported cases SIDEP in dashed grey. Risk is assessed considering weekly mobility, vaccination rates and efficacy, strains presence and transmissibility, incidence and seroprevalence in each unit.

Epidemic Spreading on Vaccinated Networks

V. Blavatska[1,2], *Yu. Holovatch*[1,2,3]. *(1) Institute for Condensed Matter Physics of the National Academy of Sciences of Ukraine, 1 Svientistski Str., 79011 Lviv, Ukraine, blavatskav@gmail.com. (2) L4 Collaboration & Doctoral College for the Statistical Physics of Complex Systems, Leipzig-Lorraine-Lviv-Coventry, Europe. (3) Centre for Fluid and Complex Systems, Coventry University, Coventry, CV1 5FB, United Kingdom*

We analyze infection spreading processes on so-called scale-free networks, where the distribution of node connectivity k obeys a power-law $p(k) \sim k^{-a}$ [1]. We asume, that only a fraction p of individuals can be affected by disease, while remaining $1-p$ individuals are immune. Such a picture can natutally emerge in formerly vaccinated population. To this end, we apply the synchronous cellular automata algorithm studying stationary states and spatial patterning in SI, SIS and SIR models on a scale free networks in a range of parameter a with the fraction p of active sites. Various immunization scenarious, in particular the random immunization and more effective optimized strategy targeting the nodes with highest degree [2] are considered. A concept of ``safety patterns" of susceptible agents surrounded by immune individuals naturally arises in a proposed system, which plays an important role in the course of epidemic processes under consideration. Estimates for the threshold values of the basic reproduction numbers as functions of active agents fraction p and parameter a are obtained and efficiency of vrious vaccination techniques are analyzed quantitatively.

References

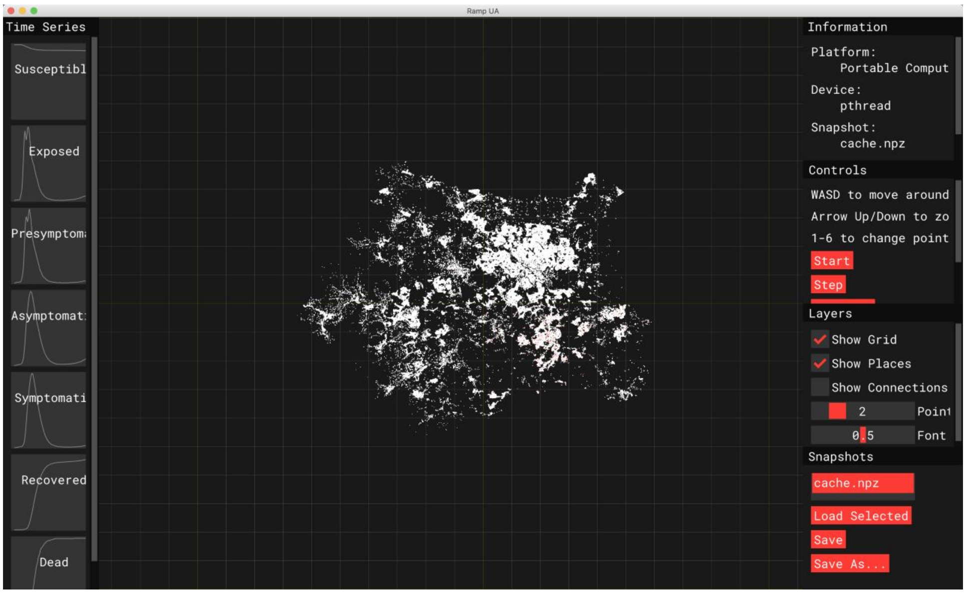
[1] A.-L. Barabasi and R. Albert, *Science* 286, (1999) 509

[2] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. *Rev. Lett.* 86, (2001) 3682.

National Scale Agent-Based Model for Epidemics

***Hadrien Salat**[1], **Anna Zanchetta**[1], Fernando Benitez[1], Man Luo[1], M. Birkin[1,2] (1) The Alan Turing Institute, 96 Euston Road NW1 2DB, London, UK, hsalat@turing.ac.uk. (2) The University of Leeds, UK.*

The Dynamic Microsimulation for Epidemics (DyME) model has been developed by the Turing in collaboration with several universities and external partners. It introduces a SEIR-type model into a pre-existing model of human social behaviour and contacts, combining synthetic population microsimulation (SPENSER) and transport flows modelling (QUANT). Initially calibrated for the County of Devon in England in 2020, the model has now been extended to the national scale. It is computationally efficient, flexible and modular: it can take any list of MSOAs (small geographical units of about 8000 inhabitants) as input, it can account for lockdown measures and can accommodate a variety of venues that potentially generate a risk of disease transmission. As such, it can be used as a basis for developing any similar model. We demonstrate how the model can be applied to assessing the impact of reopening large social gatherings at the end of a lockdown.



Platform:
Portable Comput
Device:
pthread
Snapshot:
cache.npz

WASD to move around
Arrow Up/Down to zo
1-6 to change point

☑ Show Grid
☑ Show Places
☐ Show Connections
2 Point
0.5 Font

Figure 1: Interactive results dashboard

Temporal Epidemiological Model of Community Transmission of an Airborne Transmitted Disease with Changing Network Dynamics

*A. Abbey[1], Y. Marmor[1], Y. Shahar[2], **O. Mokryn**[1]. (1) Information Systems, University of Haifa, Israel. ossimo@gmail.com. (2) Software and Information Systems Engineering, Ben Gurion University, Israel.*

Communities are complex social systems in which emergent properties - such as the dynamic nature of a viral infection - result from their members' local, uncoordinated *temporal interactions*. Models that consider Non-Pharmaceutical Interventions (NPIs) strategies and simulate them must take into account aspects that pertain to the interplay between the dynamic property in question, here the spread of the virus, and the dynamic community's characteristics, among which are the nature of the temporal interactions, and the characteristics of the participants, such as age and comorbidity information.

To address these needs, we use a known temporal dataset that has registered real-life interactions within a community for four weeks using proximity sensing. We enrich the data with virus-related information, such as a per-person susceptibility to infection, dependent on age and comorbidity. In the heart of our contagion modeling is the assumption that when interacting with infectious nodes, the length of the interaction, denoted by the weight of the edge between the participants, is correlated positively with the latent transmitted viral load, which, in turn, is positively correlated with the probability of getting infected. This is a clinically accepted assumption, such as by the European Center for Disease Prevention and Control. We examine daily the probability of getting infected as one minus the probability of not getting infected in each of the encounters while considering higher-order interactions (gatherings). Once infected, personal susceptibility determines the possibility of becoming symptomatic and then the severity of the infection. Thus, we can model viral transmission within a community while giving a detailed account of projected hospitalizations and deaths.

We examine the effect of different age distributions and the effect of the temporal network density on the spread of the virus. We find that a reduction (increase) in the rate of daily meetings has a significant effect on the transmission and can reduce (increase) significantly the spread of the virus without shutting down activity.

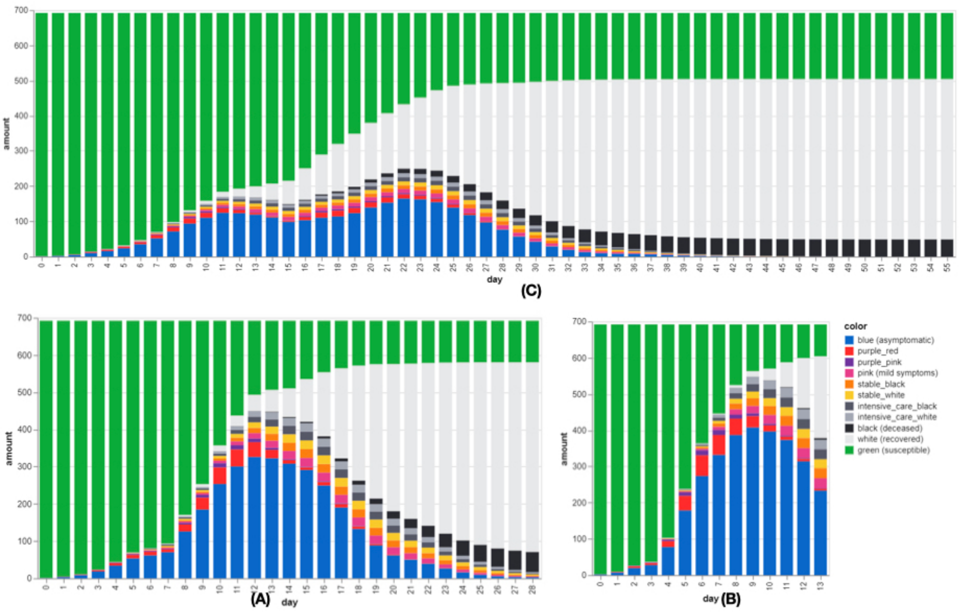


Figure 1: Transmission (A) normal Temporal Density (TD) (B) double TD (C) half TD

Robustness to outliers for Covid-19 Reproduction Number Estimation via NonSmooth Convex Proximal Optimization

Authors: B. Pascal[1], N. Pustelnik[2], R. Gribonval[3], S. Roux[2], P. Flandrin[2], P. Abry[2]
Affiliations: [1] CNRS,Cristal Université de Lille, [2] CNRS Laboratoire de Physique ENS de Lyon, [3] CNRS INRIA LIP ENS de Lyon

Abstract: The reproduction number of a pandemic, $R(t)$, is considered by epidemiologists as the critical quantity to monitor a pandemic, and to take timely and relevant sanitary policy decisions [3]. In the context of the outburst of the Covid-19 pandemic, we devised an estimator for $R(t)$, that combined the key ingredients of a well-documented incidence epidemiology model [2] with the regularity properties expected from $R(t)$, via an inverse-problem formulation [1]. This was solved by nonsmooth convex optimization, and applied to daily new infection counts, yield estimates of $R(t)$ for 200+ countries, automatically updated on a daily basis (cf. `http://perso.ens-lyon.fr/patrice.abry/`). However, because of the emergency sanitary crisis induced by the Covid-19, the data (e.g., daily infection counts) made available by the Health Authorities from almost all countries, were severely corrupted by outliers or missing values. Surprisingly, more than one year after the pandemic outburst, data are still significantly corrupted with outliers, cf. e.g., Fig. 1 top row reporting (black) the daily new infection counts for France. In [1], data were preprocessed for outlier removal prior to the estimation of $R(t)$. This two-step procedure (outlier denoising, R estimation) is sub-optimal and implies the use of a strong temporal regularization in the estimation procedure to avoid bias, at the cost of loosing the capability of detecting rapid changes in R, such as those induced e.g., by sanitary politicy decisions. Elaborating on [1], we propose here to combine outlier and reproduction number estimations into a single and joint step. This is achieved by including outlier terms in the inverse problem formulation, both into the data-model fidelity term and in the regularization term. This new functional formulation is carefully crafted to remain convex, while remaining nonsmooth and the corresponding minimization algorithm is devised. Fig. 1 illustrates the benefits (second row) of the proposed single-step estimation (red) over the two-step procedure (blue), showing that the former is far less biased by outliers than the latter and thus detect accurately changes in the pandemic dynamics. Estimated outliers are reported in the bottom raw, while the top row compares the raw infection counts (black) to the outlier-denoised infection counts (red).

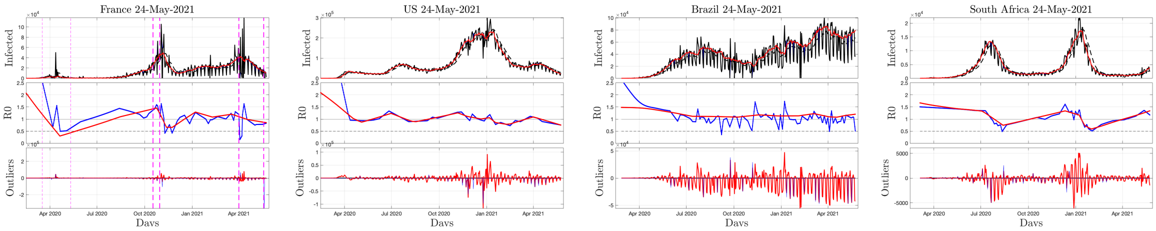


Figure 1: **Robust estimation for the reproduction numbers of 4 different countries:** Top: Daily new infections, raw (black) versus data after removal of estimated outliers (red) data ; Middle: Estimates of $R(t)$ combined (red) or not (blue) with outlier estimation ; Bottom: Estimated outliers.

References

[1] P. Abry et al. Spatial and temporal regularization to estimate covid-19 reproduction number R(t): Promoting piecewise smoothness via convex optimization. *PlosOne*, Aug. 2020. URL: `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0237901`.

[2] A. Cori, N. M. Ferguson, C. Fraser, and S. Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, 178(9):1505–1512, 2013. URL: `https://academic.oup.com/aje/article/178/9/1505/89262`.

[3] R.N. Thompson et al. Improved inference of time-varying reproduction numbers during infectious disease outbreaks. *Epidemics*, 29:100356, 2019.

Confidence intervals for Covid-19 Reproduction Number Estimation combining Nonsmooth Convex Proximal Optimization and Stochastic Sampling

Authors: Hugo Artigas[1], Gersende Fort[2], Barbara Pascal[3], Nelly Pustelnik[4], Patrice Abry[4]
Affiliations: [1] Ecole Polytechnique, [2] CNRS, Institut de Mathématiques de Toulouse, [3] CNRS, Cristal Université de Lille, [4] CNRS Laboratoire de Physique ENS de Lyon

Abstract: The reproduction number of a pandemic, $R(t)$, is considered by epidemiologists as the critical quantity to monitor a pandemic, and to take timely and relevant sanitary policy decisions. In the context of the outburst of the Covid-19 pandemic, we devised an estimator for $R(t)$, that combined the key ingredients of a well-documented incidence epidemiology model [2] with the regularity properties expected from $R(t)$, via an inverse problem formulation [1], solved by nonsmooth convex optimization. This was further improved to take into account the severe corruption of data by outliers and applied to daily new infection counts. This yields estimates of $R(t)$ for 200+ countries, automatically updated on a daily basis (cf. `http://perso.ens-lyon.fr/patrice.abry/`). However, the proposed convex optimization-based estimation procedures suffer from a major limitation for an actual practical use by epidemiologists: Being deterministic in nature, they do not naturally output confidence intervals for the estimations.

We propose here to revisit these estimation procedures (without or with outliers) within Bayesian frameworks [3], where $R(t)$ and the outliers are the parameters to estimate. A first contribution is then to propose new Markov chain Monte Carlo samplers which match nonsmooth convex optimization formulations to more efficiently assess the parameter a posteriori distribution. Second, from these estimated a posteriori distributions several Bayesian estimators are computed, and complemented with measures of uncertainties. We also show that the output of the minimization procedure in [1] corresponds to one such estimators, the Maximum a Posteriori.

Fig. 1 reports, for the USA, the estimated quantiles for the a posteriori distributions for the estimates of $R(t)$ (left) and outliers (right), superimposed to the estimates obtained from the nonsmooth proximal convex optimization procedure.

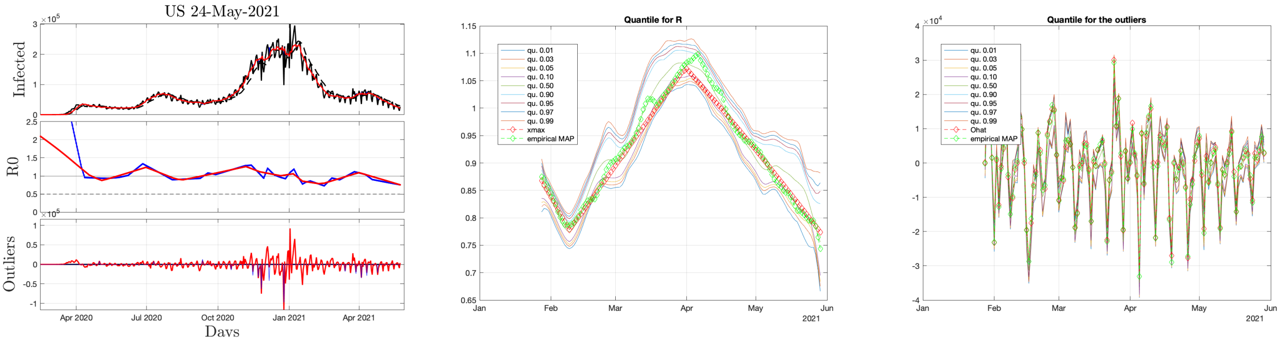


Figure 1: **Reproduction number and Outlier estimation and a posteriori distribution assessments.** Left: Top: Daily new infections, raw (black) versus data after removal of estimated outliers (red) data ; Middle: Estimates of $R(t)$ combined (red) or not (blue) with outlier estimation ; Bottom: Estimated outliers. Middle and left: Quantiles from the estimated a posteriori distributions.

References

[1] P. Abry et al. Spatial and temporal regularization to estimate covid-19 reproduction number R(t): Promoting piecewise smoothness via convex optimization. *PlosOne*, Aug. 2020. URL: `https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0237901`.

[2] A. Cori, N. M. Ferguson, C. Fraser, and S. Cauchemez. A new framework and software to estimate time-varying reproduction numbers during epidemics. *American Journal of Epidemiology*, 178(9):1505–1512, 2013. URL: `https://academic.oup.com/aje/article/178/9/1505/89262`.

[3] O. Stramer and R.L. Tweedie. Langevin-type models II: Self-targeting candidates for MCMC algorithms. *Methodology and Computing in Applied Probability*, 1:307–328, 1999.

Optimizing lock-down measures of epidemic spreading in mobility networks

***B. A. Siebert*[2]*,* *M. Asllani*[1]*, J. P. Gleeson*[2] *(1) University College Dublin, Ireland* *malborasllani@gmail.com*. *(2) University of Limerick, Ireland*

The spread of infectious diseases through human populations has long been of interest to scientists. Recently, the Coronavirus pandemic has led to renewed interest in intervention strategies to reduce the spread, especially in the absence of a vaccine. While some countries have undertaken a “zero-covid” strategy, many governments are unwilling to implement this strategy, instead preferring to reduce human mobility sufficiently to keep the spread of disease low, while still having enough mobility to protect their economies. Reducing the peak is highly desirable, as a significant number of cases can lead to severe disruption of the healthcare system. However, this comes at the cost of mobility of individuals on the network. Thus, it is desirable to reduce spread of the disease, while allowing as much mobility as possible.

Therefore, we present a reaction-diffusion model for epidemic spread on a metapopulation network, where nodes represent locations where individuals congregate, and disease spread within nodes. Individuals move between nodes via diffusion. Classically, diffusion is modelled using a random walk Laplacian on networks. However, we have modified diffusion to take the crowding on a node into consideration, that is, the transmission probability from a node to a target node depends on the spare capacity of this target node. If the target node has a very low spare capacity, but a different neighbor of the original node has a high spare capacity, then individuals will prefer to move to the node with the higher spare capacity. In effect, we are pushing individuals to avoid busy nodes with a high concentration of individuals, towards nodes with a lower concentration of individuals, effectively modelling social distancing. We show that modifying diffusion in this way reduces the peak number of infected individuals. Furthermore, we have found that due to the meta-population setting, the infection spreads in waves, modelling the second and third infection waves we have seen throughout the world.

To measure the effect of social distancing on the mobility of individuals in the network, we borrow the entropy measure from information theory which relies on the sum of the transmission probabilities at equilibrium. Using this measure, we compare two different intervention strategies. The first, is to simply lower the diffusion constant, resulting in an equal reduction in mobility for all nodes. The second strategy is to use the crowding parameter to reduce access to busier nodes more so than more empty nodes. We compare the strategies for equal values of the entropy and find that a more targeted intervention can reduce the peak number of infected individuals better than a blanket reduction in mobility, without a significant additional cost in entropy. Furthermore, we can generalize the model so that each node has its own crowding parameter, and we find that having a crowding parameter directly proportional to the degree of a node can further reduce the peak of the infection. Finally, we apply our model to a number of real-world networks, such as the contiguous USA and the London public transport network.

Individual Risk Perception and Empirical Social Structures Shape the Dynamics of Infectious Disease Outbreaks

V. d'Andrea[1]*, R. Gallotti*[1]*, N. Castaldo*[1] *and Manlio De Domenico*[1]*. (1) CoMuNe Lab, Fondazione Bruno Kessler, Via Sommarive 18, 38123 Povo (TN), Italy vdandrea@fbk.eu.*

The dynamics of a spreading disease and individual behavioral changes are entangled processes that have to be addressed together in order to effectively manage an outbreak. However, people compliance with mitigation norms is not homogeneous across a population and depends on individual risk perception. In this work we investigate the impact of people with low risk perception (risk-denier people) on epidemic spreading. In particular we: i) use a large-scale survey conducted worldwide via Facebook to relate individual level of risk perceptions to compliance with specific behavioral norms; ii) develop a mathematical model based on a suitable generalization of a Susceptibles-Exposed-Infectious-Recovered (SEIR) model with compartmental states that explicitly takes into account for the adoption of the selected mitigation norms; iii) evaluate epidemic dynamics using both mean-field approximation and agent based simulations and we study the effects of different amounts of non-compliant, risk-denier people on epidemics; iv) to understand how network features typical of human interaction patterns impact the spread of a disease, we first run simulations on multiple synthetic network models characterized by increasing complexity (e.g., level of clustering, modular structure) and then perform a set of numerical experiments on more than 200 real-world networks. In both synthetic and real contact patterns we find that community structures and degree heterogeneity hinder epidemic spreading, but this reduction is most effective at low percentages of risk-denier people.

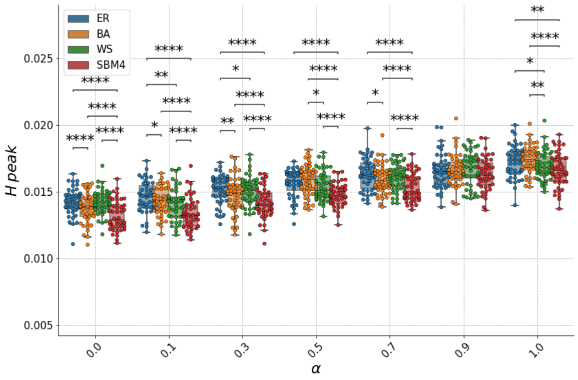


Figure 1: Results on synthetic networks: epidemic size (evaluated as peak of hospitalized people) as a function of the percentage of risk-denier people (α)

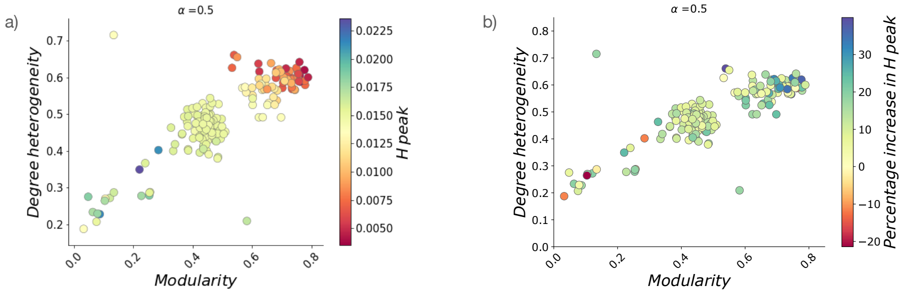


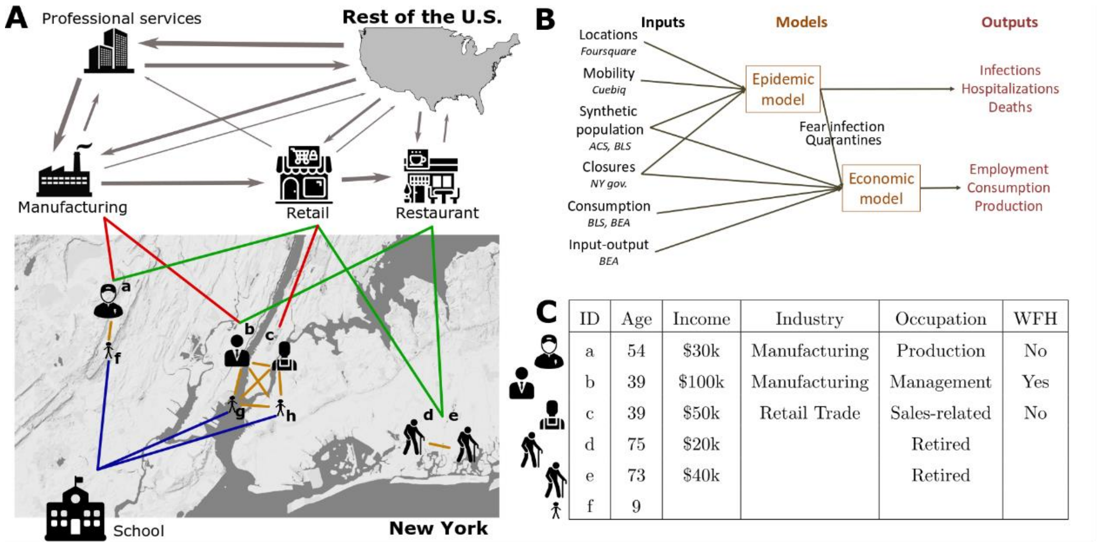
Figure 2: Results on real networks, at fixed percentage of risk-denier people: a) hospitalized peak and b) hospitalized peak normalized to the value at α = 0.0.

Modeling the distributional epidemic-economic effects of the Covid-19 pandemic

Marco Pangallo[1], *Alberto Aleta*[2], *R. Maria del Rio-Chanona*[3], *Doyne Farmer*[3,4], *François Lafond*[3], *David Martin-Corral*[5], *Yamir Moreno*[6,2], *Esteban Moro*[5,7], *Anton Pichler*[3], *Alessandro Vespignani*[8,2], *(1) Sant'Anna School of Advanced Studies, Pisa, Italy, marco.pangallo@santannapisa.it. (2) Institute for Scientific Interchange Foundation, Turin, Italy (3) University of Oxford, UK (4) Santa Fe Institute, Santa Fe, US (5) Universidad Carlos III de Madrid, Leganes, Spain (6) University of Zaragoza, Spain (7) MIT, Cambridge, US (8) Northeastern University, Boston, US*

Understanding the joint epidemic and economic effects of the Covid-19 pandemic across socio-economic groups would help design targeted non-pharmaceutical interventions. However, most existing epidemic-economic models are either too aggregate to study heterogeneous outcomes across socio-economic groups, or, when sufficiently fine-grained, not grounded enough on data. In this paper, we introduce a large-scale, data-driven Agent-Based Model that simulates epidemic and economic outcomes across industries, occupations, and income levels (Figure 1).

We calibrate our model to the first wave of Covid-19 in the New York metro area. At the epidemic level, our simulations show that, during the lockdown, richer individuals who work in occupations that can be performed from home got less infected than poorer individuals working in person in essential industries, although they got more infected before the lockdown. Most infections occur in a few "customer-facing" industries, such as retail, entertainment and bars and restaurants. At the economic level, we also find that richer individuals got less unemployed but reduced consumption more than poorer individuals; this is a well-established empirical pattern that lacked theoretical justification. Unemployment is concentrated in industries whose work activities cannot be performed from home, and in poorer areas such as the Queens and the Bronx.



ID	Age	Income	Industry	Occupation	WFH
a	54	$30k	Manufacturing	Production	No
b	39	$100k	Manufacturing	Management	Yes
c	39	$50k	Retail Trade	Sales-related	No
d	75	$20k		Retired	
e	73	$40k		Retired	
f	9				

Figure 1: Schematic of the epidemic-economic model. A: Epidemic and economic linkages in the New York (NY) metro area. Individuals (a-h) are connected to members of their household (orange links), to school if children (blue), to a workplace if employed adults (red) and to the consumption venues that they go to (green). Industries (manufacturing, retail, etc.) are connected by an input-output network of intermediate products (grey links). B: Inputs and outputs of the epidemic and economic models. Cuebiq = privacy-enhanced anonymous mobile phone mobility data; ACS=American Community Survey; BLS=Bureau of Labor Statistics; BEA=Bureau of Economic Analysis. The epidemic model is an input to the economic model through fear of infection and quarantining. C: Socio-economic characteristics of the synthetic individuals shown in panel A. WFH=Work From Home

Agent-based model analysis of reactive vaccination against COVID-19 epidemic

***B. Faucher*[1]*, R. Assab[1], J. Roux[2], D. Levy-Bruhl[3], C. Tran Kiem[4], S. Cauchemez[4], L. Zanetti[5], V. Colizza[1], P.Y. Boëlle[1], C. Poletto[1]*, *(1) INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health, Paris, France, benjamin.faucher@iplesp.upmc.fr. (2) Univ Rennes, EHESP, REPERES « Recherche en Pharmaco-Epidémiologie et Recours aux Soins » – EA 7449, 15 avenue du Professeur-Léon-Bernard, CS 74312, 35043 Rennes, France, (3) Santé Publique France, Saint Maurice, France, (4) Mathematical Modelling of Infectious Diseases Unit, Institut Pasteur, UMR2000, CNRS, Paris, France, (5) Haute Autorité de Santé, Paris, France*

One year after COVID-19 was declared a pandemic, a dozen vaccines have already been approved for use. Fast and targeted vaccination has become the main goal to reduce morbidity and mortality, often with prioritization criteria based on the risk of developing severe infection – i.e. the elderly, individuals with comorbidities. However, once vaccines are secured to these population groups, reactive vaccination could be efficient, in combination with mass vaccination, to accelerate the decline of the epidemic or to curb local case resurgences. Reactive vaccination consists in targeting the individuals most at risk of infection, e.g. those in workplaces where virus circulation is detected. Quantifying its effect requires modelling the complex interplay between the time for the vaccine to become effective, natural history of infection, case detection, pace of vaccination, and overall changes in human contact behavior. Here, we use an agent-based model to test reactive vaccination of adults (18+ year old) targeting workplaces and schools. The model integrates demographic and contact surveys data for France to build a multi-layer temporal network of contacts, where layers represent different settings – household, school, workplace, community, transport. We simulate a COVID-19 epidemic and vaccine distribution, exploring different availability and logistic constraints. We find that with the same number of vaccine doses, the combination of reactive and random (mass) vaccination is more efficient than random vaccination alone in reducing the number of cases. Results are robust across a wide range of epidemic parameters. The model allows to quantify the cost of the strategy in terms, e.g., of daily number of workplaces that need to be vaccinated and peaks in vaccine demand. These results provide valuable information on the inclusion of this strategy in COVID-19 vaccination plans once the at-risk population is vaccinated.

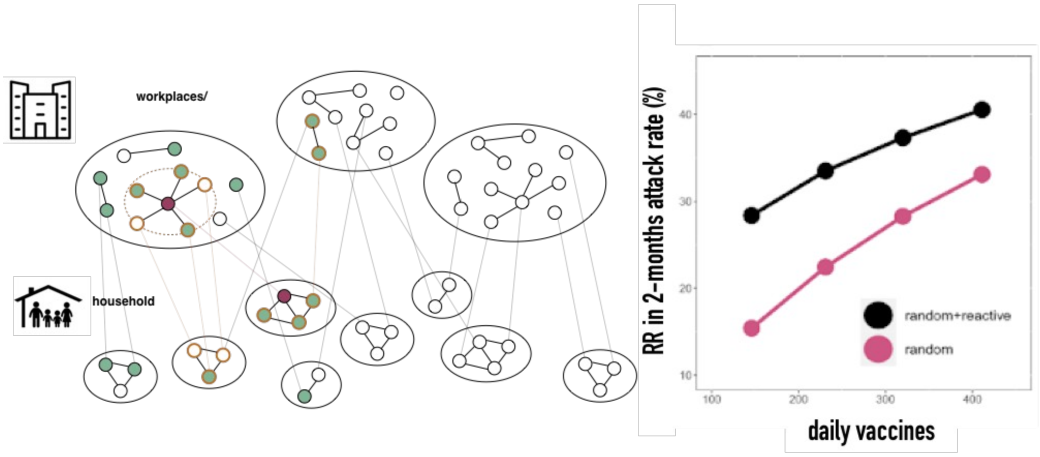


Figure 1: Left: scheme of the model. Right: Impact of vaccination on the epidemic decline. Relative reduction (RR) in the attack rate the first 2 months after the campaign start. Combined reactive and random vaccination is compared with random vaccination alone, at same number of doses administered during the period. We consider here an epidemic with reproductive ratio R= 1.2 and initial weekly incidence 150 per 100,000 inhabitants.

Socioeconomic Status Resulting in Infection Rate Disparity, an Agent Based and Mean Field Model

Sina Sajjadi[1], *Pourya Toranj*[2], *Basak Taraktas*[3], *Mehrzad Shadmangohar*[2], *Ulya Bayram*[4], *Mavi Ruiz*[5] *(1) Complexity Science Hub Vienna, Austria, sinasajjadi@protonmail.com.(2) Shahid Beheshti University, Tehran, Iran. (3) Bogazici University, Istanbul, Turkey. (4) Canakkale Onsekiz Mart University, Canakkale, Turkey. (5) Neurable, Boston, United States of America*

Official reports on the COVID-19 pandemic reveal disproportionate infection rates in lower socioeconomic groups whose economic insecurities outweigh concerns about the infection risk. Motivated by this issue, our research introduces a model for studying the effects of income inequality on infection rates during pandemics. We analyze the disparity of infection cases across different race and income groups within a context of residential segregation by constructing a novel agent based model which combines:

1. A network model for segregated communities using inter-group exposure indices. 2. A decision making model governing agents' choices to go out to work vs. quarantining themselves. 3. An SIR model describing the infection spreading.

We develop both an individual-based model (ABM) and a mean-field model (ODE) and find that their results are in agreement. We find that infection rate is directly influenced by the existing level of income. More importantly, socioeconomic inequality increases the overall infection rate of the population. Our results convey a significant message for policy makers. Governments can slow down the spread of the infection by providing financial aids to low-income households; improving the latter's financial security will enable them to stay in. This study offers a modeling framework to understand socioeconomic factors affecting the dynamics of infection.

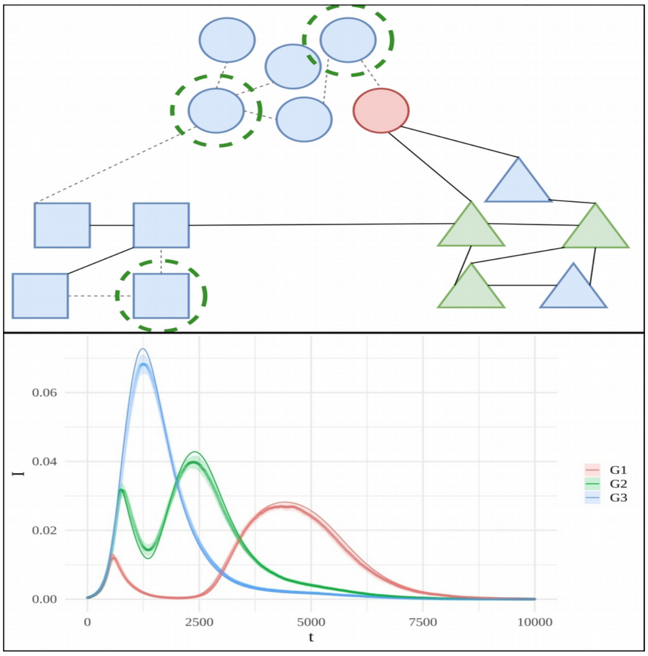


Figure 1: a) Illustration of the quarantining model, where markers denote different socioeconomic classes and colors indicate disease status (S: blue, I: red, R: green) dashed green circle indicates quarantine. b) The infection rate over time for the three different classes (low SES: blue, medium SES: green, high SES: red).

Influenza Decline During COVID-19 Pandemic: a Global Analysis Leveraging Classification and Regression Trees

***F. Bonacina**[1,2], P-Y. Boelle[1], V. Colizza[1], O. Lopez[2], M. Thomas[2], C. Poletto[1]. (1) INSERM, Sorbonne Université, Pierre Louis Institute of Epidemiology and Public Health, Paris, France, francesco.bonacina@iplesp.upmc.fr. (2) Sorbonne Université, CNRS, Laboratoire de Probabilités, Statistique et Modélisation, LPSM, Paris, France.*

The COVID-19 pandemic has caused a profound shock on the ecology of infectious diseases. Influenza offers a paradigmatic case. Its circulation reduced to very low levels or nearly zero after COVID-19 emergence. Since influenza and SARS-COV-2 share common transmission mechanisms, interaction between the two viruses can occur at many levels. In particular, change in human encounters due to non-pharmaceutical interventions (NPIs) against COVID-19 could have hindered influenza propagation. This complex interplay has been addressed so far by a number of studies focusing on specific countries. Here we address the problem at the global scale analyzing the FluNet influenza public repository for the periods before (2015-19) and during (2020-21) COVID-19 pandemic and leveraging machine learning. We map the space-time variation of influenza across countries. The percentage of positive tests decreased, globally, by 98.6% compared to the pre-COVID-19 period, despite surveillance remained broadly the same. The ratio of positivity rate in the pre-COVID-19 vs. the COVID-19 period was very heterogeneous across countries and seasons ranging from 6.3 to $1.4e^{-6}$. We thus used regression trees to quantify the impact of covariates such as COVID-19 incidence, strictness of NPIs, change in human mobility, demography, temperature and geographical region. We found that influenza decline was mainly associated with COVID-19 incidence (positive association) and region, and was also linked to median age, mobility and temperature. Our study adds important elements to the understanding of the influenza circulation in the COVID-19 era and provides a basis for monitoring future influenza re-emergence.

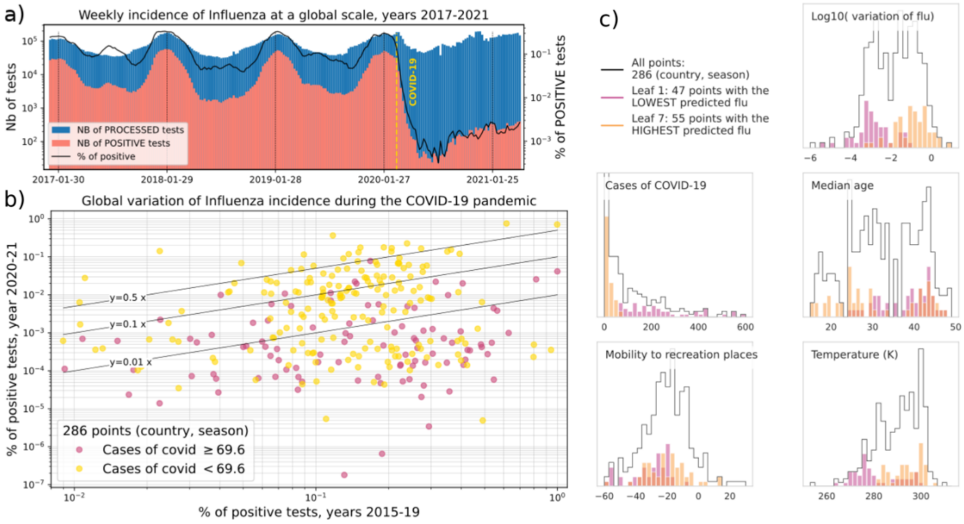


Figure 1: **a)** Weekly counts of processed and positive tests of influenza reported from 168 countries from Jan 2017 to Apr 2021. **b)** Comparison of the percentage of positive tests for the pre-COVID-19 period (2015-19) and the COVID-19 period (2020-21), considering 4 seasons and countries with at least 130 processed specimens per season. **c)** Results of the regression tree analysis. Distributions of the observed variation of influenza, daily COVID-19 cases per million inhabitants, median population age, mobility variation and temperature, for all countries/seasons and the ones with lowest and highest level of predicted influenza.

Optimising the mitigation of epidemic spreading through targeted adoption of contact tracing apps

Aleix Bassolas
Departament d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona, Spain

Andrea Santoro
Institute of Bioengineering, Center for Neuroprosthetics, EPFL, Geneva, Switzerland

Sandro Sousa, Silvia Rognone, and Vincenzo Nicosia
School of Mathematical Sciences, Queen Mary University of London, Mile End Road, E1 4NS, London (UK)

The ongoing COVID-19 pandemic is the first epidemic in human history in which digital contact-tracing has been deployed at a global scale. Tracking and quarantining all the contacts of individuals who test positive to a virus can help slowing-down an epidemic, but the impact of contact-tracing is severely limited by the generally low adoption of contact-tracing (CT) apps in the population. We consider here a classical Susceptible-Infected-Recovered model on a connected undirected graph $G = (V, E)$, with an additional contact tracing dynamics (SIR+CT), where a fraction r of the nodes have installed a contact-tracing app. We focus on the problem of determining the set $V' \subset V$ of nodes which should install CT apps in order to optimise the effect of contact tracing, i.e., to maximally slow-down spreading and reduce the incidence of a disease, under the assumption that the rate of CT app adoption is fixed.

We started from the observation that contact-tracing apps can only signal potentially infectious contacts to other people with a CT app installed, thus suggesting that one should aim at maximising the density of the contact-tracing subgraph G', consisting of all the nodes of G with a CT app installed and of all the edges among those nodes.

We provide an analytic derivation that shows that the value of the basic reproduction number of the SIR+CT dynamics can be minimised by using a generic optimisation algorithm to compute:

$$\max_{G'} \quad \mathcal{F}(G') = \sum_{\ell} \frac{1}{k_\ell} \sum_i a_{\ell i} \frac{k'_i}{k_i}(k_i - 1) \qquad (1)$$

over the ensemble of the possible choices of the contact-tracing subgraph G', where k'_i denotes the degree of node i in G'.

We show that in general the uniform random installation – which is the strategy implicitly adopted by governments when people are simply asked to install a CT app – causes a pretty slow decrease of R_0 as a function of the fraction r of nodes in G' (top panels of Fig.1). At the same time, the density of the contact-tracing subgraph G' under random installation is normally pretty low. Conversely, targeting nodes preferentially by their degree produces a contact-tracing subgraph that is in general quite dense, with a substantial reduction of R_0, at the unreasonable cost of requiring exact knowledge about the original graph G. We propose a simple heuristic strategy based on the friendship paradox, that we call Distributed Targeting Installation (DTI), and we show that DTI allows to sensibly reduce R_0 while not requiring global knowledge on G, both in synthetic and real-world networks.

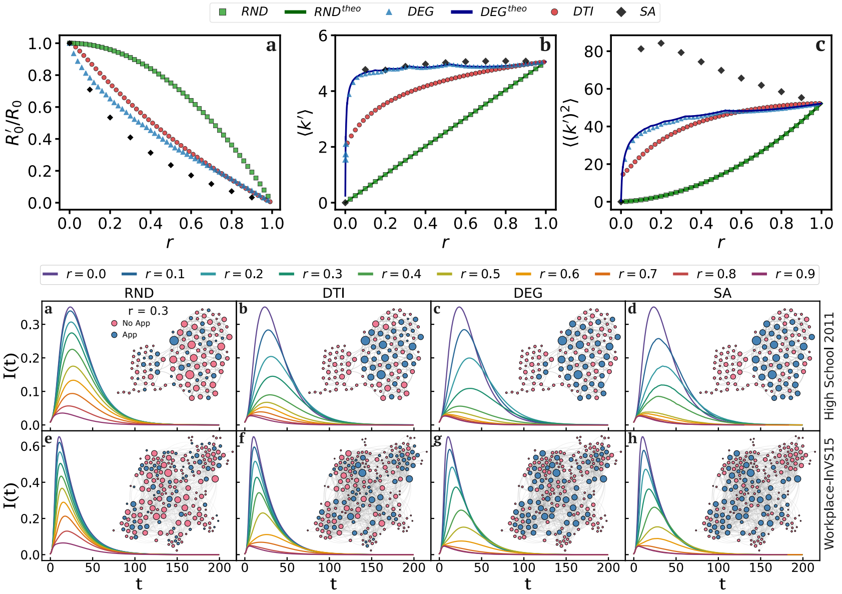


FIG. 1. The effectiveness of contact tracing in slowing down the spreading of a disease depends crucially on the actual strategy used to select the nodes that will install a contact-tracing apps. Top panels: the reduction in the basic reproduction number (left), aside with the average degree (middle) and the average squared degree (right) of the contact-tracing subgraph G'. Notice that random app installation performs quite badly. The epidemic curves and CT subgraphs in the social network of a highschool (middle panels) and of a workplace scenario (bottom panels) confirm that the proposed Distributed Targeting Installation strategy is a good trade-off between efficiency and simplicity.

These findings shed new lights on the effectiveness of contact-tracing in slowing down the spread of a disease, showing that the structure of the contact-tracing subgraph has a pivotal role, and can inform further analysis that take into account more realistic spreading models.

Vaccination Strategies for Minimizing the Number of Deaths From Emerging Diseases

T. Hiraoka, *M. Kivelä, J. Saramäki. Department of Computer Science, Aalto University, 00076 Espoo, Finland, takayuki.hiraoka@aalto.fi.*

Immunizing a population by vaccination is an effective way to protect individuals and society against infectious diseases. Network epidemiology advocates that vaccination targeting "hub" individuals generally outperforms the uniform vaccination in containing epidemics in a heterogeneously networked population, where a "hub" can be defined using degree or other centrality measures. In reality, however, it is common to target high-risk groups to be vaccinated first, where risk refers to the chance of developing grave medical conditions, including death, upon infection. The choice of strategy to be implemented primarily depends on the objective of the vaccination: Is it the cumulative number of infected individuals (epidemic size), the number of deaths, or the years of potential life lost that it tries to minimize? This question is especially important when (1) there is a large disparity in infection fatality risk (IFR) within the population and (2) the amount and/or the rate of vaccine distribution is limited.

Here, we analyze the susceptible-infected-removed (SIR) model with vaccination to investigate the difference between vaccination strategies in the effectiveness of minimizing the number of deaths. To this end, we introduce heterogeneous distributions of interactivity and IFR within the population. We assume that the population is vaccinated gradually at a time scale comparable to that of the ongoing epidemics. By numerically solving heterogeneous mean-field differential equations, we show the following: (1) Although the hub-prioritized vaccination (prioritizing those who interact more with others) is always the optimal strategy for suppressing the epidemic size, it is not always the best one to reduce the number of deaths. The vulnerable-prioritized strategy (prioritizing those who have higher IFR) performs better in this regard when the transmissibility is high, and its advantage over the interactivity-targeted strategy is pronounced when the rate of vaccine distribution is large. (2) A delayed launch of vaccination generally shrinks the margin in effectiveness between two strategies and expands the region in which the vulnerable-prioritized strategy is favorable in the parameter space. Our results highlight the importance of clarifying the objective function to be optimized when designing a vaccination strategy against an emerging disease such as COVID-19. Importantly, minimizing the epidemic size does not necessarily lead to minimizing deaths.

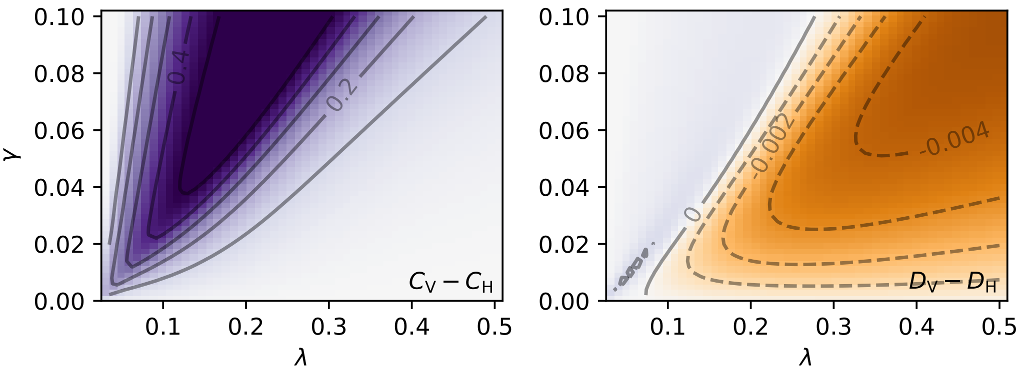


Figure 1: Differences in the epidemic size (left) and the number of deaths (right) between the vulnerable-prioritized (subscript V) and hub-prioritized (H) vaccination strategies. The purple (orange) regions are where the hub-prioritized (vulnerable-prioritized) strategy performs better.

A Sustainable Strategy for Open Streets in Post-Pandemic Cities

***D. Rhoads**[1], A.Solé-Ribalta[1,2], M.C. González[3,4,5],J. Borge-Holthoefer[1]. (1) Internet Interdisciplinary Institute, Universitat Oberta de Catalunya, Barcelona, Spain, drhoads@uoc.edu. (2) URPP Social Networks, University of Zurich, Zurich, Switzerland (3) Department of City and Regional Planning, University of California, Berkeley, CA, USA (4) Energy Technologies Area, Lawrence Berkeley National Laboratory, Berkeley, CA, USA (5) Department of Civil and Environmental Engineering, University of California, Berkeley, CA, USA*

The challenges of the past year have led to notable shifts across many fields, and urban design is no exception. Many cities have taken the opportunity presented by the pandemic to improve and expand pedestrian infrastructure, at once lifting the strain from outdoor gathering spaces, providing residents with a sense of relief, and pursuing long-standing goals to decrease automobile dependence and boost urban walkability. So far, these efforts have lacked the systems-level view afforded by treating sidewalks as a network. Here [1], using sidewalk data from ten world cities, we construct city-wide sidewalk networks –adjacent, but irreducible to their corresponding road networks. In a no-intervention scenario, we apply percolation theory to examine how well our cities' sidewalk infrastructures can provide for physical distancing in a pandemic context. The resulting breakdown of the networks, even at moderate levels of physical distancing, calls for a cautious strategy, considering the interdependencies of sidewalk and road networks: any improvement for pedestrians comes at a cost for cars. With remarkable success, we propose a shared-effort heuristic to pacify and pedestrianize road segments, delaying sidewalk connectivity breakdown while preserving road network functionality.

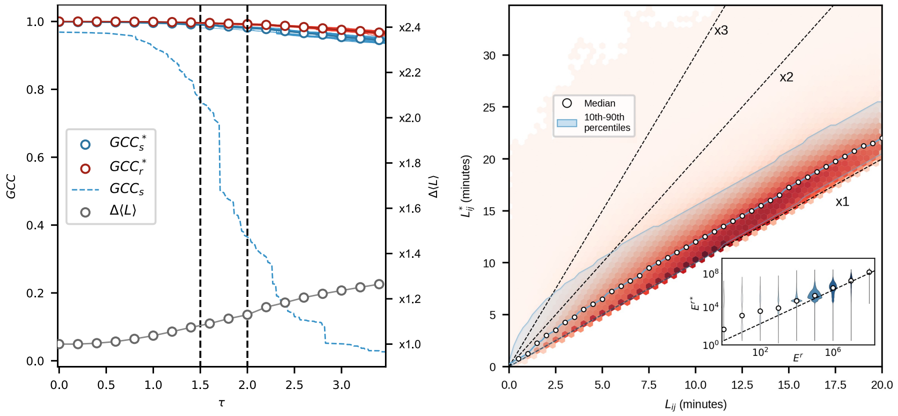


Figure 1, **Results, New York City.** *Left:* GCC_s indicates the breakdown of the largest component of the sidewalk network as the minimum effective sidewalk width τ (x-axis) rises without intervention and narrow sidewalks become dysfunctional. GCC_s^* shows improved sidewalk network connectivity post-intervention for various simulations. GCC_r^* shows the change in road network connectivity post-intervention. $\Delta\langle L\rangle$ shows the relative increase in road network average path length $\langle L\rangle$ due to road pacifications. *Right:* The distribution of all road network path lengths (i,j) pre- (L_{ij}) and post-intervention (L_{ij}^*), showing a globally constrained rise in travel times. *Inset*: Binned change in road edge betweenness pre- and post-intervention.

References

[1] Rhoads, D., Solé-Ribalta, A., González, M. C., & Borge-Holthoefer, J. (2020). Planning for sustainable Open Streets in pandemic cities. *arXiv preprint arXiv:2009.12548.*

Optimal Control of the Spatial and Temporal Allocation of COVID-19 Vaccines: Italy as a Case Study

***Joseph C. Lemaitre**[1,*]*, *Damiano Pasetto[2], Mario Zanon[3], Enrico Bertuzzo[2], Lorenzo Mari[4], Stefano Miccoli[5], Renato Casagrandi[4], Marino Gatto[4], Andrea Rinaldo[1,6]. (1) Laboratory of Ecohydrology, Ecole Polytechnique Fédérale de Lausanne, 1015 Lausanne (CH). (2) Dipartimento di Scienze Ambientali, Informatica e Statistica, Universita` Ca' Foscari Venezia, 30172 Venezia-Mestre (IT). (3) IMT Scuola Alti Studi, 55100 Lucca (IT). (4) Dipartimento di Elettronica, Informazione e Bioingegneria, Politecnico di Milano, 20133 Milan (IT). (5) Dipartimento di Meccanica, Politecnico di Milano, 20156 Milan (IT). (6) Dipartimento ICEA, Universita` di Padova, 35131 Padova (IT).*
**jo.lemaitresamra@gmail.com*

While SARS-CoV-2 vaccine distribution campaigns are underway across the world, communities face the challenge of a fair and effective distribution of limited supplies. We wonder whether suitable spatial allocation strategies might significantly improve a campaign's efficacy in averting damaging outcomes. To that end, we address the problem of optimal control of COVID-19 vaccinations in a country-wide geographic and epidemiological context characterized by strong spatial heterogeneities in transmission rate and disease history. We seek the vaccine allocation strategies in space and time that minimize the number of infections in a prescribed time horizon. We examine scenarios of unfolding disease transmission across the 107 provinces of Italy, from January to April 2021, generated by a spatially explicit compartmental COVID-19 model tailored to the Italian geographic and epidemiological context. We develop a novel optimal control framework to derive optimal vaccination strategies given the epidemiological projections and constraints on vaccine supply and distribution logistic. Optimal schemes significantly outperform simple alternative allocation strategies based on incidence, population distribution, or prevalence of susceptibles in each province. Our results suggest that the complex interplay between the mobility network and the spatial heterogeneities imply highly non-trivial prioritization of local vaccination campaigns. The extent of the overall improvements in the objectives grants further inquiry aimed at refining other possibly relevant factors so far neglected. Our work thus provides a proof-of-concept of the potential of optimal control for complex and heterogeneous epidemiological contexts at country, and possibly global, scales.

Acknowledgements (optional)

JCL and AR acknowledge funding from the Swiss National Science Foundation (snf.ch) via the project "Optimal control of intervention strategies for waterborne disease epidemics" (200021--172578). AR, EB and DP acknowledge funding from Fondazione Cassa di Risparmio di Padova e Rovigo (fondazionecariparo.it) through its grant 55722.

Predicting public health impact of COVID-19 in Mexico City using mobility and contact networks

A major issue during the COVID-19 crisis has been how to predict public health requirements, such as the amount of hospital beds and associated resources to be destined to COVID patients. In this regard, epidemiological forecasts have been an important tool for policy-makers. However, since the spread of SARS-CoV-2 is intimately tied to human mobility and congreggations of people, model parameters evolve at a rapid pace, impacting the accuracy of medium-term predictions.

Using mobile device position data, we have been able to describe and characterize daily mobility and contact patterns using network models. Using this information along with daily epidemiological data reported by the local authorities, we have deviced a long short term memory neural network model that estimates daily epidemiological tallies including hospitalizations and deaths for Mexico City. By incorporating mobility and contact network information into the model, we increase the predictive power of our model (MSE = 0.27 using only epidemiological information; MSE = 0.16 incorporating mobility and contact information).

The current model is unique in the sense that it allows decision makers to simulate non pharmaceutical interventions in the form of mobility constraints and visualize the evolution of the pandemic under different scenarios.

The Effects of the Quarantine and Vaccination on the COVID-19 Spread by Combining the Compartmental and Cellular Automaton Types of Modelling

T. Patsahan*, J. Ilnytskyi. Institute for Condensed Matter Physics, 1 Svientsitskii Street, Lviv, Ukraine,*
tarpa@icmp.lviv.ua

We consider the model for the COVID-19 spread, which emphasizes such characteristic features of the COVID-19 pandemic as: the abundance of unidentified infective individuals and gradual loss of immunity after recovery, and also includes the possibility of vaccination. The compartmental realisation of the model is examined both analytically and via numeric solution. The former provides means for finding the disease-free and endemic stationary states and their stability, as well as obtaining the basic reproductive number as the function of the model parameters. Numeric solution enables analysis of the model dynamics, of which the position and the height of the first and consecutive peaks are of the most interest. We used the scaling forms with effective exponents to fit the numeric results and, as the result, obtained simple algebraic expressions for the positions and heights of the peaks for the number of infected individuals as the functions of the model parameters. Several quarantine scenarios are examined for their effectiveness to bring the pandemic down. The cellular automaton realization of the model is considered on a square lattice of the size about 10^6 individuals, representing the population scale of Lviv city taken in our study as a case study. We consider both cases of the constant neighbourhood size q throughout the system, and of the variable neighbourhood size, which is obtained from various distributions. This type of modelling provides natural means for mimicking the quarantine (social distancing) measures, as well as the mobility of individuals, and is aimed on obtaining some important features and tendencies of the pandemic dynamics depending on the model parameters in the population of a realistic scale. We also consider various scenarios for vaccination in all cases being considered. The grid-like model is also developed, which utilizes the cellular automaton algorithm for the city divided by small city zones, whereas the compartmental type of modelling is employed inside each zone.

Acknowledgements

This work is supported by the National Research Foundation of Ukraine (grant agreement No 123/01/0338).

References

[1] D. Adam, “Special report: The simulations driving the world’s response to COVID-19”, *Nature*, 580 (2020) 316.

[2] M. Park, A.R. Cook, J.T. Lim, Y. Sun, B.L. Dickens, “A Systematic Review of COVID-19 Epidemiology Based on Current Evidence”, *Journal of Clinical Medicine*, 9 (2020) 967.

[3] J.M. Ilnytskyi, “Modeling of the COVID-19 pandemic in the limit of no acquired immunity”, *Mathematical Modeling and Computing*, 8 (2021) 282-303.

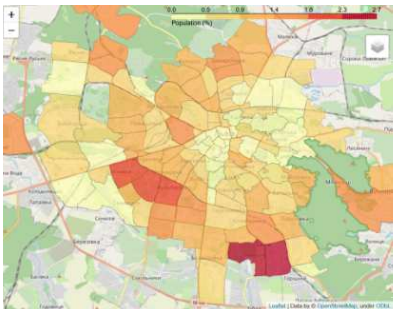
[4] J. Ilnytskyi, Y. Kozitsky, H. Ilnytskyi, O. Haiduchok, “Stationary states and spatial patterning in an SIS epidemiology model with implicit mobility”. *Physica A: Statistical Mechanics and its Applications*, 461 (2016) 36.

Modeling the Spatiotemporal Spread of the COVID-19 in an Urban Area

***T. Patsahan**, J. Ilnytskyi. Institute for Condensed Matter Physics, 1 Svientsitskii Street, Lviv, Ukraine, tarpa@icmp.lviv.ua*

Human mobility is a key factor in spreading infectious diseases such as COVID-19. Especially, it concerns urban areas where population is dense and highly mobile. Moreover, a human mobility in urban areas is strongly related with professional activities of population. For instance, a person moves from his house to workplace, where he contacts with colleagues, people in transport, in shops and in other places on his way. Then a person returns to his residence, where he contacts with his family, neighbours, friends and others. Before starting the COVID-19 vaccination a restriction of human mobility has been the only efficient method to restrain a virus spread. However, total lockdowns have a critical negative impact on economic and social life of community. Thus, various strategies of mobility restrictions are considered depending on a current epidemic situation on different areas, although they are not always well justified due to an insufficient understanding of processes of the infection spread occurring in such complex systems as contact networks of population.

Following ideas reported previously in the literature [1,2], we propose a simple approach to describe the COVID-19 infection spread in an urban area. It is based on a combination of the several extensions of compartment model SEIR [3] describing an infection transmission between individuals in different zones of urban area and the mobility network model with edges weighted according to fluxes of population movement between zones. Each node of the network represents a zone with a corresponding subpopulation. As a case study for our approach we consider Lviv city area, for which the population mobility network is built on the basis of statistical data obtained from local mobile phone operators. In figure one can see the map of Lviv divided by 124 zones, where colours depict a number of population in percent. We calculate the spatiotemporal patterns of COVID-19 spread in each zone and correlations between them in order to clarify which zones play the most important role in spreading infection.



Acknowledgements

This work is supported by the National Research Foundation of Ukraine (grant agreement No 123/01/0338).

References

[1] R. Pastor-Satorras, C. Castellano, P. Van Mieghem, A. Vespignani, "Epidemic processes in complex networks", *Reviews of Modern Physics*, 87 (2015) 925-979.

[2] S. Changruenngam, D.J. Bicout, C. Modchang, "How the individual human mobility spatio-temporally shapes the disease transmission dynamics". *Scientific Reports*, 10 (2020) 1.

[3] J.M. Ilnytskyi, "Modeling of the COVID-19 pandemic in the limit of no acquired immunity", *Mathematical Modeling and Computing*, 8 (2021) 282-303.

Attention dynamics on the Chinese social media Sina Weibo during the COVID-19 pandemic

H. Cui[1]*, J.Kertész*[1]*. (1) Department of Network and Data Science, Central European University, Quellenstrasse 51, A-1100, Vienna, Austria, cui_hao@phd.ceu.edu.*

Understanding attention dynamics on social media during pandemics could help governments minimize the effects. We focus on how COVID-19 has influenced the attention dynamics on the biggest Chinese microblogging website Sina Weibo during the first four months of the pandemic. We study the real-time Hot Search List (HSL), which provides the ranking of the most popular 50 hashtags based on the amount of Sina Weibo searches. We show how the specific events, measures and developments during the epidemic affected the emergence of different kinds of hashtags and the ranking on the HSL. A significant increase of COVID-19 related hashtags started to occur on HSL around January 20, 2020, when the transmission of the disease between humans was announced. Then very rapidly a situation was reached where COVID-related hashtags occupied 30–70% of the HSL, however, with changing content. We give an analysis of how the hashtag topics changed during the investigated time span and conclude that there are three periods separated by February 12 and March 12. In period 1, we see strong topical correlations and clustering of hashtags; in period 2, the correlations are weakened, without clustering pattern; in period 3, we see a potential of clustering while not as strong as in period 1. We further explore the dynamics of HSL by measuring the ranking dynamics and the lifetimes of hashtags on the list. This way we can obtain information about the decay of attention, which is important for decisions about the temporal placement of governmental measures to achieve permanent awareness. Furthermore, our observations indicate abnormally higher rank diversity in the top 15 ranks on HSL due to the COVID-19 related hashtags, revealing the possibility of algorithmic intervention from the platform provider.

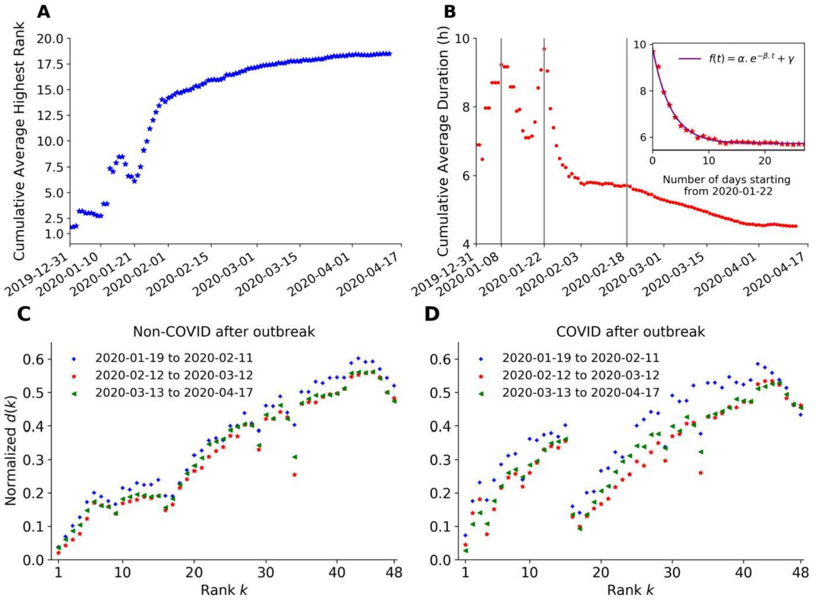


Figure 1: Attention dynamics in different periods of COVID-19. (**A**) Cumulative average highest rank of COVID-hashtags whose first appearance was in the time interval since December 31, 2019. (**B**) Cumulative average duration (hours). The inset shows a three-parameter exponential fit (α=4.13 hα=4.13 h, β=0.31 h/dayβ=0.31 h/day, γ=5.72 hγ=5.72 h) for the cumulative average duration decay after January 22, 2020. (**C**) Rank diversity of the 48 ranks on the HSL, taking all non-COVID hashtags in the three periods after the outbreak. (**D**) Rank diversity taking all COVID-hashtags in the three periods after the outbreak.

Higher Education Responses to COVID-19 in the United States: The Crucial Role of On-Campus Diagnostic Testing for Students and Staff

***B. Klein**[1,2], N. Generous[1,2], S.V. Scarpino[1], and A. Vespignani[1,2]. (1) Network Science Institute, Northeastern University, Boston, USA. (2) Laboratory for the Modeling of Biological & Socio-Technical Systems, Northeastern University, Boston, USA , b.klein@northeastern.edu.*

With a novel dataset of testing and case counts from over 1,000 institutions of higher education (IHEs) in the United States, we analyze the number of SARS-CoV-2 infections and deaths in the surrounding counties during the Fall 2020 semester (August to December 2020). During this period, we see significantly fewer reported cases and deaths in counties with IHEs that perform and report regular student testing, whereas before and after the semester, these counties had almost identical COVID-19 incidence. After controlling for several county-level demographic variables, we perform a matched comparison to isolate the timeline and magnitude of the effects of IHE policy (e.g. in-person vs. remote coursework). By further including data about local mitigation policies in place during the Fall 2020 semester (e.g. stay-at-home, business closures, limits to large gatherings, etc.), we see evidence for a potential underlying mechanism: The benefits of regularly testing students are especially pronounced in counties without local mitigation policies in place, suggesting that campus testing can itself be thought of as a mitigation policy. In sum, asymptomatic testing of students, faculty, and staff at IHEs—coupled with a robust protocol for the isolation/quarantine of individuals with positive tests and their close contacts—should be thought of as another (effective) tool that we can use to combat the spread of COVID-19. Lastly, this dataset and the analyses in this work highlight the benefits of collecting and maintaining large, centralized, and standardized data for responding to this and future pandemics.

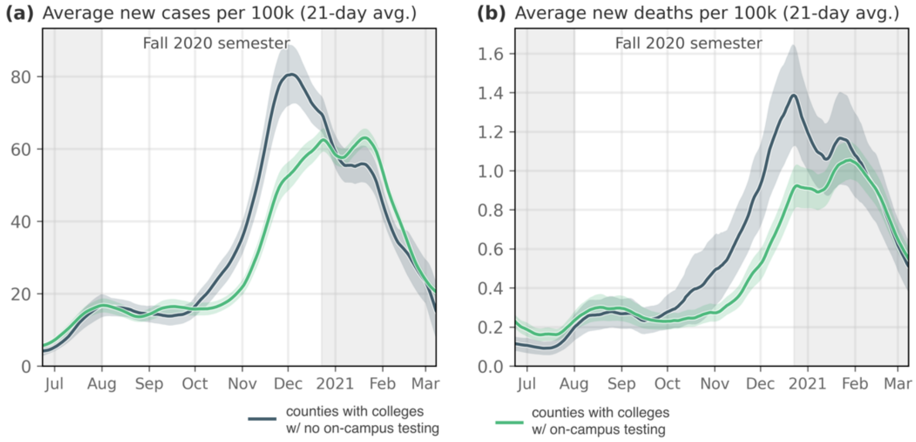


Figure 1: Average new case counts per 100,000 (a) and deaths per 100,000 (b) in counties with institutes of higher education that administer COVID tests on campus (green) compared to counties with institutes of higher education that do not conduct campus testing (black).

Field Study of the Network of Interactions between Pedestrians and Associated Risks of Viral Spread in Non-Confined Crowds

Willy GARCIA $^{(1)}$***, Simon MENDEZ*** $^{(2)}$***, Baptiste FRAY*** $^{(1,3)}$***, Alexandre NICOLAS*** $^{(1)}$

(1) Institut Lumière Matière, CNRS & Université Claude Bernard Lyon 1, Villeurbanne, France,
alexandre.nicolas@univ-lyon1.fr
(2) IMAG, CNRS & Université de Montpellier, Montpellier, France
(3) ENTPE, Vaulx-en-Velin, France

The broad patterns of contacts in a population may not be sufficient to inform detailed epidemiological models, in particular if they aim to be applied to a specific situation. Instead of these contact rates, we show that detailed field-data about pedestrian interactions in given settings (in particular, outdoors in the frame of our study) can be collected empirically. These empirical data are used to build a detailed network of interactions in the crowd and, by combining it with spatially resolved models of viral transmission via respiratory droplets, to infer the risks of new infections raised in each situation [1]. Recently, we have improved these coarse-grained models to better anchor them in Computational Fluid Dynamical (CFD) simulations of the propagation of droplets during breathing (**Figure 1**) and talking, which affords a more accurate study of the spatial dependence of transmission. Finally, these modelling efforts have allowed us to assess the mitigation efficiency of redesigning strategies, for instance switching from two-way foot traffic to one-way foot traffic.

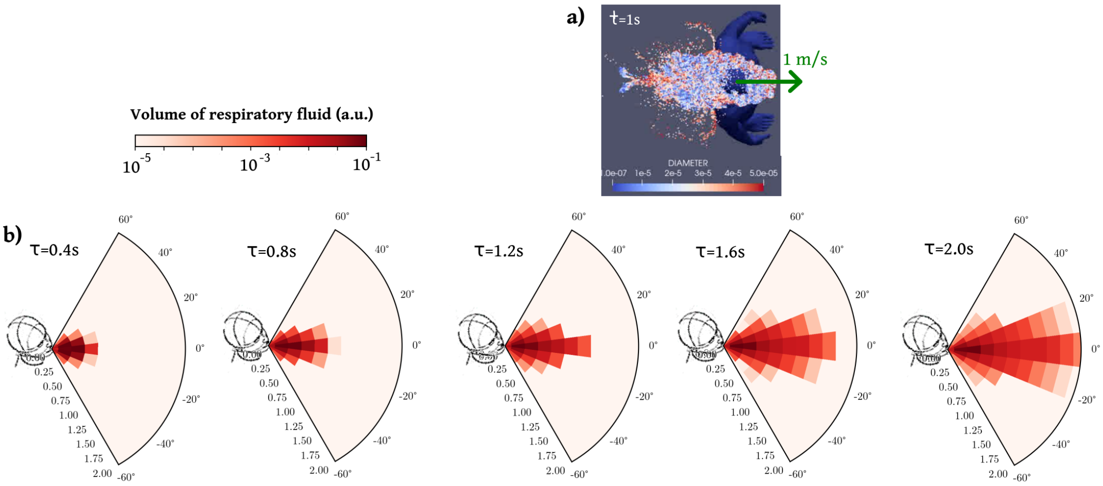


Figure 1: Propagation of respiratory droplets during breathing. (a) Microscopic CFD simulations, (b) Coarse-grained spatio-temporal model derived from these data.

Acknowledgements

We acknowledge funding from Agence Nationale de la Recherche under the programme Recherche-Action COVID-19 and from the GENCI Computing facility.

References (optional)

[1] Willy Garcia, Simon Mendez, Baptiste Fray, Alexandre Nicolas. Assessment of the risks of viral transmission in non-confined crowds. 2020, https://hal.archives-ouvertes.fr/hal-03060324v2/

9 Methodologies

Contents

First-passage times to quantify and compare structural correlations and heterogeneity in complex systems

A.Bassolas[1], *V.Nicosia*[2]. *(1) Universitat Rovira I Virgili, Tarragona, Spain, aleix.bassolas@urv.cat (2) Queen Mary University of London, London E1 4NS, United Kingdom*

We propose here [1] a principled methodology to quantify the presence of correlations and heterogeneity in the distribution of classes or categories in a complex system. The method is based on the statistics of passage times of a uniform random walk on the graph of interconnections among the units of the system. For instance, in the case of a social system we can construct a graph among individuals based on the observed relations or contacts among them. By normalising class mean first passage times with respect to a null-model where classes are reassigned to nodes uniformly at random, we can effectively quantify and compare the heterogeneity of class distributions in systems of different nature, size, and shape. We first test the framework on a variety of systems with simple geometries and ad-hoc class assignments, and then we use it in three real-world scenarios, namely the quantification of polarisation in the Brexit referendum and in the US Congress since 1926, the role of face-to-face interactions among individuals in the spread of an epidemics (See Figure 1), and the relation between economic segregation and prevalence of crime in the 53 US cities with more than one million citizens.

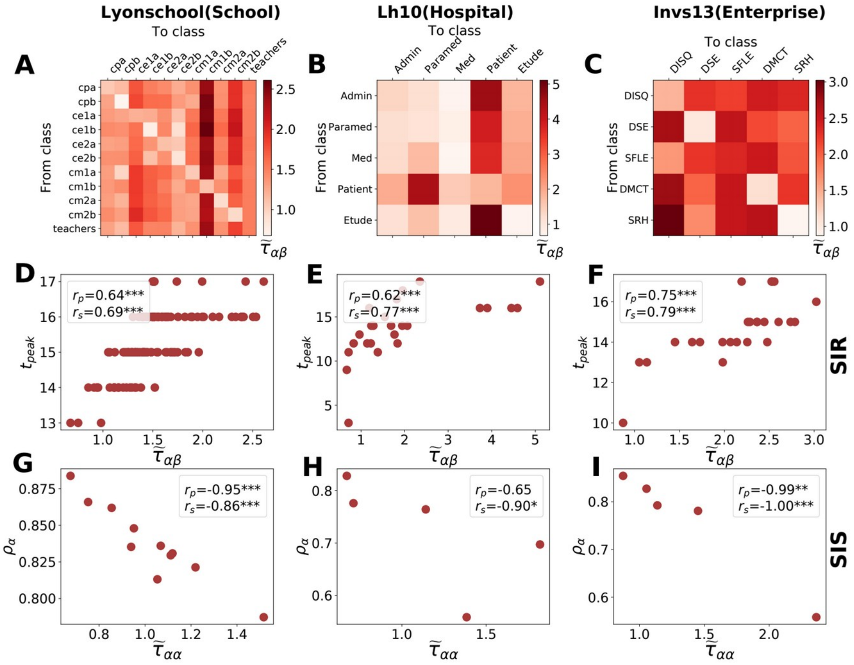


Figure 1: Class mean first passage times and the spread of epidemics in contact networks. (A-C) CMFPT in contact networks corresponding to a school, hospital and an enterprise. (D-F) Relation between $\widetilde{\tau}_{\alpha\beta}$ and the time steps until the peak of the epidemic t_{peak} in a SIR model. (G-I). Relation between the return times $\widetilde{\tau}_{\alpha\alpha}$ and the fraction of infected individuals in each class in the stationary state.

References

[1] Aleix Bassolas and Vincenzo Nicosia. "First-passage times to quantify and compare structural correlations and heterogeneity in complex systems." *Communications Physics* 4, 1, 1-14 (2021).

Mining Stochasticity of Community Detection to Refine Global and Local Organization of Networks

*D. Lee[1], B. J. Kim[1], and H. Kim[2,3], **S. H. Lee**[4,*]. (1) Sungkyunkwan University, Suwon 16419, Korea (2) Korea Institute of Energy Technology, Naju 58322, Korea (3) Universidad del Desarrollo, Santiago 7610658, Chile (4) Gyeongsang National University, Jinju 52725, Korea, lshlj82@gnu.ac.kr. (*presenting author)*

The concept of community detection has long been used as a key device for handling the mesoscale structures in networks. Suitably conducted community detection reveals various embedded informative substructures of network topology. A number of algorithms have been developed to determine the most likely community structures in networks. Such a probabilistic or stochastic nature of this problem can naturally involve the ambiguity in resultant community structures. More specifically, stochastic algorithms can result in different community structures for each realization in principle. Throughout the series of work [1,2], instead of trying to "solve" this community degeneracy problem, we turn the tables by taking the degeneracy as a chance to quantify how strong companionship each node has with other nodes. In addition, regarding the practical usage of community detection, it has always been a tricky problem to assign a reasonable community resolution for networks of interest. Because of the absence of the unanimously accepted criterion, most of the previous studies utilized rather *ad hoc* heuristics to decide the community resolution. We harness the concept of consistency in community structures of networks to provide the overall community resolution landscape of networks, which we eventually take to quantify the reliability of detected communities for a given resolution parameter. More precisely, we exploit the ambiguity in the results of stochastic detection algorithms and suggest a method that denotes the relative validity of community structures regarding their stability of global and local inconsistency measures using multiple detection processes. Applying our framework to synthetic and real networks, we confirm that it effectively displays insightful fundamental aspects of community structures.

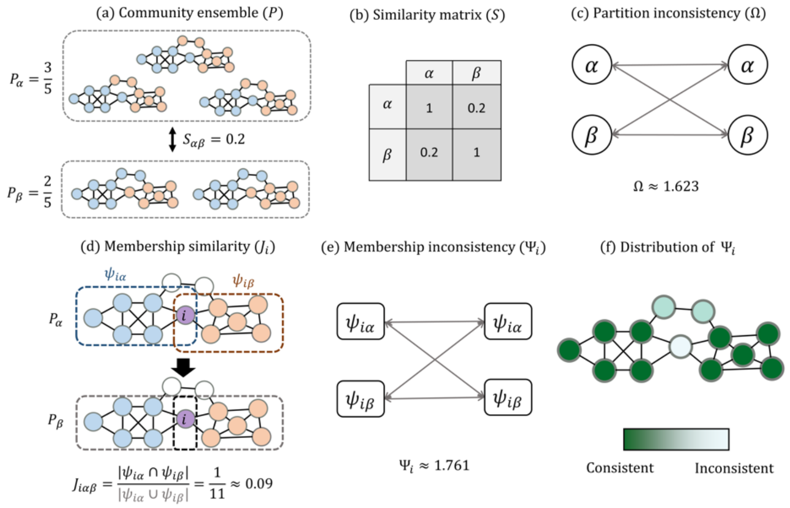


Figure 1: The process of quantifying global and local inconsistency measures of stochastic network community detection

References

[1] H. Kim and S. H. Lee, *Physical Review E*, 100 (2019) 022311.

[2] D. Lee, S. H. Lee, B. J. Kim, and H. Kim, *Physical Review E*, 103 (2021) 052306.

Modelling IR: An Agent-Based Model Approach to The Emergence of International Orders

E. Dacrema[1]*, A. Zotti*[2]*. (1) World Food Programme, Headquarters, Rome, Italy, eugenio.dacrema@gmail.coml.*
(2)Catholic University of the Sacred Heart, Largo Gemelli 1, Milan, Italy

This paper investigates the development of the international order governing inter-state relations based on an entirely innovative take, one applying an agent-based modelling approach, on the assumptions and predictions of three major school of thoughts in the field of international relations (IR) – Realism, Liberalism, and Constructivism (Dunne et al 2013). We model the behavior of international state-actors as characterized by different sets of mixed strategies comprised between total cooperativeness and total mutually-exclusive interest maximization. By simulating hundreds of simultaneous interactions between agents, we analyze which kinds of strategies prevail over time and which tend to disappear to be substituted by fitter ones. Unlike what is most commonly maintained in the field of IR (Zhang 2017, Ikenberry 2011, Young 1992), we argue that a new international order does not emerge from the top-down action by an international leader – e.g., one emerging from a major war or crisis (Gilpin 1991, Modelski 1987) – but as the result of local interactions among agents governed by structural features of the agents – such as propensity to cooperate or to pursue self-interest as well as tolerance towards non-compliance – and of the system – such as the initial distribution of power and the rewards resulting from cooperation or non-cooperation among agents. The evolutionary mechanism embedded in our agent-based model show that the fitness of diverse strategies varies at each stage of the simulations. Although after the start of the simulations in most cases non-cooperative strategies are dominant for long periods of time, a new cooperative order can still emerge as the heterogeneity of the evolutionary mechanism allows for the occurrence of a change of phase determined by the random establishment of new networks of alliances between newly-created cooperative actors.

References

[1] B. Zhang, 'Hans Morgenthau, Realist Theory of International Leadership, and the Future of Global Order', *Chinese Political Science Review*, 2.4 (2017) 512-530.

[2] G. Modelski, *Long Cycles in World Politics*, United States, 1990.

[3] G.J. Ikenberry, *Liberal Leviathan: The Origins, Crisis, and Transformation of the merican World Order*, US, 2011

[4] O. Young, 'The Effectiveness of International Institutions: Hard Cases and Critical Variables' in J.N. Rosenau and E. Czempiel, (eds.), *Governance without Government: Order and Change in World Politics*, United States, 1992, 160-194.

[5] R. Gilpin, *War and Change in World Politics*, United States, 1981

[6] T. Dunne, M. Kurki, and S. Smith, (eds.) *International relations theories*. United Kingdom, 2013.

Generalizing the Radiation Model with Exogenous Drivers

S. Raimondo[1,2], *M. De Domenico*[2].
(1) University of Trento, Department of Mathematics, Via Sommarive 9, Povo (TN), Italy, sraimondo@fbk.eu.
(2) Fondazione Bruno Kessler, CoMuNe Lab, Via Sommarive 18, Povo (TN), Italy, mdedomenico@fbk.eu.

– ***Background*** – Human mobility and migration are complex phenomena emerging from the interplay between social and environmental conditions. The radiation model is a celebrated mathematical model for human migration which is physically based and parameter free [1]. This model does not explicitly consider external drivers, such as socioeconomical and environmental conditions. According to its original formulation the probability of a movement from site *i* to site *j* depends only on their populations, and is given by

$$P(1 \mid m_i, n_j, s_{ij}) = \int_0^{\infty} m_i P(z)^{m_i - 1} \frac{dP(z)}{dz} P(z)^{s_{ij}} [1 - P(z)]^{n_j} dz = \frac{m_i n_j}{(m_i + s_{ij})(m_i + n_j + s_{ij})} \quad (1)$$

where m_i and n_j are the populations at site *i* and *j* and s_{ij} is the total population between *i* and *j*. The drivers of migration are implicitly encoded in the populations and cannot be disentangled. The central term in eq.1 describes the radiation-absorption process, where the random variable z distributed as $P(z)$ describes the "benefit" of the sites. Leaving aside the technical interpretation, the important point is that the distribution $P(z)$ is assumed to remain the same for all sites, regardless of the external conditions. Therefore, through integration, the influence of $P(z)$ disappears.

– ***Model*** – In our study [2], we propose a generalization of this model, by including explicitly the external drivers of migration, allowing for prediction under different socioeconomic and environmental scenarios. In particular, the drivers of migration modify, through an appropriate function, the location and scale parameters of the distribution $P(z)$, so that better external conditions correspond to higher attractiveness of the site, population being equal. The physical process remains the same, but the influence of population is now balanced by the external drivers. The mathematical model becomes:

$$P\left(1 \mid m_i, n_j, s_{ij}\right) =$$
$$\int_0^{\infty} \left[m_i P(o_i|\theta_i)^{m_i - 1} \frac{dP(o_i|\theta_i)}{do_i} \right] \left[n_j P(d_j|\theta_j)^{n_j - 1} \frac{dP(d_j|\theta_j)}{dd_j} \right] P(o_i < d_j) \prod_{k \in L_{ij}} P\left(o_i > d_k\right) do_i dd_j$$

where o_i and d_j are the benefits at origin and destination respectively and L_{ij} are the nodes in-between *i* and *j*. The crucial point is that the benefit distributions are differentiated through the parameter vectors $\theta_i = [\mu_i, \sigma_i]$ which contain the location and scale parameters and varies with the conditions at each site.

– ***Results*** – We present the details of the theory and some application of our "driven" radiation model using a set of synthetic-world sites, with a fractal spatial distribution and a realistic power law distribution of population. We show that our model generalizes the basic radiation so that if the benefit distribution parameters are the same in all sites, the results coincide. Instead, if the benefit of some sites are altered, maintaining the population constant, the migration flows are driven accordingly. Finally, we show the results obtained with real world data of bilateral migration flows.

References

[1] F. Simini *et al.* (2011) *A universal model for mobility and migration patterns*
[2] S. Raimondo, M. De Domenico, (2021) *Generalizing the Radiation Model with exogenous drivers,* (to be submitted).

Microstates estimation from macro data using localised ensemble Kalman filters

B. Kolic*[1] *(1) Institute for New Economic Thinking, University of Oxford, Manor Road Building, Manor Road, Oxford, United Kingdom blas.kolic@maths.ox.ac.uk.

In socio-economic systems, micro-level interactions play an important role in their macro behaviour. Often, we posses sensible interactions rules governing their dynamics; however, observations of the system are sparse and only available in aggregation. In this paper, we show how to estimate the states of a dynamical system represented by a model (eg. agent-based model, initial-value problem, etc) of individual interacting agents. These agents may be embedded in some topology (eg. a network, geographical space, etc), and we assume only macro observations are available. To do so, we apply an ensemble Kalman Filter (EnKF) adapted with model-space covariance localisation to assimilate the states of the agents sequentially. Localisation is a technique in which we exploit the short-range interactions of a system to have more realistic and accurate estimations, and we introduce the network localisation: an extension of the covariance localisation for systems with network topologies. Additionally, we formulate the state estimation problem from a Bayesian perspective and, under several approximations, derive the EnKF equations. This approach is useful for initialising agent-based models because it needs a small number of model simulations and treats the dynamical system as a black box. We demonstrate the effectiveness of this method on a high-dimensional approximation of the chaotic Mackey-Glass system and in the Hegselmann-Krause nonlinear agent-based model for opinion dynamics. In the latter model, we show that covariance localisation greatly improves the accuracy of the EnKF. Furthermore, we show that the EnKF can adapt its microstates estimates to exogenous factors in the observations, such as shocks (see Figure 1).

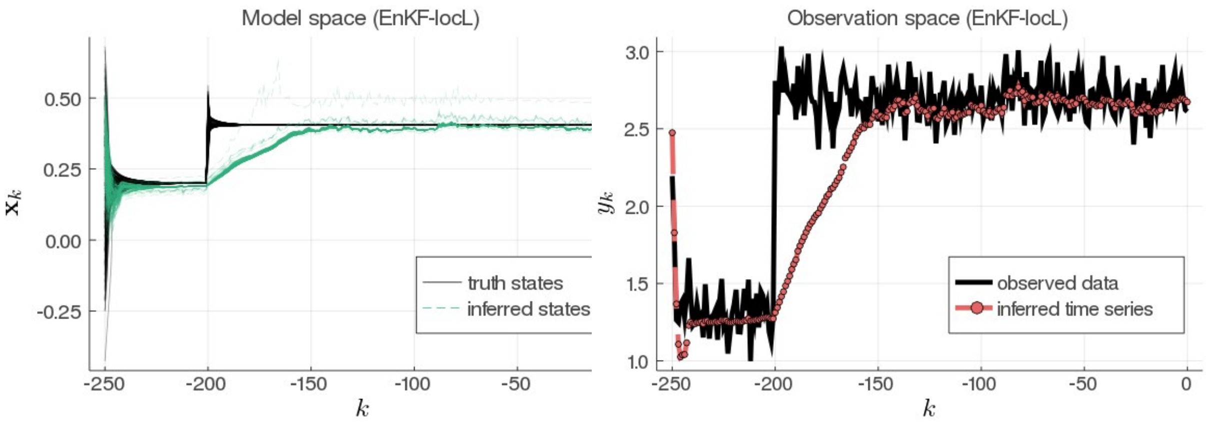


Figure 1: Using an ensemble Kalman filter (EnKF), we infer microstates (left) from aggregate noisy observations (right) of the Hegselmann-Krause opinion dynamics model. Our EnKF incorporates the network topology of the system, significantly improving the accuracy of our estimates. Additionally, we simulate an exogenous shock on the system at time t = 50, and we observe that the EnKF adapt the microstates to the new state of the system accurately.

Acknowledgements (optional)

Acknowledgements to Doyne Farmer, Fabian Aguirre, and Luca Mungo for very insightful comments and discussions.

Scalable Learning of Effective Spreading Models on Networks

Mateusz Wilinski[1] and Andrey Lokhov[1]

[1]Los Alamos National Laboratory, Los Alamos, 87545 New Mexico, United States of America

Spreading processes on networks play an increasingly important role in modeling infectious diseases, regulatory networks, marketing and opinion setting. Events like Coronavirus or Cambridge Analytica affair further highlight the need for prediction, optimization, and control of diffusion dynamics. To tackle these tasks, it is essential to learn the effective spreading model and transmission probabilities across the network of interactions. Unfortunately, in most cases the transmission rates are unknown and need to be inferred from the spreading data. Additionally, it is rarely the case that we have full observation of the dynamics. As a result, typical maximum likelihood approach quickly becomes intractable for large network instances.

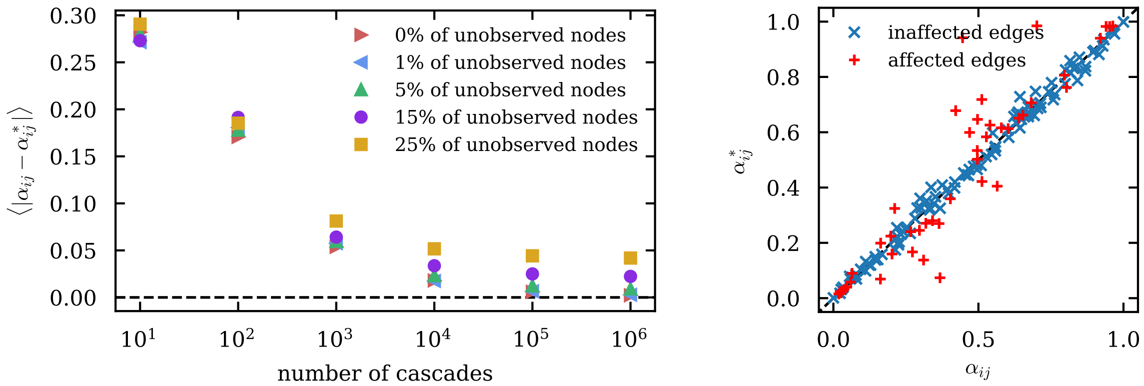


Figure 1: *Left*: Average distance between real and estimated transmission parameters α_{ij}, for a random 3-regular graph with $N = 100$ nodes, as a function of the number of available cascades. *Right:* Example of parameters estimation for a random 3-regular graph with $N = 100$ nodes, $M = 10^4$ cascades and 15% unobserved nodes. Each point represents a different parameter α_{ij}. The red crosses represent edges affected by an unobserved node.

We introduce an efficient algorithm, based on a mean-field approximation and the dynamic message-passing approach, which is able to reconstruct parameters of the effective spreading model given only limited information on the activation times of nodes in the network. The proposed method can be easily generalized to a large class of dynamic models on networks and beyond. Exemplary results for Independent Cascade model on a random regular graph, are shown in figure 1. Number of cascades refer to the number of observed spreading realisations. For each simulation we randomly pick unobserved nodes and transmission probabilities. Note that we analyse a general case, where transmission probability can be different for each edge, but the method can easily be adjusted to simpler settings. Our algorithm is linear in both cascade length (number of spreading steps) and size of the network (number of edges), which makes it scalable and efficient when dealing with huge empirical networks. Additionally, we present a scheme to infer the effective model, which is able to approximate marginal activation probabilities even for settings where message passing is usually inaccurate.

Random graph interpretation of directed graph Laplacians

__Xue Gong__[1,2], Desmond John Higham [1], Konstantinos Zygalakis[1]. (1) School of Mathematics, University of Edinburgh, Edinburgh, UK, s1998345@ed.ac.uk. (2) Maxwell Institute Graduate School, Edinburgh, UK

We consider algorithms that map and reorder the nodes in a directed network to uncover hidden structure. We develop a general framework that allows us to associate optimisation formulations with maximum likelihood problems on random graphs. We focus on two existing spectral approaches, the Magnetic Laplacian algorithm [1] and the Trophic Laplacian algorithm [2], that build and analyse Laplacian-style matrices via the minimisation of frustration and trophic incoherence. These algorithms aim to reveal periodic and linear directed hierarchies, respectively. We show that the two algorithms are associated with new classes of directed random graph models. Using this random graph setting, we are able to compare the two algorithms on a given network. This provides a new computational tool to quantify which structure is more likely. We illustrate the approach on synthetic and real networks, and discuss practical implementation issues.

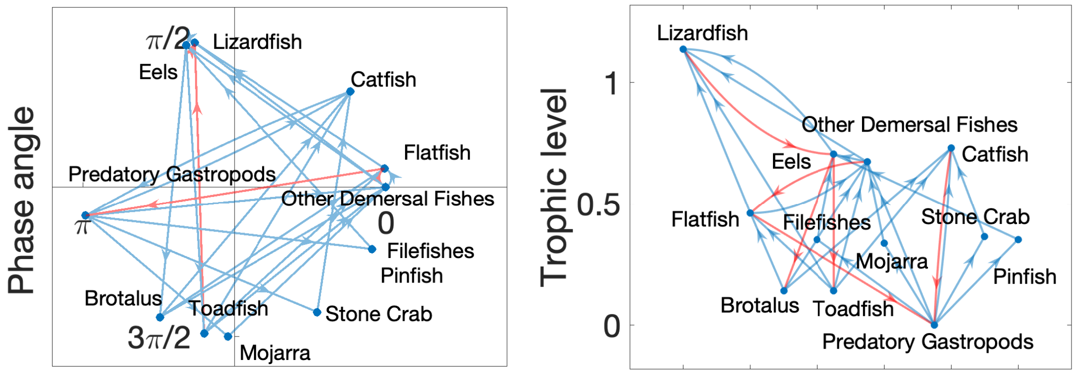


Figure 1. A subgraph of the Florida Bay food web [3] mapped using two spectral algorithms. Left: The Magnetic Laplacian algorithm [1] maps the nodes to phase angles and recovers directed periodic structures. Right: The Trophic Laplacian algorithm [2] maps the nodes to trophic levels and recovers linear hierarchies. Directed edges that respect the desired structure are shown in blue. Here, after calibrating the network to the new random graph models, the periodic version is found to have a higher likelihood.

Acknowledgements

Desmond John Higham was supported by EPSRC Programme Grant EP/P020720/1.

References

[1] Michaël Fanuel, Carlos M Alaíz, and Johan A. K Suykens. Magnetic eigenmaps for community detection in directed networks. *Physical Review. E,* 95(2):022302–022302, 2017.

[2] Robert S Mackay, Samuel Johnson, and Benedict Sansom. How directed is a directed network? *Royal Society Open Science*, 7(9):201138–201138, 2020.

[3] Robert E Ulanowicz and Donald L DeAngelis. Network analysis of trophic dynamics in south Florida ecosystems. *US Geological Survey Program on the South Florida Ecosystem*, 114:45, 2005.

Meta-validation of bipartite network projections

Andrea Zaccaria[1,3], *Alessandro Carra*[2], *Luca Didomenicantonio*[2], *and Giulio Cimini*[4,3].

(1) Institute for Complex Systems (CNR) - UoS Sapienza, 00185 Rome (Italy), andrea.zaccaria@cnr.it ; (2) Physics Department, Sapienza University of Rome, 00185 Rome (Italy); (3) Enrico Fermi Center for Study and Research, 00184 Rome (Italy); (4) Physics Department and INFN, University of Rome Tor Vergata, 00133 Rome (Italy)

Monopartite projections of bipartite networks are key tools to model indirect interactions in complex systems. The standard approach to extract informative patterns from these systems is to statistically validate them using a suitable null network model. A popular choice of null model is the configuration model (CM), built by constraining the degrees of the network and randomizing everything else. However different CM formulations exist, depending on how the degree constraints are imposed and for which nodes of the bipartite network, and the specific choice of the p-value threshold is also arbitrary. Here we systematically compare the application of the various CM formulations in filtering the same networks, showing that they lead to remarkably different results even when the *same* statistical threshold is adopted.

Furthermore, we propose a framework to reconcile these techniques. The starting point is the observation that more similar results are obtained for the same density of statistically significant links. In particular, we show that a common community structure may emerge in a specific range of *density* values. This criterion provides a quantitative approach to build a *meta-validated network projection*: we identify model-specific statistical thresholds for which the different filtering techniques find a maximum agreement and the signal is strongest, and at the same time we obtain results independent from both the way in which the null hypothesis is formulated and the specific p-value threshold. We illustrate this procedure using three unrelated case studies: the network of countries and scientific fields they show a competitive advantage in; the network of mammalian species and the viruses they can be host for; and the Movielens network of users and the movies they rated.

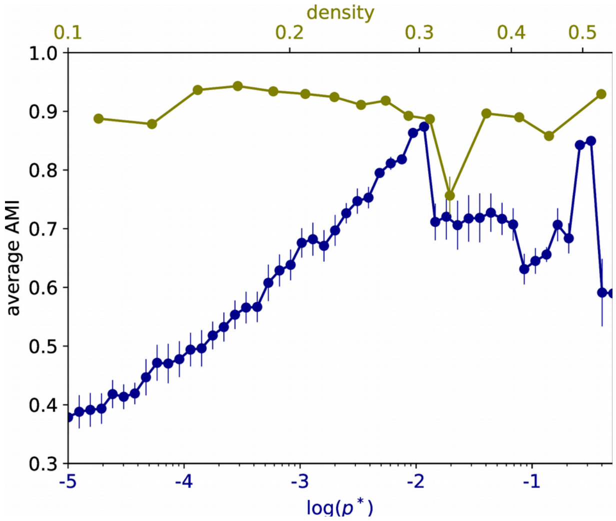


Figure: Average Adjusted Mutual Information (AMI) between the best partitions of the network validated by the various null models. Values are plotted for filtered networks obtained with the same p-value threshold (blue circles) or of equal density (olive circles). The latter option reveals a higher concordance among the null models.

References

[1] Cimini, G., Carra, A., Didomenicantonio, L., and Zaccaria, A. (2021). *Meta-validation of bipartite network projections*. **arXiv**:2105.03391.

Mitigating Limited Observability via Proxy Utilities in Game Theoretic Control

***Samuel D. Johnson**, Dana Warmsley, Charles Martin. HRL Laboratories, 3011 Malibu Canyon Road, Malibu, US, sdjohnson@hrl.com*

In this study, we investigate complex multi-agent optimization scenarios in which autonomous agents are tasked with achieving a system-level objective in the presence of limited communication between agents. Game Theoretic Control (GTC) [1] approaches this coordination problem by: 1) defining agents' local utility functions such that equilibria correspond to local optima of a system objective; and 2) specifying a learning algorithm for agents to use to update their action selections so that equilibrium is guaranteed to obtain. Because an agent's local utility functions generally depend upon both their own action as well as the actions of others, agents must be able to "observe" the actions of each other when evaluating the utility of their own (candidate) actions. However, such observability may not be guaranteed in practice, and we show that equilibrium can fail to obtain when even a small number of agent pairs are unable to observe one another (Figure 1, top).

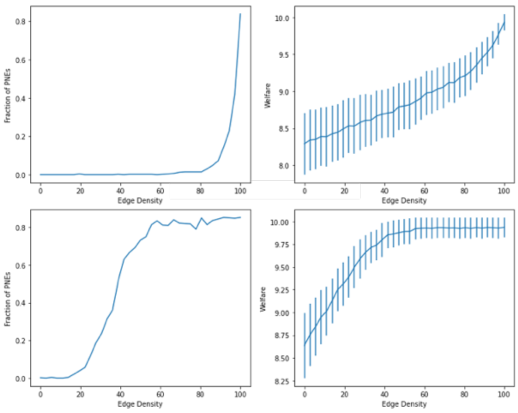


Figure 1. Settings where agents use local utility estimates only (top) and proxy utility computations (bottom). Panels on the left show the likelihood that an equilibrium is reached (y-axis) relative to the density of the inter-agent graph (0 = tree; 100 = complete graph). Panels on the right show the overall system performance.

Recent work to overcome these coordination failures due to the unobservability of agents' chosen actions have focused primarily on approximating agents' utility functions by ones that insert "default" actions in place of those which cannot be observed [2,3]. We extend these local utility approximations by incorporating proxy utility computations from neighboring agents. In our approach, inter-agent observability is represented in a graph where an edge (i, j) indicates that agents i and j can observe one another's actions – i.e., i can observe the actions in its neighborhood N_i. An agent i chooses an action a_i based on a weighted combination of their own local utility calculation $\hat{u}_{ii}(a_{N_i})$ and proxy utility estimates $\hat{u}_{ji}(a_{N_j})$ from its neighboring agents j via the function $\bar{u}_i(a) = (1-\alpha)\,\hat{u}_{ii}(a_{N_i}) + \alpha\, f_{agg}(\hat{u}_{ji}(a_{N_j}), \hat{u}_{ki}(a_{N_k}), \ \ldots)$, where f_{agg} aggregates proxy estimates across neighbors and $0 \leq \alpha \leq 1$ is a constant. By having each neighbor estimate i's utility based on their local information, i is able to indirectly glean information about the actions of agents it cannot observe. We demonstrate our technique via the multi-vehicle target assignment problem (TAP) and observe that proxy estimates can significantly increase the likelihood that equilibrium is reached and system performance maintained even when agents' pairwise observability is significantly diminished (Figure 1, bottom).

References

[1] J. Marden, J. Shamma, *Handbook of Game Theory with Economic Applications*, 4 (2015) 681—699.

[2] P. Brown, J. Marden, *Proc. of 18th European Conference on Control (ECC 2019)*, IEEE, 2019, 3396—3401.

[3] D. Grimsman, J. Seaton, J. Marden, P. Brown, *Proc. of 59th IEEE Confrence on Decision and Control (CDC 2020)*, IEEE, 2020, 1666—1671.

Unlocking heterogeneous node activation in Boolean networks through dynamic programming

G.Torrisi *[1], A.Annibale [2], R. Kühn [2]. (1) King's College London, Strand,London ,United Kingdom , giuseppe.torrisi@kcl.ac.uk. (2) King's College London, Strand,London ,United Kingdom*

We consider the dynamics of node activation in networks of Boolean linear threshold units with fully asymmetric connectivity in the presence of noise, for which the dynamic cavity method provides the most efficient way to evaluate probabilities of dynamic trajectories. However, the complexity of the cavity approach grows exponentially with the in-degrees of nodes, which creates a de-facto barrier for the successful analysis of systems with fat-tailed in-degree distributions. We present a dynamic programming algorithm that overcomes this barrier by reducing the computational complexity in the in-degrees from exponential to quadratic, whenever couplings are chosen randomly from (or can be approximated in terms of) a discrete set of equidistant values. As a case study, we analyse the dynamics of a random Boolean network with a fat-tailed degree distribution and fully asymmetric binary $\pm J$ couplings, and we use the power of the algorithm to unlock the noise dependent heterogeneity of stationary node activation patterns in such a system.

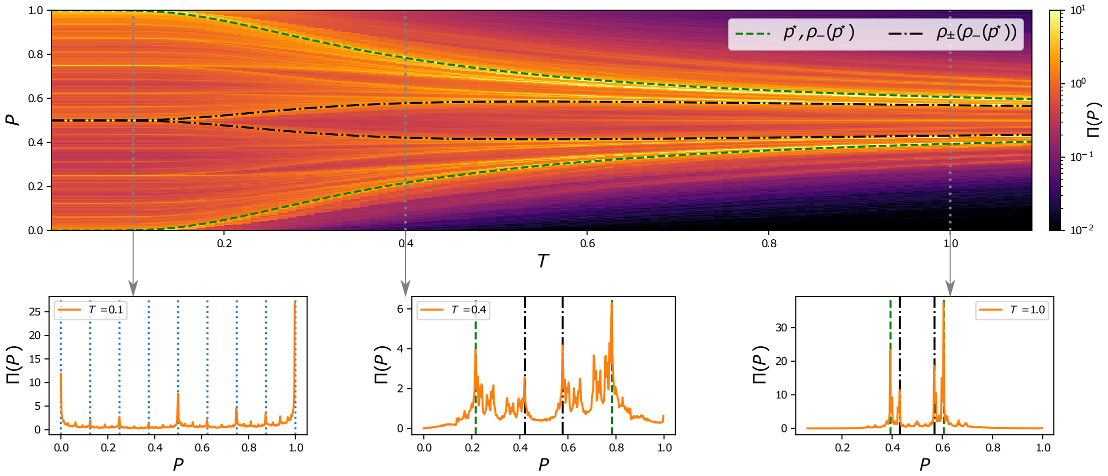


Figure 1: (Top) Heat-map of the frequency density $\Pi(P)$ of heterogeneous node activation probabilities at different values of noise parameter T. The dashed lines show theoretical predictions that can be conceptualized in terms of nodes with degree 1. The vertical dotted lines indicate the values of T at which the histograms of $\Pi(P)$, in the lower panels, are computed.

(Bottom) Line chart of $\Pi(P)$ at different noise levels (low, medium, and high), i.e., $T = 0.1$ (left), $T = 0.4$ (middle), and $T = 1.0$ (right). The vertical dotted lines on the left panel are located at $\left(0, \frac{1}{8}, \frac{1}{4}, \frac{3}{8}, \ldots\ 1\right)$.

Fundamental Biases in influence maximization methods

Vedran Sekara[1,2], Ivan Dotu[3], Esteban Moro[4,5] & Manuel Garcia-Herranz[1]

[1] *IT University of Copenhagen*
[2] *United Nations Children's Fund (UNICEF)*
[3] *Moirai Biodesign*
[4] *MIT Media Lab*
[5] *Universidad Carlos III de Madrid*

Social connections are a conduit through which individuals communicate, information propagates, and diseases spread[1–3]. Identifying individuals within social networks which are more likely to adopt ideas or technologies (or diseases) and spread them to others is essential in order to develop effective information campaigns, fight epidemics, and to maximize the reach of limited resources. Consequently a lot of work has focused on identifying sets of influencers, with a wealth of different heuristics being proposed[4–12]. Inferring sets of influential individuals, however, has previously only been considered a computational problem, with little focus on which demographics these methods reach, and which communities they leave behind. Here we show that seeding information using influence maximization methods, including traditional and state-of-the art methods, only benefit connected and central individuals, consistently leaving the most vulnerable behind. Studying information cascades in synthetic and real-world social networks, we find that up to 59.5% of individuals in a network are disadvantaged with respect to the amount of information they receive, and more than 77.5% of individuals are disadvantaged with respect to how fast they receive information. Our results reveal troublesome outcomes of applying influencer maximization algorithms as they do not disseminate information equally and threaten to create an increasingly unequal society. To bridge the information gap we develop a multi-objective method which selects nodes by optimizing influence maximization, while at the same minimizing information inequality. Our result show that for 5 real world networks (face-to-face network, email network, blog network, Facebook friendship network, collaboration network) not only does our method lower the number of vulnerable nodes, it also improves influence maximization in general. These findings provide a basis for thinking about fairness and effectiveness in influence maximization problems.

References

1. Granovetter, M. S. The strength of weak ties. In *Social networks*, 347–367 (Elsevier, 1977).
2. Pastor-Satorras, R. & Vespignani, A. Epidemic spreading in scale-free networks. *Physical review letters* **86**, 3200 (2001).
3. Rogers, E. M. *Diffusion of innovations* (Simon and Schuster, 2010).
4. Kempe, D., Kleinberg, J. & Tardos, É. Maximizing the spread of influence through a social network. In *Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining*, 137–146 (ACM, 2003).
5. Chen, Y., Paul, G., Havlin, S., Liljeros, F. & Stanley, H. E. Finding a better immunization strategy. *Physical review letters* **101**, 058701 (2008).
6. Banerjee, A., Chandrasekhar, A. G., Duflo, E. & Jackson, M. O. The diffusion of microfinance. *Science* **341**, 1236498 (2013).
7. Morone, F. & Makse, H. A. Influence maximization in complex networks through optimal percolation. *Nature* **524**, 65–68 (2015).
8. Zdeborová, L., Zhang, P. & Zhou, H.-J. Fast and simple decycling and dismantling of networks. *Scientific reports* **6** (2016).
9. Clusella, P., Grassberger, P., Pérez-Reche, F. J. & Politi, A. Immunization and targeted destruction of networks using explosive percolation. *Physical review letters* **117**, 208301 (2016).
10. Braunstein, A., Dall'Asta, L., Semerjian, G. & Zdeborová, L. Network dismantling. *Proceedings of the National Academy of Sciences* 201605083 (2016).
11. Lokhov, A. Y. & Saad, D. Optimal deployment of resources for maximizing impact in spreading processes. *Proceedings of the National Academy of Sciences* 201614694 (2017).

12. Aral, S. & Dhillon, P. S. Social influence maximization under empirical influence models. *Nature Human Behaviour* 1 (2018).
13. Chen, W., Wang, Y. & Yang, S. Efficient influence maximization in social networks. In *Proceedings of the 15th ACM SIGKDD international conference on Knowledge discovery and data mining*, 199–208 (ACM, 2009).

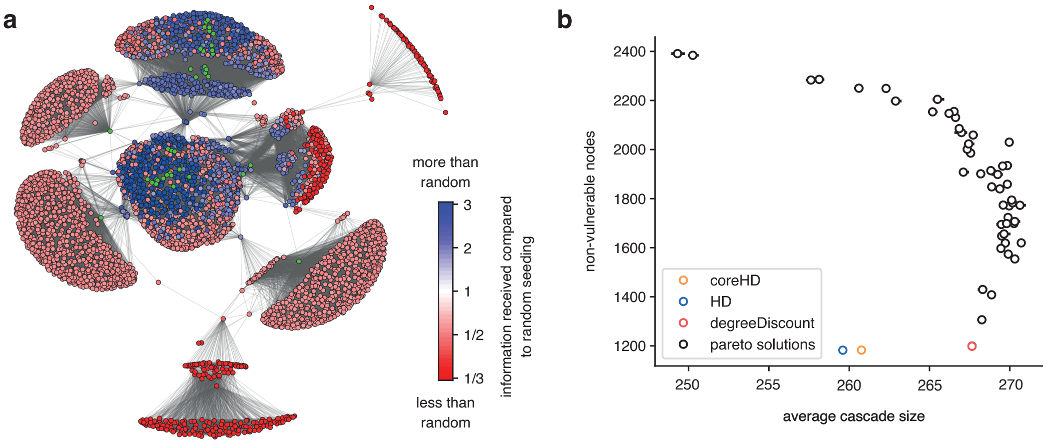


Figure 1: **Inequalities in diffusion of information**. **a**, Numerical simulations for a Facebook friendship network where information has been seeded at 1% of the nodes (colored in green) identified here by using the coreHD heuristic[8]. Information diffusion is simulated using an independent cascade model and we compare how effective and fair the seed sets identified by different influence maximization algorithm are at spreading information. As a reference model we compare the algorithms to a process where influencers are selected at random. The colorbar denotes the fraction of information packets a node receives when information is seeded through a influencer set compared to random. **b**, Solution space for seed sets, illustrating how effective seed sets are at reaching nodes (cascade size) and at being fair in distributing information. CoreHD[8] and degree discount[13] are state-of-the-art methods for identifying influential node sets. They are here compared to node sets selected according to highest degree (HD) and our solutions lying on the pareto front identified by our multi-objective optimization heuristic. Our method identifies multiple better seed sets which are both more effective in distributing information, and more fair.

Spatiotemporal dynamics characterization: brain connectivity similarity profiles

A. Pidnebesna[1,2]*, S. Jiricek*[1,2,3]*, V. Koudelka*[2]*, P. Sanda*[1]*, K. Vlcek*[4]*, J. Hammer*[5]*, J. Hlinka*[1,2]*. (1) Institute of Computer Science, the Czech Academy of Sciences, Pod Vodárenskou věží 271/2, 182 00, Prague, Czech Republic, pidnebesna@cs.cas.cz. (2) National Institute of Mental Health, Topolová 748, 250 67, Klecany, Czech Republic. (3) Czech Technical University in Prague, Technická 1902/2, 166 27, Prague, Czech Republic. (4) Institute of Physiology of the Czech Academy of Sciences, Vídeňská 1083, 142 20, Prague, Czech Republic. (5) Charles University, 2nd School of Medicine, University Hospital Motol, V Úvalu 84, 150 06, Prague, Czech Republic.*

Availability of the data of different modalities or even different nature has rapidly increased during the last years. Every measurement type has its advantages and disadvantages, and often the analysis could benefit from a combination of several data types. However, comparing different properties of the data might be needed. In our study, the visual discrimination task experiment was performed for several subjects, measured by intracranial EEG (iEEG) and by high-density EEG (hdEEG), but not simultaneously [1]. Therefore, for being able to use/compare both modalities, we proposed an approach, allowing us to compare the spatiotemporal dynamics of the data, estimated by several dynamical connectivity measures. For our analysis, we selected several examples of multivariate time series data and a set of connectivity measures. Data included real data (scalp EEG, iEEG, hdEEG with source reconstruction) and simulated time series (synaptic input model, stationary and non-stationary vector autoregressive process, white noise). For connectivity analysis, we selected several functional connectivity measures (Pearson's correlation, partial Pearson's correlation, phase-locking value), and effective connectivity measures, respectively (Granger causality, Direct Transfer Function and Partial Directed Coherence) [2].

For every mentioned data, all connectivity matrices, computed in a sliding window, were put together. Then, the first principal component was computed as the reference point for similarity profile computation. Obtained temporal profile (scores) characterize main temporal dynamics presented in estimated connectivity; spatial profile (coefficients) indicate how representative particular temporal dynamics is for a selected pair of regions. The percentage of the explained variance indicates how informative every component is. Those three objects were used to construct the Connectivity Typicality Profile (CTP) by correlating them with corresponding objects, computed for every connectivity separately. CTP was used to characterize how typical every measure from the set is. For every analyzed data, CTPs were collected together in Connectivity Typicality Fingerprint (CTF) to describe the data by spatial, temporal and informative profiles of connectivity measures. Described CTFs were used for comparing the data measured/simulated with different techniques. The proposed approach allows defining clusters on datasets that have comparable CTF, and whose connectivity measures have similar properties inside the cluster. We showed that hdEEG with source reconstruction are closer to the iEEG by their spatiotemporal dynamics than to any of the modelled data we used, thus it could be a reasonable approximation of intracranial EEG during the experiment.

Acknowledgements

This study was supported by the Czech Science Foundation grant GA19-11753S.

References

[1] Vlcek, K. *et al.* (2020). Mapping the Scene and Object Processing Networks by Intracranial EEG. Frontiers in Human Neuroscience, 14, 1–21.

[2] Dauwels, J. *et al.* (2010). A comparative study of synchrony measures for the early diagnosis of Alzheimer's disease based on EEG. NeuroImage, 49(1), 668–693.

Modelling Time-Varying Interactions in Complex Systems: the Score-Driven Kinetic Ising Model

C. Campajola[1], *D. Di Gangi*[2], *F. Lillo*[2,3], *D. Tantari*[3]. *(1) URPP Social Networks and UZH Blockchain Center, University of Zurich, Andreasstrasse 15, 8050 Zurich, Switzerland, carlo.campajola@uzh.ch. (2) Scuola Normale Superiore, p. dei Cavalieri 7, 56126 Pisa, Italy. (3) University of Bologna, p. di Porta S. Donato 5, 40126 Bologna, Italy*

A common issue when analysing complex systems is that the interactions that characterise them often change in time. This can make it difficult to find optimal models that describe this evolution and can be estimated from data, particularly when the driving dynamics is unknown. In this work [1] we offer a new perspective on the development of models for time-varying interactions introducing a generalisation of the Kinetic Ising Model (KIM), a minimalistic pairwise constant interactions model which has found applications in multiple disciplines. To do so we adopt the Score-Driven methodology, which is designed to keep arbitrary choices of dynamics at a minimum and to achieve information theoretical optimality. In this framework the model's time-varying parameters are deterministic functions of the observations through the model's log-likelihood function: this significantly reduces the computational complexity of the estimation and the exposure to misspecification errors, resulting in less stringent sample selection criteria. The flexibility of our approach allows to tailor the model to a wide range of applications: we show that the model can be employed to forecast stocks volatility at high-frequency, to separate the response to external drivers from endogenous interactions in trading strategies or neuron populations, and to generalise Temporal Exponential Random Graph Models (TERGM) for social networks analysis.

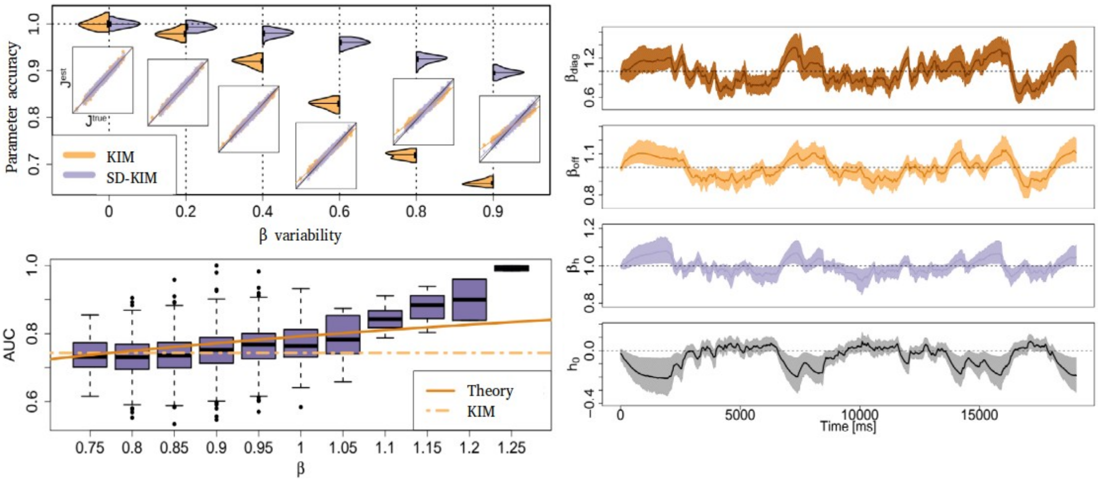


Figure: (top-left) Correction of misspecification error with dynamic inverse temperature β; (bottom-left) High-frequency volatility forecasting accuracy measured with AUC as function of score-driven inverse temperature β, compared to theoretical prediction and stationary KIM; (right) Separation of effects of autocorrelation (β_{diag}), cross-correlation (β_{off}), baseline firing rate (β_h) and input response (h_0) for a population of salamander retinal neurons.

References

[1] **C. Campajola**, D. Di Gangi, F. Lillo and D. Tantari, *Modelling time-varying interactions in complex systems: the Score-Driven Kinetic Ising Model*, preprint arXiv:2007.15545 (2020).

Variational Graph Autoencoders for Multiview Canonical Correlation Analysis

Yacouba Kaloga[1], *Pierre Borgnat*[1], *Sundeep Prabhakar Chepuri*[2], *Patrice Abry*[1], *Amaury Habrard*[3] *(1) Univ Lyon, Ens de Lyon, Univ. Claude Bernard, CNRS, Laboratoire de Physique, Lyon, France, yacouba.kaloga@ens-lyon.fr.*
(2) Department of Electrical and Communication Engineering, Indian Institute of Science, Bangalore, India.
(3) University of Lyon, UJM-Saint-Etienne, CNRS, Laboratoire Hubert Curien, UMR 5516, France.

Digital societies generate large amounts of structured data that frequently stem from observing a common set of objects through different modalities. Such multiview datasets are also encountered in many different fields like computational biology, acoustics, surveillance, or social networks. An example of multiview dataset is provided in Fig. 1. Many methods have been developed in order to perform various machine learning task on these datasets, e.g. clustering or classification. They are based on linear Canonical Correlation Analysis (CCA), which is a linear method which performs a dimensional reduction for two-view datasets. Linear CCA has been extended to deal with dataset with more than 2 views, or to be non-linear, or scalable, or to considers potential graph structure such as in Fig. 1. Yet, as of today, not a single method has all these properties at the same time. In the presents work, we develop a novel approach for multiview CCA, that we call MVGCCA, based on a variational graph neural network. This nonlinear model can consider prior graph-based geometric constraints and is scalable to large-scale datasets with multiple views. It combines the probabilistic interpretation of CCA with autoencoder architecture based on graph convolutional neural network layers. In addition, it is robust to corrupted dataset where some views are missing (for example some pictures of green instances in Figure 1 can be missing). The model can also generate new samples of missing views dataset, or even regenerate a missing view from samples.

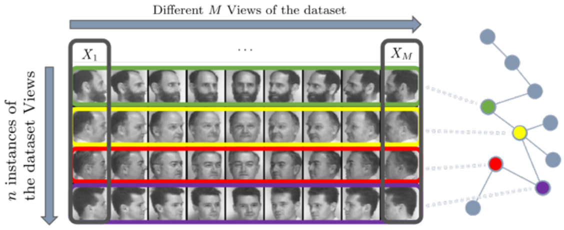


Figure 1: Multiview dataset illustration with potential geometric structure.

We conduct experiments to assess our claim about the performance of the proposed model. One such experiment is about a recommendation task on twitter dataset. This dataset is composed of multiview data where we have 6 features for several twitter users, corresponding to some aspect of their social activities. The goal is to perform a task of friend recommendation. We compare our methods to other ones (see Fig. 2) and we are competitive with the state of the start except on recall metrics. But our methods, thanks to its scalabity, can process a larger version of this data which brings recall metrics performance close to the state of the art.

Dataset	**Recommendation**		
Metric	Prec.	Recall	Mrr
PCA	0.1511	0.0795	0.3450
GPCA	0.1578	0.0831	0.3649
MCCA	0.0815	0.0429	0.2225
GMCCA	**0.2290**	**0.1206**	**0.4471**
MVGCCA	0.1753	0.0583	0.4432

Dataset	**Recommendation Large**		
Metric	Prec.	Recall	Mrr
MVGCCA	0.1745	0.0960	0.4301

Figure 2: Twitter recommendation task: MVGCCA is the only graph-aware, scalable method. It allows us to run the experiments on a dataset 5 times larger than competing methods.
Reference: Extended abstract available on arXiv:2010.16132 (2020).

On the use of Null Models to Assess Statistical Significance in Bipartite Networks

M.J. Palazzi[1], *A. Solé-Ribalta*[1,2], *J. Borge-Holthoefer*[1], *(1) Internet Interdisciplinary Institute (IN3), Universitat Oberta de Catalunya, Barcelona, Spain, mariapalazzi@uoc.edu. (2) University of Zurich, Zurich, Switzerland.*

The use of null models has been a cornerstone to assess the emergence of many network properties at different levels of organization (micro-, meso- or macroscales). Notwithstanding, the debate around which is the most appropriate null model for a given problem is far from being over. Within the ecological community, for example, the discussion around whether nestedness is -or not- ubiquitous in natural systems [1], and under which assumptions, remains open [2]. Yet, efforts have been devoted to exploring to what extent current models are vulnerable to statistical errors or to introduce new models that employ different randomization procedures. Here, we attempt to show that assessing several descriptors under a single model may produce ambiguous results, which difficult the comparison regarding the join emergence of different arrangements within a single network. To this aim, we analyze the statistical significance -in terms of z-scores- for nestedness and modularity with three different null models for a set of ~10k synthetic bipartite networks with varying levels of nestedness [4]. We show that, in some scenarios, this approach can consistently over-or underestimate the presence of significant structural patterns. In light of these ambiguities, we introduce an alternative model (The SBP probabilistic model) that helps to reduce the observed under- and overestimation and highlight the need for the development of new frameworks that take into account those biases.

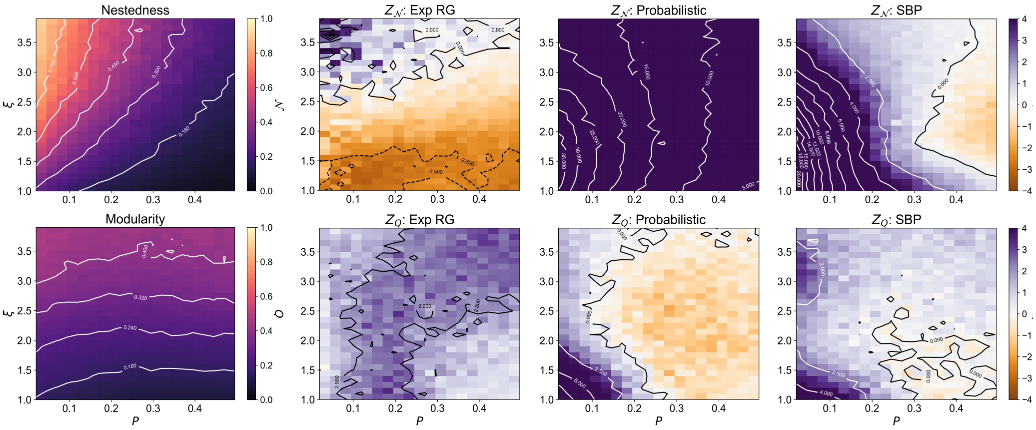


Figure 1: 2-dimensional plots in the ξ - P parameter space showing: Nestedness (top-left) and Modularity (bottom-left) values, z-scores for $\mathcal{N}$ (top-middle) and Q (bottom-middle) under the entropy-based null model [3], z-scores for $\mathcal{N}$ (top-right) and Q (bottom-right) under the probabilistic model [1], and z-scores for $\mathcal{N}$ (top-right) and Q (bottom-right) under the SBP (Solé, Borge, Palazzi) probabilistic model for a set of ~10k synthetic bipartite networks with nested structure. The ξ parameter controls the slimness of the nested structure, and indirectly, the network connectance. The noise parameter P randomly reassigns links outside the perfect nestedness [4].

References

[1] J. Bascompte, P. Jordano, et al., *PNAS,* **100**, 9383 - 9387, (2003).
[2] Molina, C., and L. Stone. *Ecology* **101**(3):e02945 (2020).
[3] C. Payrató-Borràs, L. Hernández, and Y. Moreno *Phys. Rev. X* **9**, 031024
[4] M. J. Palazzi, J. Borge-Holthoefer, et al., *J.R. Soc. Interface* **16**, 20190553 (2019).

Identifying Phase Transitions in Categorical Time-Series Data

B. Ricca[1,2]*. (1) St. John Fisher College, Rochester, NY, USA, bricca@sjfc.edu. (2) Lyda Hill Institute for Human Resilience, Colorado Springs, CO, USA*

Entropy- and information-related approaches to identify phase transitions and construct early warning indicators have been studied for more than 30 years. Social science research has recently used these approaches with categorical variables [1, 2]. However, categorical time-series are less well-studied than their numeric counterparts, and no agreement currently exists regarding the appropriate calculation of windowed entropy or mutual information for such data, resulting in conflicting boundary identification between different techniques. Three issues contribute to the difficulties in using entropy-related approaches for phase boundary identification. First, sliding window entropy is not always appropriate for categorical variable time-series. Second, social science time-series are often relatively short. Third, phase transitions in many social systems include an intermediate regime between two relatively stable phases. (See, for example, Piaget's idea of disequilibration [3].) This study examines these issues and proposes a recursive approach to determining phase boundaries in categorical time series. A comparison of different approaches to phase boundary identification with synthetic and real-world data is presented.

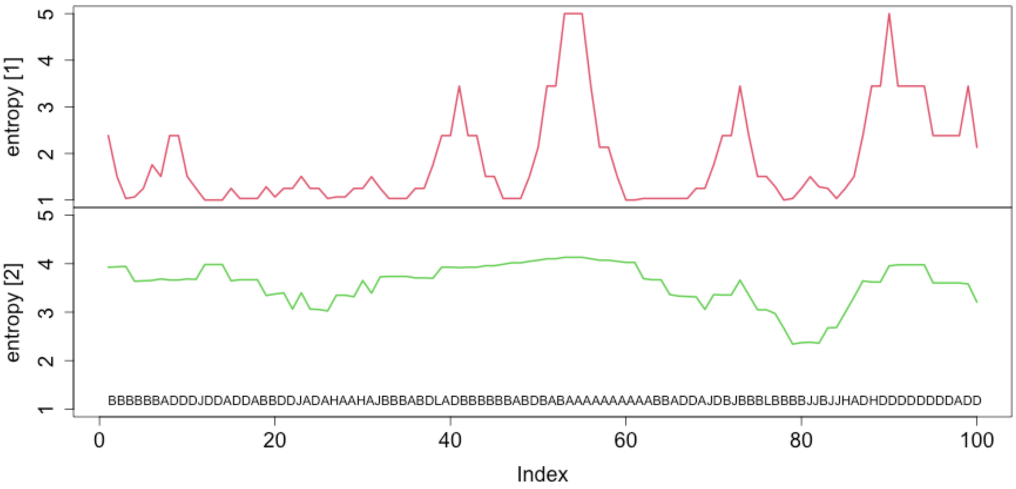


Figure 1: Entropy calculations for a single data set, using the methods from [1] (top graph) and [2] (bottom graph). Categorical time-series data (from [2]) is shown on the lower plot.

References

[1] T. Wiltshire, J. Butner, S. Fiore, *Cognitive Science*, 42 (2017) 129.

[2] B. Ricca, M. Jordan, N. Bowers, *Journal of Experimental Education*, 88 (2020) 431.

[3] J. Gallagher, D. Reid, *The learning theory of Piaget and Inhelder*, 2002.

Programming Memories and Computations in Recurrent Neural Networks Without Training

Jason Z. Kim[1], Zhixin Lu[1], Danielle S. Bassett[1-6]

(1) Department of Bioengineering, University of Pennsylvania, Philadelphia, PA, 19104 USA
(2) Department of Electrical and Systems Engineering, University of Pennsylvania, Philadelphia, PA, 19104 USA
(3) Department of Physics & Astronomy, University of Pennsylvania, Philadelphia, PA, 19104 USA
(4) Department of Neurology, University of Pennsylvania, Philadelphia, PA, 19104 USA
(5) Department of Psychiatry, University of Pennsylvania, Philadelphia, PA, 19104 USA
(6) Santa Fe Institute, Santa Fe, NM, USA

Recurrent neural networks (RNNs) possess unique and powerful computational capabilities due to the internal dynamics of their hidden states. Such dynamics are crucial for cognitive processes that rely on the formation and manipulation of mental representations such as working memory, spatial navigation, and speech. To take advantage of these processes, substantial efforts have been made to train representations in RNNs. However, we still lack a fundamental understanding of the native language of recurrent and distributed representations. We provide such an understanding and language to the extent of programming the RNN connections to store and modify internal memories without training. Specifically, we use tools from dynamical systems and control to track the dynamical variables of inputs as they propagate through the RNN. We then use the distributed representation of variables in the neurons to program the RNN connection weights to store memories and precise modifications to these memories without actually simulating or training the RNN. Equipped with this language of distributed representation, we answer fundamental questions about the learnability of memories and storage capacity of RNNs as a function of the number of neurons and initial connectivity architecture, thereby enabling the principled design of RNNs for specific applications.

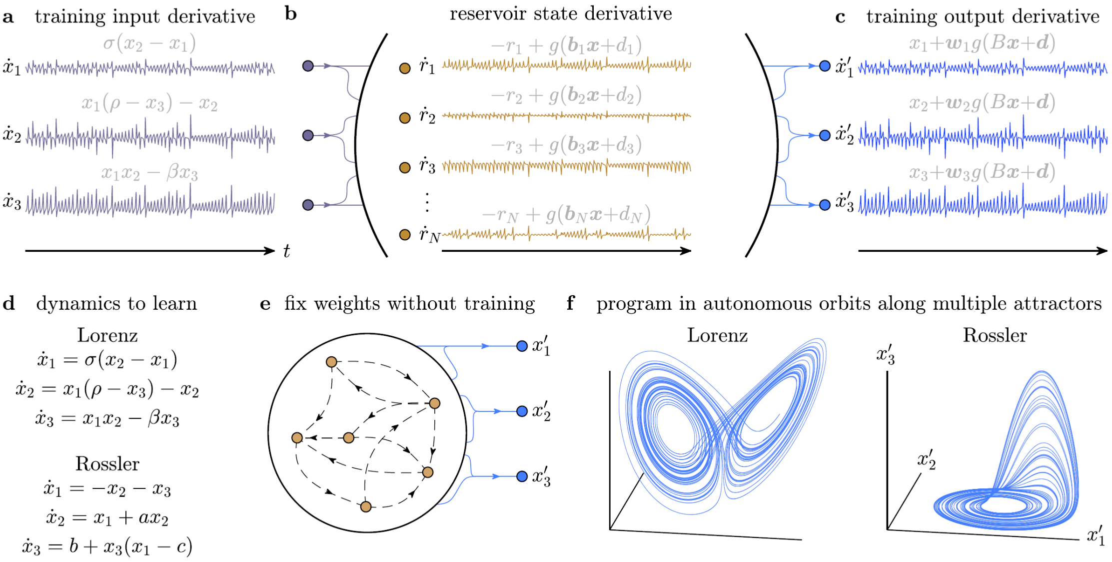


Figure 1: (a-c) A typical attractor learning scheme, where (a) a chaotic input signal is used to drive (b) an RNN, whose states are used to (c) train an output to reproduce the signal. (d-f) Instead of running simulations and training, we can analytically track the (d) dynamical variables through the differential equations of the RNN dynamics, and (e) fix the RNN connectivity to (f) *program* in the desired function. In this case, the weights were programmed to store both the Lorenz and the Rossler attractors as memories in the RNN.

Measuring Randomness Quality Applying Complex Networks and Computational Geometry

***M. Alves**, Bruno. Odemir. (1) USP Alumni, Gent, Flanders, Belgium, alves.m@mail.com. (2) São Carlos Institute of Physics - IFSC/USP, Av. Trab. São Carlense, 400 - Parque Arnold Schimidt, São Carlos - SP, Brazil*

With the aim of mimicking the randomness of natural phenomena, computers are used to generate sequences that look as random as possible running pseudo-random number generators (PRNG). PRNGs have diverse applications in information security, digital games, simulations, modeling, gambling, arts, among others. Despite this fact, existing methods of randomness measure do not offer a definitive solution to the need of classifying PRNGs. The main methods of randomness measurement consist of statistical tests through which the sequences generated by the algorithms are analyzed. Once PRNGs are analyzed by these test suits, they are considered (un)satisfactory random. This work explores an aspect that has been neglected in statistical tests: the spatial distribution of pseudo-random sequences. It is conjectured that this distribution is the source of undisclosed patterns in test suites. One way to study this arrangement in more depth is by using models that explore the relation of values and iterations. The relation of elements in space has to be the basis of the paradigm. This work applies the theory of complex networks and computational geometry methods to find patterns in pseudo-random sequences. The analyzed sequences are plotted in a Cartesian plan generating a set of points that are converted into graphs considering the Euclidean distance between the points within a radius. The best combination of descriptors formed by measurements of graphs and geometric properties is selected. When patterns emerge, one can point out flaws in the widely used methods for measuring PRNGs quality. It is intended to suggest a complementary approach for the evaluation of PRNGs, contributing to a better classification of the PRNGs and, consequently, to cause improvements in the studies on information security. This includes identifying patterns in sequences generated by algorithms that are considered pseudo-random by current statistical tests and thus identifying the limitations of these assessments.

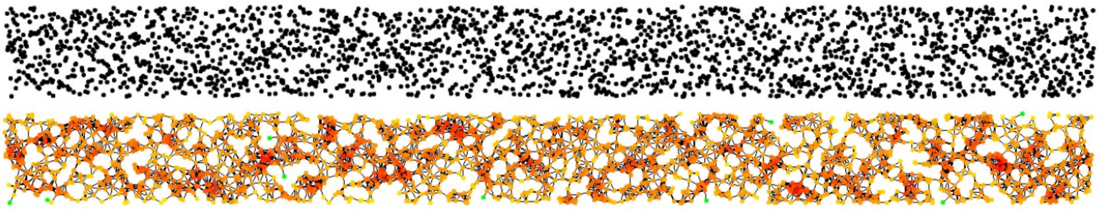


Figure 1: Conversion of a pseud-random plot into a 2D graph to be analyzed by Complex Network and Computational Geometry measurements

References (optional)

[1] Alves, Marcela Lopes. Classification of pseudo-random number generators by complex networks and computational geometry analysis [dissertação]. São Carlos: Universidade de São Paulo, Instituto de Ciências Matemáticas e de Computação; 2019 [citado 2021-05-30]. doi:10.11606/D.55.2020.tde-06012020-171245.

A Spatial Perspective on Burstiness Measure for Spatiotemporal Event Analysis

Eun-Kyeong Kim^{1,2}*. (1) Department of Geography, University of Zurich, Winterthurerstrasse 190, 8057 Zurich, Switzerland, eun-kyeong.kim@uzh.ch. (2) GeoVISTA Center, Department of Geography, Pennsylvania State University, PA 16802, USA.*

Exploring space-time structures is essential to understanding dynamic geographic/environmental events that are localized in space and time. A *bursty* pattern, where extremely long inter-event times alternate with many short inter-event times, is one of the characteristic temporal patterns of events that has been paid attention to in the field of complex systems [1]. The concept and measures of *bursts* have been developed and applied in many theoretical and applied studies [2-3,6]. However, spatial perspectives have been seldom integrated with analyzing the bursty phenomena, although spatial dimension of events including a spatial scale and the environmental context of an event can affect temporal dynamics of phenomena [4]. In measuring patterns of geographic/environmental events, there is the modifiable areal unit problem (MAUP) that a spatial aggregation unit choice may affect measurement outcomes, well-known in the field of geography [5]. Particularly, the shape and scale of spatial aggregation units or spatial sampling methods of events can potentially influence the detected temporal patterns. The same applies to the burstiness measure (e.g., [2,6]); it is crucial to understand how different settings of a spatial dimension of measurements changes the burstiness measurement results even with the same input data. In this study, I explore the impact of various spatial parameters including the shape and size of spatial aggregation units and geographical sampling methods on burstiness measures of Kim & Jo (2016) [2] and Goh & Barabasi (2008) [6], and provide guidelines for constructing a spatial aggregation unit as a spatial container of events and selecting spatial sampling methods for those burstiness indicators. As a proof-of-concept, I apply the proposed guideline to the burstiness analysis of traffic accident events in Switzerland.

References

[1] Barabasi, A.-L. (2005). The origin of bursts and heavy tails in human dynamics. *Nature, 435*(7039), 207-211.

[2] Kim, E.-K., & Jo, H.-H. (2016). Measuring burstiness for finite event sequences. *Physical Review E, 94*(3), 032311.

[3] Karsai, M., Jo, H.-H., & Kaski, K. (2018). *Bursty human dynamics*. Springer International Publishing.

[4] Kim, E.-K. (2018). *Local indicators of temporal burstiness for spatio-temporal event analysis* (doctoral dissertation). The Pennsylvania State University, University Park, PA, USA.

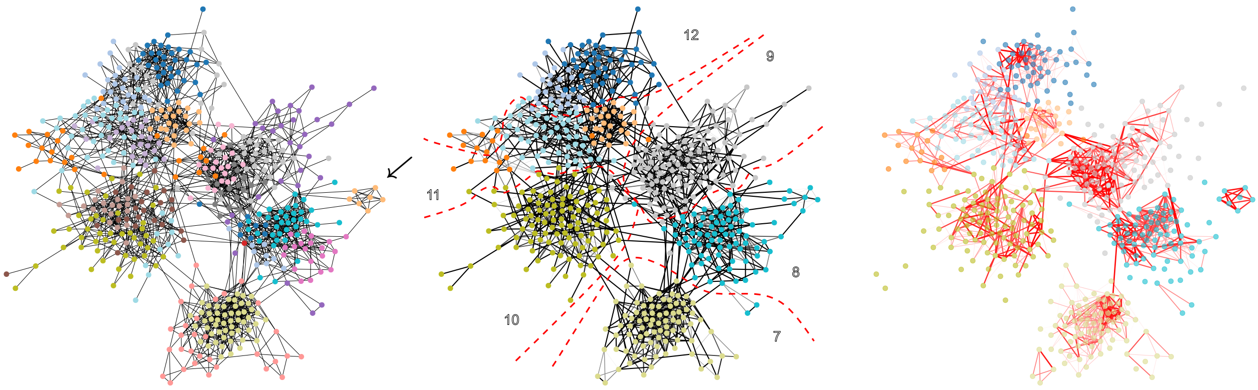
[5] Openshaw, S. (1984) *The modifiable areal unit problem*. Concepts and Techniques in Modern Geography No. 38. Geobooks, Norwich, England.

[6] Goh, K. I., & Barabási, A.-L. (2008). Burstiness and memory in complex systems. *EPL (Europhysics Letters)*, *81*(4), 48002.

Disentangling homophily, community structure and triadic closure in networks

Tiago P. Peixoto
Department of Network and Data Science, Central European University, 1100 Vienna, Austria

Network homophily, the tendency of similar nodes to be connected, and transitivity, the tendency of two nodes being connected if they share a common neighbor, are conflated properties in network analysis, since one mechanism can drive the other. Here we present a generative model and corresponding inference procedure that is capable of distinguishing between both mechanisms. Our approach [1] is based on a variation of the stochastic block model (SBM) with the addition of triadic closure edges, and its inference can identify the most plausible mechanism responsible for the existence of every edge in the network, in addition to the underlying community structure itself. We show how the method can evade the detection of spurious communities caused solely by the formation of triangles in the network, and how it can improve the performance of link prediction when compared to the pure version of the SBM without triadic closure.



Seminal edges Triadic closure edges

(a) SBM ($\Sigma_{\text{SBM}} = 8757.0$ nats) (b) SBM/TC ($\Sigma_{\text{SBM/TC}} = 8456.3$ nats)

Figure 1. Network of friendships between high school students — Adolescent health (comm26) [2]. (a) Fit of the SBM, yielding $B = 26$ communities. (b) Fit of the SBM/TC, uncovering $B = 9$ communities, with seminal (black) and triadic closure (red) edges shown separately in the left and right figures. The thickness of the edges correspond to the marginal probabilities π_{ij} and $1 - \pi_{ij}$ for the seminal and closure edges, respectively. The red dashed lines delineate the known divisions into grades, as numbered.

---

[1] Tiago P. Peixoto, “Disentangling homophily, community structure and triadic closure in networks,” arXiv:2101.02510 [physics, stat] (2021), arXiv: 2101.02510.

[2] James Moody, “Peer influence groups: identifying dense clusters in large networks,” Social Networks **23**, 261–283 (2001).

10 Inter/Transdisciplinarity, applied research and calls for collaboration

Contents

A complex systems perspective on flood risk management

A. Zischg[1]*. (1) University of Bern, Institute of Geography, Oeschger Centre for Climate Change Research, Mobiliar Lab for Natural Risks. Hallerstrasse 12, 3012 Bern, Switzerland,* [andreas.zischg@giub.unibe .ch](mailto:andreas.zischg@giub.unibe.ch)

Flood risk management aims at regulating and controlling risks from natural hazard processes to societal activities and infrastructure. Risk is the probability of a loss, in geosciences often expressed as a function of hazard, exposure, and vulnerability. Based on flood risk analysis and the analysis of the local circumstances, the best combination of measures to reduce risks are found in a decision-making process. Measures for risk reduction can be the construction of river levees or reservoirs for flow regulation as well as land use regulation policies for avoiding the construction of infrastructure in hazard areas or measures to reduce the physical vulnerability of exposed infrastructure. Some measures show effects only in the long-term while others have short-term effects on flood risk reduction.

However, in a constantly changing and interconnected world, flood risk management becomes increasingly complex. Structural flood protection measures have long lifecycles and thus the solutions of problems of the past (e.g., early 19th century river regulations) are still determining nowadays problems and decisions. This path dependency must also be considered for our nowadays solutions for flood risk management and their effects on future problems. Old solutions (lateral levees, dams) are reaching their lifetime and might be expected to fail during the next flood event. Moreover, flood risks become systemic, e.g., the dependence of our society on the functioning of traffic networks and power grids or communication networks is increasing. New risks emerge with socio-technical developments and developments of floodplains. Drivers of flood risk change are interconnected and have feedbacks between each other. Flood risk management is sensitive to context. The implementation of flood risk reduction measures often depends on local conditions of nature protection, water quality acts, hydropower installations, preservation of historical monuments, traffic network, NGO’s and local societal powers or elites. Flood risk management is implemented together by different levels of societal and political organisation: the local community, the regional governmental authorities, the national government, as well as international regulations influence flood risk management. Local awareness varies and therefore patterns emerge between communities that are prioritizing prevention and communities that are preferentially reacting to flood events. Thus, also adaptive behaviour and patterns of self-organization are to be expected. In addition, climate change will increase uncertainties and limits long-term planning by incomplete knowledge of future developments and deep uncertainties.

Complexity theory offers a perspective and a framework for managing flood risk changes in this era of global changes and provides methods for confronting these deep uncertainties. Floodplains can be seen as coupled human and natural systems and as complex adaptive systems. Therefore, flood risk management practice can be adapted in future by incorporating elements from complex system sciences. Inter-, trans-, or post-disciplinarity play a key role for the implementation into practice.

Promoting Complexity in Scientific and Interdisciplinary Group Discussion: The development of the 'Relatoscope' and 'Observatron' methods

Ana Teixeira de Melo[1], *(1) Centre for Social Studies, University of Coimbra, Portugal, E-mail: anamelopsi@gmail.com;*

In many areas of science, group discussions and debates are a fundamental activity. Increasingly, Science embraces Interdisciplinarity, where knowledge production could be framed in terms of the degree of integration and emergence of novelty in theoretical, methodological or pragmatic/problem-solving terms. Scientific meetings often take place where a variety of perspectives are presented but with limited conditions for the type of deep, non-linear interactions from which something truly new may emerge (qualitatively different and not reducible to the sum of the parts). The degree of interpenetration of ideas and the potential of such discussions and debates to promote true scientific breakthroughs is oftentimes constrained by insufficient attention to the intentional facilitation of such interactions, namely towards increasing the complexity (differentiation, integration, recursivity, emergence) of their outcomes. A Complex Thinking perspective could be applied to these situations, informing the intentional management and facilitation of the interactions in a group towards positive emergent outcomes in terms of the novelty and relevance of the knowledge produced and the fitness with the system(s) of interest. Scientific groups, as complex systems, can be understood and managed with a focus on the internal group dynamics and their constraints, both in relation to interindividual emotional exchanges and in terms of the interaction of ideas, concepts and methods deriving from the individual's disciplines and their cognitive constructions. From a complex systems perspective, the nature of their relations between the individuals' contributions is as important as their nature and of the constraints posed by their own emergent products and contexts. A recently proposed framework for the operationalisation of the notion of Complex Thinking [1] postulates that the enactment, at the level of the thinking, of a set of properties such as those organising complex systems may, in given circumstances, lead to the emergence of new ideas or hypotheses as well as richer group experiences. Applied to the dynamics of group discussions such collective enactment could support higher levels of integration and the emergence of sufficient novelty to build new and more complex forms of knowledge. Building from an initial proposal of a relational thinking method we have been developing a suite of strategies to promote the complexity of the thinking with applications to Scientific and Interdisciplinary Group Discussions. In this paper, we discuss the development, applications and challenges to the further refinement and evaluation of the 'Relatoscope' and 'Observatron' methods in promoting dimensions such as the structural and dynamic complexity of the thinking, dialogic and explanatory complexity, and the complexity of the observer in supporting creative, second-order or emergent complex thinking in science.

Acknowledgements

Ana Teixeira de Melo is supported by the Portuguese Foundation for Science and Technology through the Transition Norm (DL57/2016/CP1341/CT0011)

References

[1] Melo, A. T. (2020). *Performing Complexity: Building Foundations for the Practice of Complex Thinking*. Springer Nature.

Combining Viral Genetic And Animal Mobility Network Data To Unravel *Peste Des Petits Ruminants* Transmission Dynamics In West Africa

***Andrea Apolloni*[1,2]** *,Arnaud Bataille[1,2], Habib Salami[1,2,3†], Ismaïla Seck[4,5†] , Modou Moustapha Lo[3], Aminata Ba[3], Mariame Diop[3], Baba Sall[4] , Coumba Faye[4], Mbargou Lo[4], Lanceï Kaba[6], Youssouf Sidime[6], Mohamed Keyra[6], Alpha Oumar Sily Diallo[6], Mamadou Niang[7], Cheick Abou Kounta Sidibe[7], Amadou Sery[7], Martin Dakouo[7], Ahmed Bezeid El Mamy[8], Ahmed Salem El Arbi[8], Yahya Barry[8], Ekaterina Isselmou[8], Habiboullah Habiboullah[8], Abdellahi Salem Lella[8], Baba Doumbia[8], Mohamed Gueya[8], Caroline Coste[1,2], Cécile Squarzoni Diaw[1,9], Olivier Kwiatek[1,2], Geneviève Libeau[1,2]*

(1) ASTRE, Univ Montpellier, CIRAD, INRAE, F-34398, Montpellier, France, andrea.apolloni@cirad.fr (2) CIRAD, UMR ASTRE, F-34398 Montpellier, France (3) Institut sénégalais de Recherches agricoles, Laboratoire national de l'Elevage et de Recherches vétérinaires (LNERV), BP 2057 Dakar-Hann, Sénégal.(4)Direction des Services vétérinaires, Dakar, Senegal (5) FAO, ECTAD Regional Office for Africa, 2 Gamel Abdul Nasser Road, PO Box GP 1628 Accra, Ghana (6)Institut supérieur des Sciences et Médecine Vétérinaires, BP 2201 Dalaba, Guinea (7) Laboratoire central vétérinaire (LCV), BP 2295, Bamako, Mali.(8) Office national de Recherches et de Développement de l'Elevage (ONARDEL), BP 167 Nouakchott, Mauritania.(9) CIRAD, UMR ASTRE, F-97491 Ste-Clotilde, La Réunion, France

Peste des Petits Ruminants (PPR) is a deadly viral disease that mainly affects small domestic ruminants. This disease threatens global food security and rural economy but its control is complicated notably because of extensive, poorly monitored animal movements in infected regions.

We combined the largest PPR virus genetic and animal mobility network data ever collected in a single region to improve our understanding of PPR endemic transmission dynamics in West African countries.

Phylogenetic analyses identified the presence of multiple PPRV genetic clades that may be considered as part of different transmission networks evolving in parallel in West Africa. A strong correlation was found between virus genetic distance and network-related distances. Viruses sampled within the same mobility communities are significantly more likely to belong to the same genetic clade.

These results provide evidence for the importance of animal mobility in PPR transmission in the region. Some nodes of the network were associated with PPRV sequences belonging to different clades, representing potential "hotspots" for PPR circulation.

Our results suggest that combining genetic and mobility network data could help identifying sites that are key for virus entrance and spread in specific areas. Such information could enhance our capacity to develop locally adapted control and surveillance strategies, using among other risk factors, information on animal mobility.

Outreach in Complexity Studies for transforming ‘everyday’ modes of thinking: The Complex Thinking Academy

Ana Teixeira de Melo[1], *Leo Simon Dominic Caves*[2] *(1) Centre for Social Studies, University of Coimbra, Portugal, E-mail: anamelopsi@gmail.com; (2) Independent researcher and consultant, S. Félix da Marinha, Portugal; Associate of the York Cross-Disciplinary Centre for Systems Analysis (YCCSA) of the University of York, United Kingdom and collaborator of the Centre of Philosophy of Sciences of the University of Lisbon, Portugal*

Complexity Studies have paved the way for the development of new modes of thinking and doing in Science, with significant implications in terms of applied interventions with real-world complex systems. The history of Science is as much influenced by the evolution of human cultures and their intrinsic philosophies and modes of thinking as it has contributed to shaping them. Complexity Studies, including the Science(s) and the Philosophy(ies) of Complexity have an important role to play not only in shaping new modes of scientific exploration and for managing the complexity of the world but also in shaping ‘everyday’ modes of thinking and our ways of being and thinking together, as a part of that complexity. There is scope for the development of educational interventions aiming to promote deeper transformations in the worldviews that frame the modes of thinking and practices in our ‘everyday; and ‘real-world’ lives, affecting our modes of being, living and acting together in our local communities. In this paper, we present the work developed under the umbrella of the Complex Thinking Academy (CTA) operating through workshops aiming to disseminate and communicate key concepts of a paradigm of complexity and of the practice of Complex Thinking. Pilot workshops have been run, from 2018, in the UK, and in Portugal, with groups with different characteristics including the general population, interdisciplinary group of academics and secondary school students. They were designed to be isomorphic, in process terms, with the principles they aim to highlight, affording an experience *in* Complexity and practice of some key dimensions organising the Complexity of the Thinking. The sessions rely on an engaging and experiential methodology inviting participants to enter an imaginary role playing scenario set-up in a space colony- ComplexCity- in the year of 2118. Participants are invited to reflect on the problems facing ComplexCity, as well as on the nature of the descriptions, explanations and solutions proposed to deal with them. In a ‘training’ section, they are introduced to some key complexity notions and then invited to experiment with using tools designed to support the practice of some key structural dimensions of Complex Thinking as well as of the complexity of the observer. These pilot experiences and their preliminary evaluations hint at the potential of the CTA to support transformational learning, impacting on the participant’s worldviews and, potentially, on their way of being and acting in relation to others, the world and its challenges. In this paper, we synthesize the methods used and the core results of the qualitative evaluations of previous CTA. We discuss the potential of the methodologies used and the prospects for expanding the CTA into intervention programs in Complexity, aiming to explore and amplify its potential to support meaningful transformations on individuals and their communities. We discuss different ways that the CTA can develop, e.g. formats, contexts and the challenges ahead in terms of their implementation and evaluation. We open a call for collaborations.

Acknowledgements

Ana Teixeira de Melo is supported by the Portuguese Foundation for Science and Technology through the Transition Norm (DL57/2016/CP1341/CT0011)

The Need for a Paradigm Shift in Sustainable and Healthy Nutrition Research

A. Aleta[1]*, on behalf of the SNSB. (1) ISI Foundation, Turin, Italy, albertoaleta@gmail.com*

The research in the field of Sustainable and Healthy Nutrition is calling for the application of the latest advances in seemingly unrelated domains such as complex systems and network sciences on the one hand, and big data and artificial intelligence on the other. The confluence of these fields, whose methodologies have experienced explosive growth in the last few years, promises to solve some of the more challenging problems in sustainable and healthy nutrition, i.e., integrating food- and behavioral-based dietary guidelines. Focusing on nutrition and health, we argue what kind of methodological shift is needed to open current disciplinary borders to the methods, languages and knowledge of the digital era and a systems thinking approach. Specifically, we advocate for the adoption of interdisciplinary, complex-systems-based research to tackle the huge challenge of dealing with an evolving interdependent system in which there are multiple scales – from the metabolome to the population level -, heterogeneous – and more often that not – incomplete data, and population changes subject to many behavioral and environmental pressures. To illustrate the importance of this methodological innovation, we explore state-of-the-art food composition databases and show how unreliable they can be. Finally, we address the problem of finding healthy and sustainable diets while tacking into account the cultural aspects of nutrition.

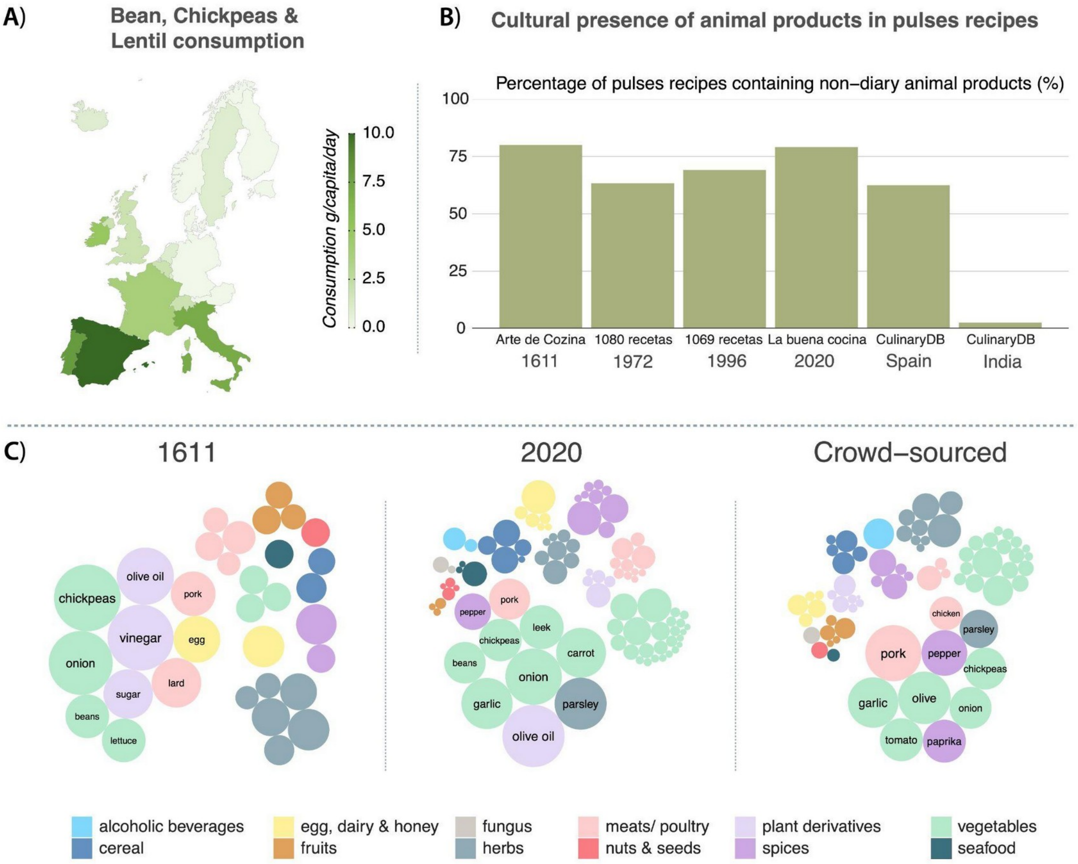


Figure 1: A complex systems view of food composition. Panel A shows a comparison of the intake of legumes in Western and Northern Europe, which reveals the role of sociocultural factors. Looking at the ingredients that are commonly paired with these legumes in Spanish cuisine, one finds that this consumption is heavily related to the consumption of non-dairy animal products, although this is not true for the Indian cuisine sample (B).

More Complexity and Less Difficulty in Music Assisted Language Learning

X. Chen, *Université Bourgogne Franche-Comté, Faculté des sciences Mirande, 9, Av. A. Savary, 21000 Dijon, France E-mail: Xiao.chen-perdereau@u-bourgogne.fr*

We will present some aspects of our innovative music assisted language learning (MALL) method. Why introduce music which makes the language learning even more complex? On the practical sides, more complexity implies more interactions interconnecting the involved cognitive systems. We do not however consider that more complexity leads necessarily to more learning difficulty. In a previous work we studied the additional complexity in the framework of social cognition, and found that it allowed significantly improved speech proficiency [1,2]. As well, the added music-induced complexity was found to help beginners in Chinese Mandarin. Indeed we need to understand how the cognitive systems interact with each other to take advantage of this complexity. Although quantitative studies in the research labs attract more and more attention, practical, more qualitative complementary approaches are still necessary to provide research insights, materials, view angles and verification.

This work focuses in the oral production of comprehensive phrases. The difficulties manifested by the learners are mainly tonal pronunciation, phonetic memory and syllables grouping. Learning materials were extracted from [3], one of the first dedicated MALL method in which the music and language coherence was inspired from concepts developed in the anterior theoretical works on music and language structures [4]. The learning procedure and the observations are described in [5]. Music and speech recordings will be available during the poster presentation to show specific conditions and results.

Three stages have been identified in which music offers benefits on reducing learning difficulties: 1. in the perception stage through collective singing trained by familiarized rhythms and melodies [6] and quick individual phonetic self-correction. 2. in the consolidation stage through rewarded repetition, the pronunciation and the phrases fluency were improved consequently. This result support the recent studies on neuroscience confirming that music can act as a reward in the learning process [7]. 3. in the speech sequences reproduction stage it provides an additional cue for memory retrieval before reproduction. This last role played by music is interesting for further investigation as its action is not limited to language learning. The role of music in souvenir retrieval has indeed been demonstrated in the field of health science [8].

References

[1] X. Chen (Perdereau), Proceedings of the European Conference on Complex Systems, (2012), Springer, 671

[2] X. Chen (Perdereau), 15th International Conference on Cognitive Neural Systems, Boston, USA (2011).

[3] C. Matin Rouge, “Soirée Zou”, Edition Yin et Langues, Dijon (2009)

[4] F. Lerdahl, and R. Jackendorf, “A Generative theory of tonal music”, MA, USA, MIT Press (1983)

[5] X. Chen, “Music assisted language learning”, 2021, https://hal.archieves-ouvertes.fr/hal-03207582

[6] X. Chen (Perdereau), “Speech rhythm processing”, invited talk, International Conference on Acoustics, Pittsburgh, USA, 2015. The abstract in JASA 137, 2272 (2015).

[7] B. P. Gold, E. Mas-Herrero, Y. Zeighami, M. Benovoy, A. Dagher and R. J. Zatorre, PNAS, (2019) 116, 8, 3310-3315

[8] A. Moussard, E. Bigand, S. Belleville, I. Peretz, doi: 10.1080/09602011.2014.917982, Epub 2014 Jun 2.

Rhythm And Form In Music, A Complex Systems Approach.

Kolic B.[1], *Rodríguez Cervantes M. T.*[2], *Padilla Longoria P.*[2,3], *Knights F.*[3] *(1) University of Oxford, Oxford, United Kingdom, blas.kolic@maths.ox.ac.uk , (2) UNAM, Ciudad de México, México, (3) University of Cambridge, Cambridge, United Kingdom.*

Form, in a general sense, consists of shape and structure. In music, shape results from changes in attributes of sound in time, while structure is the relations among sounds and sound-configurations [1]. How these two are generated is always related to rhythm. Based on an information and a network-theoretical approach, we develop three measures of the rhythmic complexity of a musical piece and its parts: 1) the *rhythmic heterogeneity*, based on the permutation entropy [2], where high values indicate a high variety of rhythmic patterns; 2) the *disequilibrium*, based on the distribution of on-beat onsets, where high values indicate high syncopation; and 3) the *component extractor*, based on the communities of a visibility graph of the rhythmic figures over time [3], where we identify structural components that constitute the piece. Under the same parameters, our measures are comparable within a piece or between pieces. In Fig. 1, we show the rhythmic complexity of two contrasting examples: a minuet from Haydn string quartet and the Super Mario Bros theme. Within the pieces, we observe patterns that accurately correspond to their musical sructural divisions. Between the pieces, we find that the Super Mario theme is significantly more complex than the Haydn minuet, as the formal analysis of the works would already suggest.

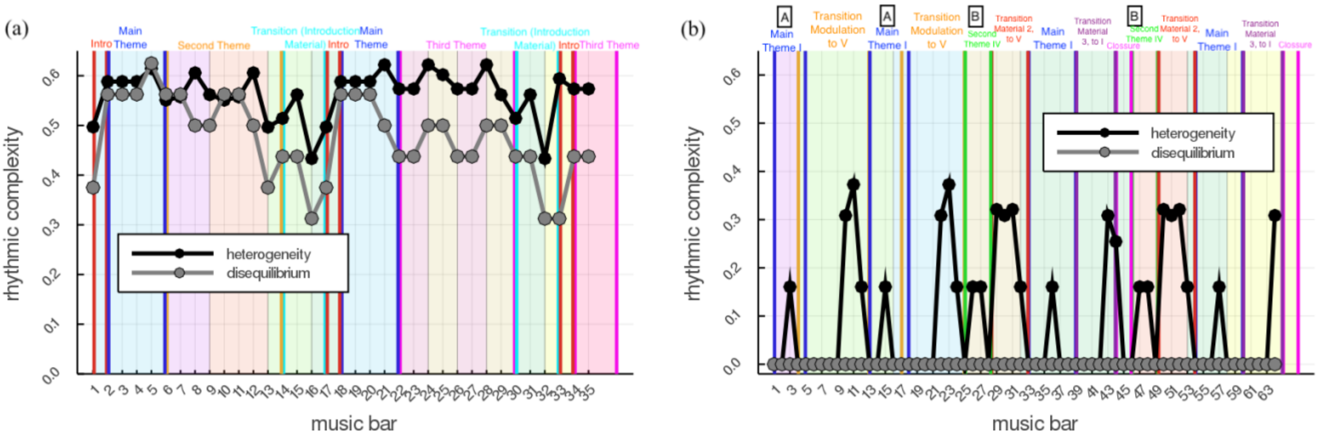


Figure 1: **Rhythmic complexity measures (see legend) for (a) the Super Mario theme and (b) the Haydn Minuet from String Quartet Op.17 No.2**. The complexity values at bar *i* consider the rhythmic patterns within bars *i* and *i+1*. Solid vertical lines indicate structural divisions of the piece, as indicated by their color. The background colouring indicates the community structure obtained with the *component extractor*. The total *heterogeneity* (*disequilibrium*) values are 0.78 (0.48) and 0.56 (0) –out of 1 (1)– for the Super Mario theme and the Haydn minuet respectively.

References

[1] Tenney, James. *From Scratch: Writings in Music Theory*. (p. 150) University of Illinois Press, 2014, pp. 150-165.

[2] A Tlaie, I Leyva, R Sevilla-Escoboza, VP Vera-Avila, and I Sendiña-Nadal. Dynamical complexity as a proxy for the network degree distribution. Physical Review E, 2019.

[3] Lucas Lacasa, Bartolo Luque, Fernando Ballesteros, Jordi Luque, and Juan Carlos Nuno. From time series to complex networks: The visibility graph. Proceedings of the National Academy of Sciences, 105(13):4972–4975, 2008.

A Proposal for Measuring and Visualizing Complexity of Musical Information Using Networks

A. Alcalá[1], *P. Padilla*[2]. *(1) Faculty of Sciences, UNAM, Mexico,* albertoalcala@ciencias.unam.mx.
(2) IIMAS, UNAM, Mexico.

We present a framework and method for constructing a graph or network from essential musical information for structural analysis. These graphs englobe statistical information about musical elements, such as notes, chords, rhythms, intervals, etc., and the relations among them, and so become helpful in visualizing and understanding features of a musical fragment. To obtain these networks we parse musical data out of a digital symbolic music file. This data undergoes different analytical procedures from Graph Theory, such as measuring centrality of nodes, community detection and entropy calculation. The resulting information reflects important structural characteristics of the fragment in question: predominant elements, connectivity between them, and complexity of the information contained in it. We analyze pieces of music coming from different styles, and compare some of our results with traditional music analysis to show the consistence, and thus the potential utility, of this method for music analysis.

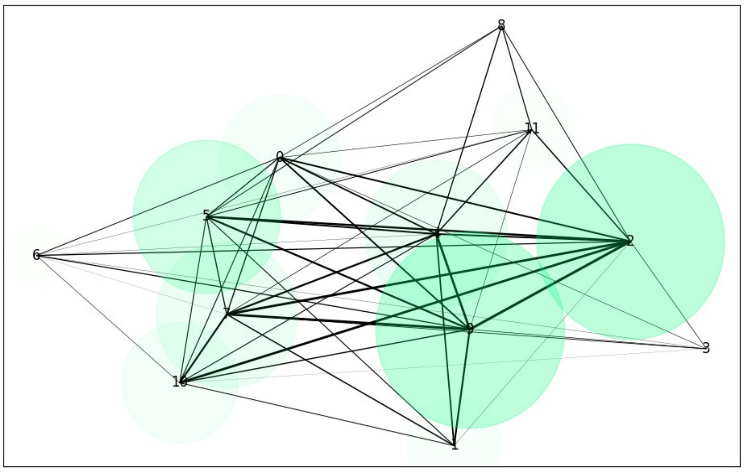


Figure 1: Graph showing relations among pitch classes of

J.S. Bach's Contrapunctus I from The Art of Fugue.

References

[1] M. Newman, *Networks*, 2nd ed., Oxford University Press, United Kingdom, 2018.

[2] Mowshowitz, A.; Dehmer, M. Entropy and the Complexity of Graphs Revisited. Entropy 2012, 14, 559-570. https://doi.org/10.3390/e14030559

[3] Albini, Giovanni & Bernardi, Marco. (2019). Graph Theory and Music: A Mathematical Tool for Musicians. 10.1007/978-3-319-95588-9_162.

Emergence of Abrupt Phase Transitions in Urban Congestion Localization

A. Lampo[1], *J. Borge-Holthoefer*[2], *S. Gómez*[1], *A. Solé-RIbalta*[1,3]. *(1) Universitat Rovira i Virgili, Tarragona, Catalonia, nello.lampo@gmail.com. (2) Universitat Oberta de Catalunya, Catalonia. (3) University of Zurich, Zurich, Switzerland.*

The sustainability of urban life represents one of the greatest challenges of our time. In this context, the engineering and optimization of road transportation networks stands up as a crucial task to ensure the efficiency of traffic dynamics and to control the related congestion emergence. In this work we analyze almost one hundred of real road networks worldwide and look into the geographic location where the congestion onset occurs. We show that congestion points experience spatial abrupt transitions, namely they shift away from the city center as larger areas are incorporated in the urban system [Fig. 1 (a)-(b)]. To explain this spatial phenomenon of congestion (evaluated in terms of maximum betweenness centrality) we introduce a family of random planar network models composed by a dense center surrounded by an arboreal periphery [Fig. 1 (c)]. Our model allows us to analytically, numerically and experimentally describe the congestion behavior and to single out the fundamental ingredients behind the unveiled transitions [Fig. 1 (d)]. Elaborating on the implications of our results, we show that they may have influence in the design and optimization of road networks regarding urban growth and the management of daily traffic dynamics.

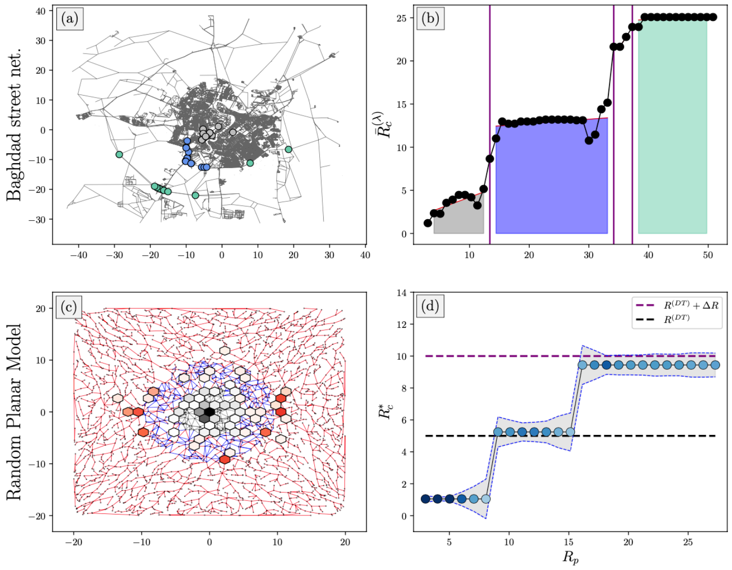


Figure 1: Spatial behavior of congestion nodes for the Baghdad empirical street network (a) and the random planar model (c). In the panel (b) and (d) we present the dependence of the congestion radius on the size of a selected city area, for the empiric and synthetic network respectively.

References

[1] A. Lampo, J. Borge-Holthoefer, S. Gómez, A. Solé-Ribalta. Multiple abrupt phase transitions in urban transport congestion. *Physical Review Research*, 3 013267 (2021).

[2] A. Lampo, J. Borge-Holthoefer, S. Gómez, A. Solé-Ribalta, Emergence of spatial transitions in urban congestion dynamics. *Journal of Applied Network Science* (in press, 2021) arXiv: 2103.04833

Description of Mesoscale Pattern Formation in Shallow Convective Cloud Fields by using Time-dependent Ginzburg-Landau and Swift-Hohenberg Stochastic Equations

***Diana L. Monroy**, Gerardo G. Naumis[1] dial_mm@ciencias.unan.mx. (1) Departamento de Sistemas Complejos, Instituto de Física, Universidad Nacional Autónoma de México, Apartado Postal 20-364, 01000, Ciudad de México, CDMX, Mexico*

The time-dependent Ginzburg-Landau (or Allen-Cahn) equation and the Swift-Hohenberg equation, both added with a stochastic term, are proposed to describe cloud pattern formation and cloud regime phase transitions of shallow convective clouds organized in mesoscale systems. The starting point is the Hottovy-Stechmann linear spatiotemporal stochastic model for tropical precipitation, used to describe the dynamics of water vapor and tropical convection. By taking into account that shallow stratiform clouds are close to a self-organized criticality and that water vapor content is the order parameter, it is observed that sources must have nonlinear terms in the equation to include the dynamical feedback due to precipitation and evaporation. The nonlinear terms are derived by using the known mean field of the Ising model, as the Hottovy-Stechmann linear model presents the same probability distribution. The inclusion of this nonlinearity leads to a kind of time-dependent Ginzburg-Landau stochastic equation, originally used to describe superconductivity phases. By performing numerical simulations, pattern formation is observed. These patterns are better compared with real satellite observations than the pure linear model. This is done by comparing the spatial Fourier transform of real and numerical cloud fields. However, for highly ordered cellular convective phases, considered as a form of Rayleigh-Bénard convection in moist atmospheric air, the Ginzburg-Landau model does not allow us to reproduce such patterns. Therefore, a change in the form of the small-scale flux convergence term in the equation for moist atmospheric air is proposed. This allows us to derive a Swift-Hohenberg equation. In the case of closed cellular and roll convection, the resulting patterns are much more organized than the ones obtained from the Ginzburg-Landau equation and better reproduce satellite observations as, for example, horizontal convective fields.

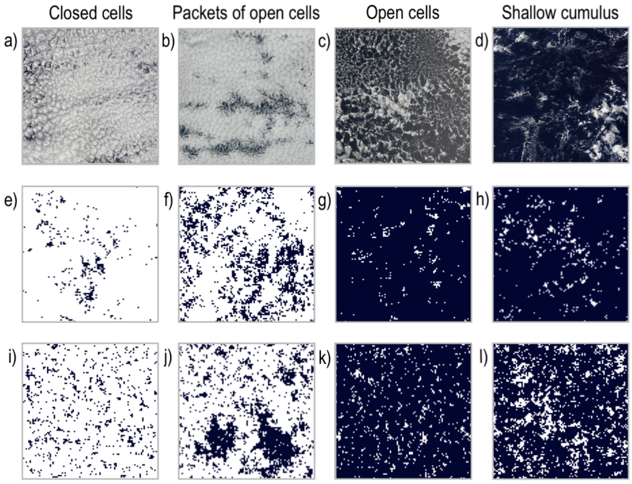


Figure 1: The four distinctive phases of shallow cloud organization.

References

[1] D. L. Monroy and G. G. Naumis, Description of mesoscale pattern formation in shallow convective cloud fields by using time-dependent ginzburg-landau and swift-hohenberg stochastic equations, Phys. Rev. E103,032312 (2021)

Burstiness and 1/f noise in the physical activities of children and adults

M. Takeuchi[1,2]*, Y. Sano*[1]*, (1) University of Tsukuba, Tsukuba, Ibaraki, Japan, takemak@gmail. (2) CyberAgent, Inc., Tokyo, Japan*

Burstiness in human behavior has been observed at various levels, from individual behavior to collective social phenomena [1]. Many studies related to burstiness in human dynamics use web-based data in various contexts [2]. In such cases, the state and attributes (e.g., demographic information) of the individuals are often unknown. On the other hand, there are still few studies on human behavior, particularly on small children using physical activity data, where subjects' states and attributes are clear. Previous research on human physical behavior has shown that the distribution of inter-event times (IETs) changes with mental illness [3]. Burstiness of human behavior may depend on the various states of the individual. Thus, investigating how burstiness is affected by an individual's state and attributes will help to clarify the rules behind human behavior and apply medical implications.

In this study, we measured human physical activities using accelerometers, along with recording the subjects' states (e.g., active or resting), and investigated those burstiness and temporal correlation. Here, subjects are an infant (2-month-old), children (3-year-old), and adults.

We found that children showed a lower scaling exponent in the cumulative distribution of IETs (Fig. 1). In other words, children have higher burstiness than adults. Also, some of the physical activity data show 1/*f* noise, which has also been identified in the brain activities [4]. We will discuss the point processes that may generate our observed features. Moreover, we are planning to measure children and their parents' activities simultaneously to explore the impact of their individual burstiness on their temporal communication network in the future.

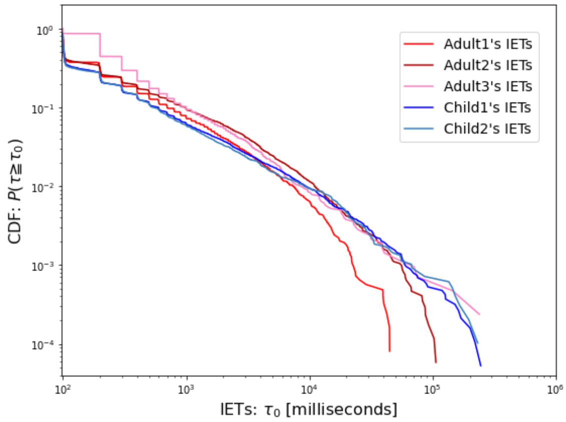


Figure 1: CDF of IETs for adults and children.

References

[1] M. Karsai, et al., *Bursty Human Dynamics, Springer*, 2018.

[2] A.-L.Barabási, Nature435, 207–211 (2005).

[3] T. Nakamura, et al., *PLoS ONE,* 3(4), e2050 (2008).

[4] E. Novikov, et al., *Phys. Rev. E* 56, R2387(R) (1997).

Change of network structure of elastomers under elongation

M. Ito[1], *Y. Amamoto*[2], *T. Ohnishi*[1]. *(1) Rikkyo University, Toshima, Tokyo, Japan, ito.mariko@rikkyo.ac.jp*
(2) Kyushu University, Motooka, Fukuoka, Japan

Elastomers are polymers with elasticity and have a molecular structure where polymer chains are cross-linked. Such a structure can be represented by a network composed of nodes each of which stands for a crosslinking point. Each link can be regarded as a connection by a polymer chain between two crosslinking points. Amamoto et al. defined *modified centrality* of a crosslinking point as closeness centrality on a weighted network of elastomer where each link has a weight induced by spatial distance between crosslinking points [1]. They showed that physical property around a crosslinking point under elongation of elastomer is significantly related to its modified centrality before elongation, by analyzing data generated through coarse-grained molecular dynamics simulations.

In this study, we ask how the weighted network structure of elastomer changes along with elongation—though the network topology of elastomer is unchanged, its weighted network structure should evolve during elongation since it reflects the distance between crosslinking points. We defined a centrality of a link based on the previous study. For the aforementioned simulation data, we examined transition of the centrality during elongation for each crosslinking point. Stability of the transition of centrality depended on the concentration under which crosslinking reaction was carried out. Centralities of links fluctuated widely till a certain elongation ratio at high concentration, whereas they were relatively stable during elongation at low concentration. This implicates that, during elongation, the weighted network structure of elastomer changed qualitatively differently according to the reaction concentration. As a future perspective, we may be able to find a characteristic of elastomer's deformation by investigating on coevolution of the centrality and the weight of each link.

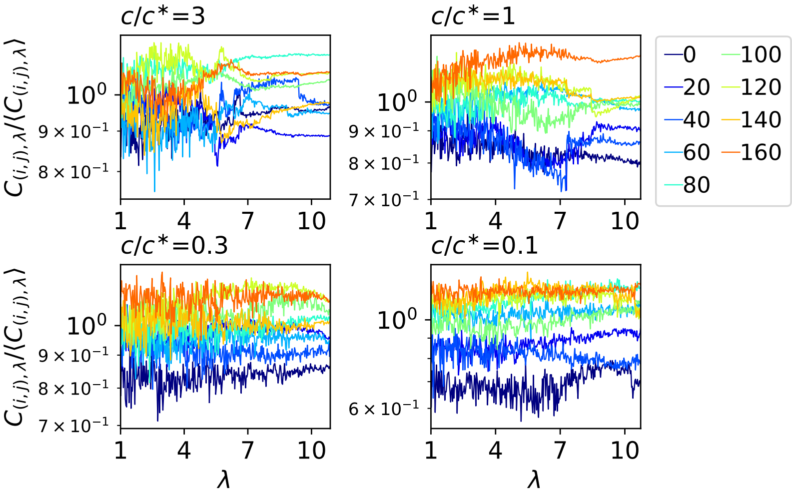


Figure 1: Transition of centralities during elongation. Each panel exhibits the transition of centrality for 9 links in a sample for each reaction concentration c/c^* (c^* denotes the overlap concentration). The vertical and the horizontal axes exhibit elongation ratio and centrality, respectively. Plots are colored according to the ascending order of centralities before elongation.

Acknowledgements This work was supported by JSPS KAKENHI Grant Number JP17H06460, JP17H06468, JP20H04644 and JP20H02800.

Reference [1] Y. Amamoto, T. Ohnishi, et al., *Patterns*, 1 (2020) 100135.

Emergence of Popular Hashtags and Ranking Dynamics on the Chinese Microblogging Complex System

***H. Cui*[1]**, *J.Kertész[1]. (1) Department of Network and Data Science, Central European University, Quellenstrasse 51, A-1100, Vienna, Austria, cui_hao@phd.ceu.edu.*

Microblogging websites are complex systems where the users interact with each other and generate contents usually marked by hashtags. Some hashtags become popular by being reposted by a large number of users. As time passes, new popular contents emerge, and the popularity of the old ones vanishes. On the Chinese microblogging website Sina Weibo, the real-time Hot Search List (HSL) is a reflection of hashtag popularity among all users in the whole system. The patterns of the HSL in the Weibo system are influenced by the circadian behaviors of the users and external factors. In the first part of the talk, we present observed daily patterns in the cumulative number of hashtags, search volume indices and rank diversity. We show a case study on how the COVID-19 pandemic has influenced the emergence of different kinds of hashtags and the ranking on the HSL. In the second part of the talk, we present two emergence mechanisms of the hot topics by showing their network evolution patterns before appearing on the HSL: In one case ("Born in Rome") the hashtag almost immediately gets to the HSL, while in the other case (“Sleeping Beauty”) there is a considerable time lag before that happens. Accounts with a large number of followers have to be involved to create sufficient popularity for a hashtag so that it can reach HSL. We also investigate the factors that influence the success of the hot topics to become popular and enter the HSL. We observe the importance of the timing of the creation as well as the contents of the posts and technique of presentation. Finally, we quantify the fostering effects of the HSL on hashtags which reach that level.

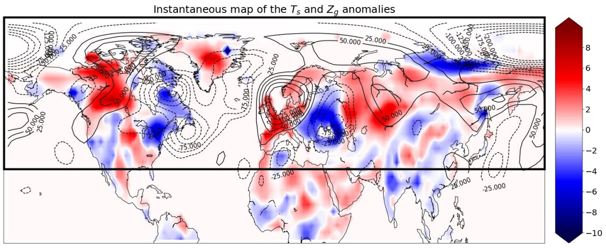
Deep Learning based Extreme Heatwave Forecast

V. Jacques-Dumas[1], *F. Ragone*[1,2], *F. Bouchet*[1], *P. Borgnat*[1], *P. Abry*[1] *(1) Laboratoire de Physique (UMR 5672), CNRS, ENS de Lyon, Univ Lyon, France (contact : pierre.borgnat@ens-lyon.fr). (2) Earth and Life Institute, Université Catholique de Louvain, Louvain-La-Neuve, Belgium*

Forecasting the occurrence of heatwaves is a challenging issue, which involves complex (nonlinear, nonstationary) data modeling and a challenge is the elaboration of advanced and innovative forecasting methods. A specificity is that extreme events, of major societal stake, are not often observed and (very) costly to simulate from physics-driven numerical models. For the problem of forecast of extreme heatwave (typically, the 5% to the 1% most important), we propose Deep Learning architectures as viable strategies for prediction, even from a quite limited amount of available relevant climate data. The present work uses 1000 years of the PlaSim (Planete Simulator Climate Model) data from [1]; we say that there is an anomaly temperature of duration D above some area $\mathscr{A}$ and of strength a, and time t when the average of the temperature fluctuations exceed a:

$$Y(t) = \frac{1}{D}\int_t^{t+D} \frac{1}{\mathcal{A}}\int_{\mathcal{A}} \left(T_s - \langle T_s \rangle\right)(\vec{r}, u)\mathrm{d}\vec{r}\mathrm{d}u > a.$$

Then, to forecast the existence of such an anomaly, we train a deep learning neural network (based on CNN) that aggregates climate data of different natures (surface temperature and/or geopotential height), and we consider the problem as a supervised classification problem (*Y(t)* is equal to 1 during heatwave, else 0) trained with the cross entropy as loss function. The important point, and a novelty, is to show how to cope with the class-size imbalance that is intrinsically associated with rare event prediction. Undersampling [2] allows to correct a part of this imbalance. Yet, really extreme events (1.25% in our study, e.g. associated to a temperature anomaly of more than 4.2 K over France, for D=14 days in summer) are still difficulty to predict. We improve the training of the predictor thanks to transfer learning [3] that leverage the nested nature of extreme events (naturally included in less extreme ones). This combination of large-class undersampling and transfer learning achieves good performance in forecasting the occurrence of extreme heatwaves, as early as 15 days in advance from the observation, for a single time.



Without transfer learning	Stacked FT of T_s and Z_g		
Levels	5%	2.5%	1.25%
Average MCC	0.44	0.37	0.29
Median MCC	0.44	0.37	0.29
Std MCC	0.03	0.03	0.06
With transfer learning			
Average MCC	–	0.40	0.35
Median MCC	–	0.40	0.36
Std MCC	–	0.04	0.05

Figure 1: Example of heatwave (the maps of temperature and geopotential anomalies). Right: performance (in term of the Matthews Correlation Coefficient, MCC) of the prediction, with or without transfer learning, with aggregated (stacked) fields and for intensity ranging from 5% to 1.25%. With transfer learning, we obtain a good prediction for the most extreme events (1.25%).

Acknowledgements: Work supported by the ACADEMICS Grant of IDEXLYON, Univ. Lyon, PIA ANR-16-IDEX-0005.

References. Full article available online: arXiv preprint arXiv:2103.09743 (2021)

[1] F. Ragone, J. Wouters, and F. Bouchet. *PNAS*, 115(1):24–29, 2018.

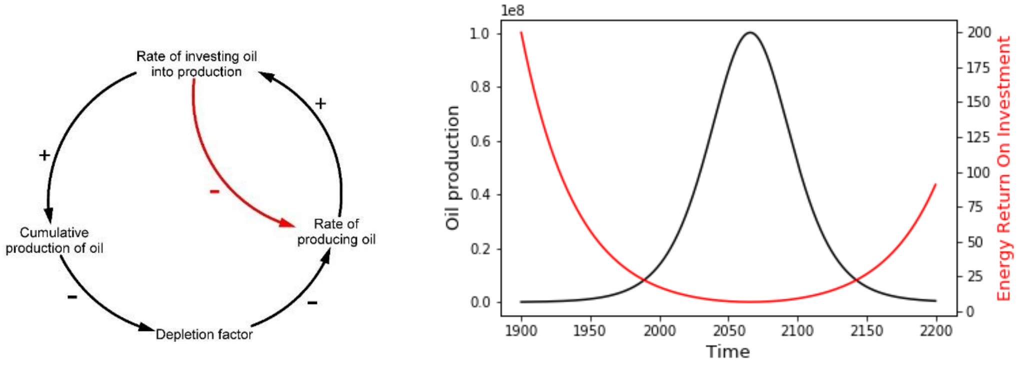
[2] J.M. Johnson and T.M. Khoshgoftaar. *Information Systems Front.*, (22):1113–1131, 2020.

[3] L. Y. Pratt. *Advances in Neural Information Processing Systems*, 5:204–211, 1993.

Peak Oil and the Minimum EROI for Society

S. Roman[1]. *(1) Centre for the Study of Existential Risk, University of Cambridge, sr911@cam.ac.uk*

We propose a class of low-dimensional dynamical systems to model the production of non-renewable energy that highlight one key insight: the peak of production coincides with the minimum of the energy return on investment (EROI). Since Hubbert proposed the theory of peak oil [1] there has been a growing concern for when worldwide production of crude oil would reach its maximum and then start declining. The first decade of the 21st century saw a flurry of predictions [2, 3] and increased distress regarding the topic which has since subsided. The extraction of non-conventional oil resources has helped mitigate the problem [4] but does not address the issue of declining EROI. There is a minimum EROI necessary for society with its modern facilities (e.g., healthcare, education) to function properly [5] or otherwise face the possibility of collapse. Our model indicates that the concern for peak oil was misplaced: society can experience low, even unsustainable EROI, well before global production would ever peak.



(a) (b)

Figure 1: (a) Feedback mechanisms in the class of models we propose that generalize Hubbert's theory of oil production. The models include two stocks which are themselves rates of change. The structure of the causal loop diagram reflects the general feedback mechanisms found in the long-term dynamics of societies that underly their sustainability or collapse [6]. (b) Illustrative run of the model that shows how EROI can decrease substantially well before the peak of oil production.

References

[1] Hubbert, M. K. (1959, March). Techniques of prediction with application to the petroleum industry. In *44th Annual meeting of the American Association of Petroleum Geologists* (Vol. 43). Shell Development Company Dallas, TX.
[2] Deffeyes, K. S. (2001). *Hubbert's peak: the impending world oil shortage*. Princeton Univ. Press.
[3] Hirsch, R. L. (2005). The inevitable peaking of world oil production. *Bulletin of the Atlantic Council of the United States*, *16*(3), 1-10.
[4] Bardi, U. (2019). Peak oil, 20 years later: failed prediction or useful insight?. *Energy research & social science*, *48*, 257-261.
[5] Hall, C. A., Balogh, S., & Murphy, D. J. (2009). What is the minimum EROI that a sustainable society must have?. *Energies*, *2*(1), 25-47.
[6] Roman, S. (2020). Long-term feedback mechanisms underlying societal collapse. *Conference on Complex System 2020.* DOI: https://doi.org/10.5281/zenodo.4427919

Prediction of New Outlinks on the World Wide Web

T. K. N. Dang[1]*, D. Bucur*[1]*, B. Atil*[2]*, G. Pitel*[3]*, F. Ruis*[1]*, N. Litvak*[1,4]*. (1) University of Twente, Drienerlolaan 5, 7522 NB, Enschede, The Netherlands, t.k.n.dang@utwente.nl. (2) Bogazici University, Bebek, 34342 Istanbul, Turkey. (3) Exensa, 2e Rue Perier, 92120 Montrouge, France. (4) Eindhoven University of Technology, Den Dolech 2, 5612 AZ, Eindhoven, The Netherlands*

Detecting change on the World Wide Web is essential for web crawling in order to keep the web page repository up to date. In this work, we introduce prediction models for two metrics that are important in organizing the crawling order of pages: change rate of *internal* (i.e., from the same domain and HTTP communication protocol as the page) or *external* (i.e., from different domains or protocols) outlinks, and the number of new *internal/external* outlinks. We consider four sets of features. Static page features (SP) include, e.g., content size, and the current number of outlinks. Dynamic page features (DP) are historical values of those SP features which can change, and the number of new outlinks. In the experiments we have included DP features up to 8 weeks. Static network features (SN) include, e.g., the TrustRank of the page, and the number of inlinks. Dynamic network features (DN) simply are the historical values of SN features. The results show that SP features have a high predicting value for both *internal* outlink change rate and new *internal* outlinks. Specifically, 65% of the variance in the outlink change rate is explained with a low mean absolute error of 14%. The change rate pattern can be visualized with only two important SP features, Fig. 1 (left). For the classification model, we compare to a simple baseline: the change next week is predicted to be the same as in the previous week, which also gives good predictions, Fig. 1 (right). In addition, we observe that prediction can be made successfully based on only recently historical information, Fig. 1 (right). For simplicity, we omit the results of *external* outlinks. We also make probabilistic predictions of new outlinks by adapting the recent method NGBoost to learn parameters from our full set of features. This offers the means to contrast the real data to the Poisson model and also evaluate other probabilistic models like zero-inflated Poisson and negative binomial.

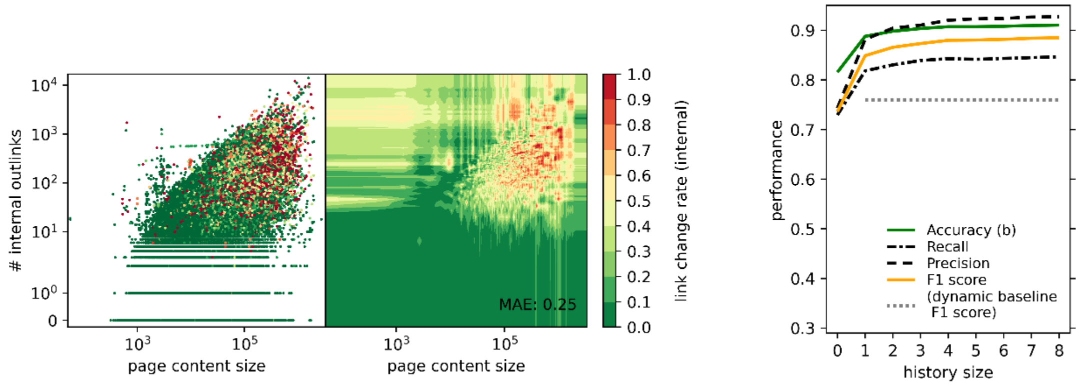


Figure 1: Left: Predicting the change rate of *internal* outlinks using two most important SP features. (1^{st} panel) Empirical data and (2^{nd} panel) regressors trained with the SP features. For unavailable empirical data the regressors create extrapolated values. Right: Classification scores with different history sizes of target new *internal* outlinks with all features

Acknowledgements

This work is supported by the project Eurostar E!113204 WebInsight.

Alphabetical Index

Index