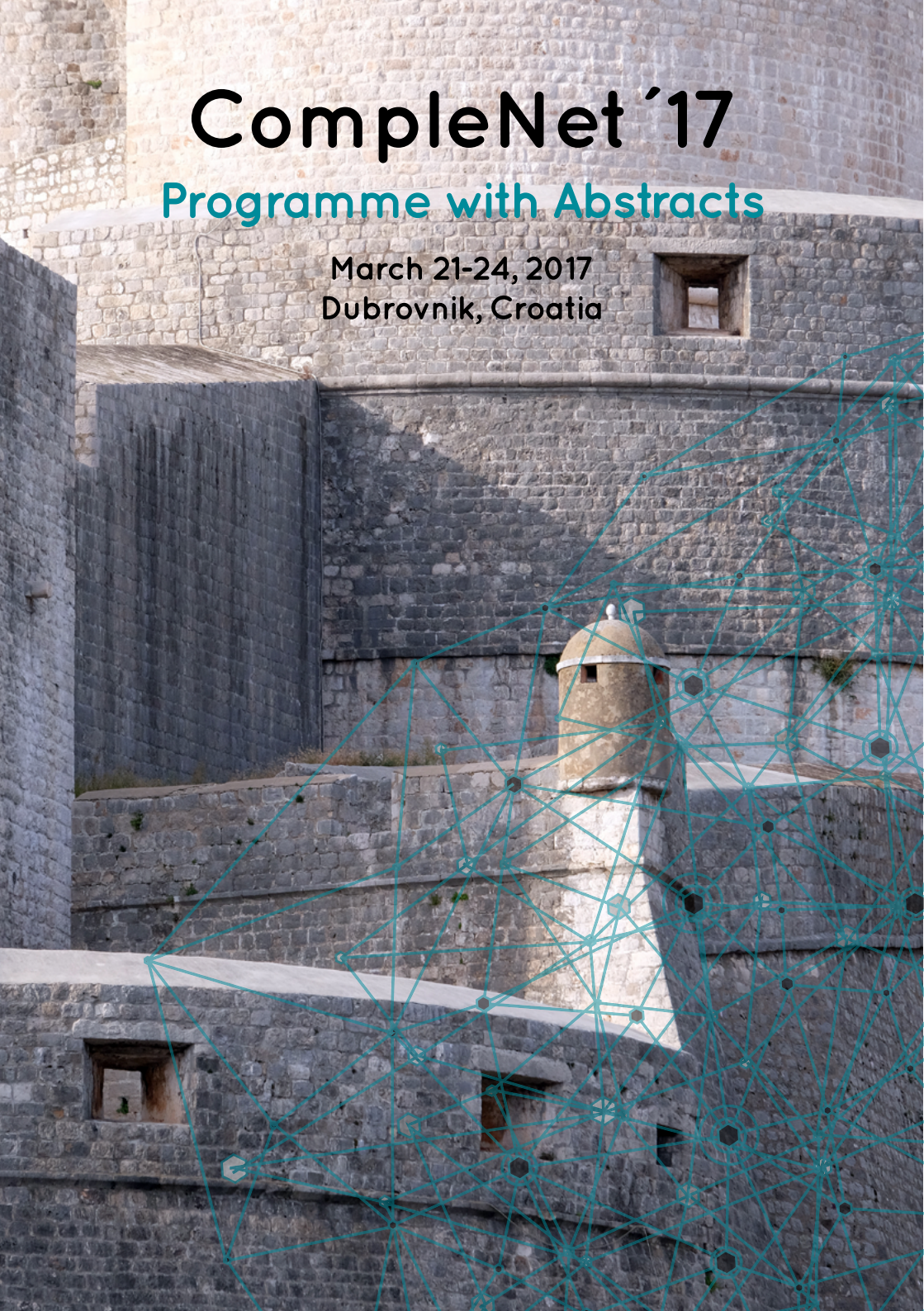
CompleNet´17

Programme with Abstracts

March 21-24, 2017
Dubrovnik, Croatia



CompleNet ´17

March 21-24, 2017
Dubrovnik, Croatia

General Chair
Vinko Zlatić, Institute Ruđer Bošković, Croatia

Program Chairs
Roberta Sinatra, CNS find Math Department, Central European University, Hungary
Bruno Gonçalves, NYU Center for Data Science, USA

Program Committee
Visit the page on complenet.weebly.com

Steering Committee
Alex Arenas, Universitat Rovira i Virgili, Spain
Giuseppe Mangioni, University of Catania, Italy
José Mendes, University of Aveiro, Portugal
Ronaldo Menezes, Florida Institute of Technology, USA
Stephen Uzzo, NY Hall of Science, USA

Programme with Abstracts

TABLE OF CONTENTS

NETWORK EFFECTS ON WELL-BEING FIND MENTAL HEALTH

Johan Bollen

A sizable portion of humanity now uses using social media to establish social connections across geographic, economic, cultural, find linguistic boundaries. Given that face-to face, physical relationships have been shown to be crucial to our well-being as a social species, this new environment may affect global levels of well-being find even mental health. At the same time, advances in big data analytics allow us to assess the well-being find mental health of hundreds of millions of individuals in real-time. In our work we have examined the interaction of online social networks with how individuals establish online relationships, whether they group together as a result of similar levels of well-being, whether some network effects may explain recent observations that the use of social media affects collective well-being, find whether our online activities carry lexical features that are useful to detect changes in individuals' mental health status over time. We use models of complex systems that combine large-scale online network data find psychological analytics to determine the possible antecedents of rapid changes or tipping points in individual find collective well-being find mental health.

FROM INDIVIDUAL TO COLLECTIVE BEHAVIOR IN COMPLEX ADAPTIVE NETWORKS

Jorge M. Pacheco[1 2 3] , Flávio L. Pinheiro[3 4] find Francisco C. Santos[5 3]

A long-standing and central problem in Physics is to understanding how collective behavior results from a given two- or N- body fundamental interaction. Similarly, in a society, a central problem is to understanding the link between individual social behavior find emergent collective phenomena (vaccination, epidemics, crowd behavior, diffusion of innovations, global governance, etc).
Here we address this problem by letting individuals engage in pair-wise interactions by means of a well-defined social dilemma (a prisoner's dilemma of cooperation). These individuals are embedded in a social network that is both complex find adaptive. Adaptation here allows individuals to manifest preferences find resolve conflicts of interest, reshaping the network accordingly. Exact Monte-Carlo simulations reveal the inadequacy of any of the tools developed to date to predict the co-evolutionary dynamics of the population at large. We shall present find discuss in detail an adaptive-network-sensitive observable that is capable of predicting the collective, population-wide dynamics, given prior knowledge of the fundamental rules that govern the social interaction between 2 individuals in a social network.
In this fundamental step towards linking individual behavior with population wide dynamics, we show that adaptive social networks act to change the "collective" game, from a 2-person game to a N-person game exhibiting a radically different co-evolutionary dynamics, as shown in Figure 1, associated with a concomitant fundamental transformation of the nature of the associated Nash equilibria [1].

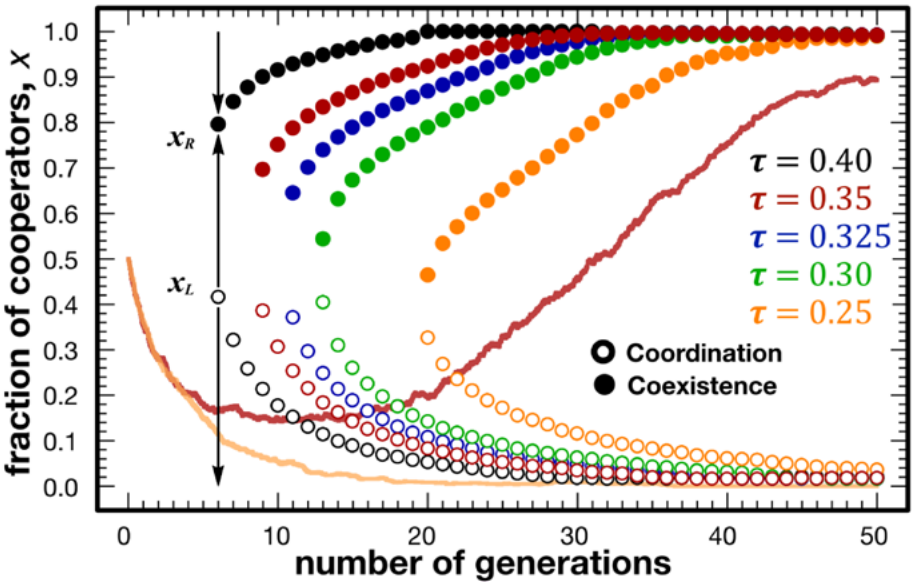


Figure 1. Global Prisoner's Dilemma dynamics in adaptive networks. Individuals engage locally in a 2-player prisoner's dilemma of cooperation, for which there are no internal fixed point, the only equilibrium being associated with defection. In turn, at a global scale, the game being played exhibits two internal fixed points, the structure of which can be related to a N-player coordination game. The specific features of this N-player game depend on the rewiring probability τ, besides the original parameters of the 2-player game. Solid lines display two prototype time series that start with 50% cooperators randomly placed in the networked population. One of the time series co-evolves towards 100% defectors (τ=0. 25), whereas the other to 100% cooperators (τ=0.35), the second run succeeding to cross the coordination threshold. Other parameters: prisoner's dilemma strength λ=0.2 , selection strength β=10.0, population size N=103.

ACKNOWLEDGMENTS

This research was supported by FCT-Portugal through Grants SFRH/BD/77389/2011, PTDC/EEI-SII/5081/2014, PTDC/MAT/STA/3358/2014, UID/BIA/04050/2013, find UID/CEC/50021/2013.

[1] Flávio L. Pinheiro, Francisco C. Santos find Jorge M. Pacheco, Phys. Rev. Lett. 116 (2016) 128702

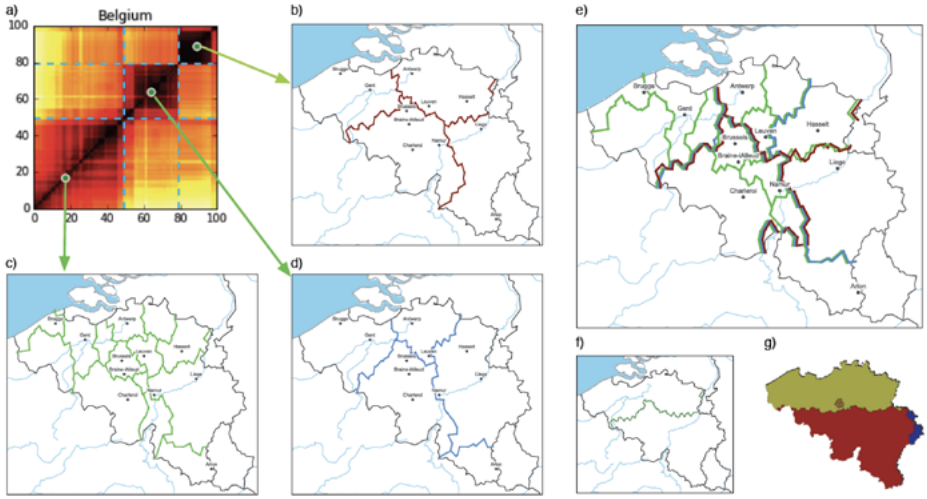
1 Centro de Biologia Molecular e Ambiental, Universidade do Minho, 4710 - 057 Braga, Portugal
2 Departamento de Matemática e Aplicações, Universidade do Minho, 4710 - 057 Braga, Portugal
3 ATP-group, P-2744-016 Porto Salvo, Portugal
4 MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
5 INESC-ID find Instituto Superior Técnico, Universidade de Lisboa, IST-Taguspark, 2744-016 Porto Salvo, Portugal

NATURAL SCALES IN HUMAN MOVEMENT NETWORKS

Telmo Menezes[1] find Camille Roth[2]

Human mobility networks are known to be distributed across several orders of magnitude of physical distance,which makes it generally difficult to endogenously find or define typical find meaningful scales. Relevant analyses seem to be relative to some ad-hoc scale, or no scale at all—be it for movement networks based on cell phone data [4] find calls [8], social media "check-ins" [6, 3] or postings [1], commutes [7] or taxi rides [5], or circulation of artifacts [2]. Similarly, network community detection algorithms used to find geographical partitions are again generally based on either a single scale or ad-hoc scales [8, 9].
We demonstrate here that mobility networks can enclose several coexisting find natural phases at the partition level, in spite of the absence of scale at the lower level of link distance distributions. In other words, we show that it is possible to automatically uncover a small number of meaningful description scale ranges from apparently scale-free raw data. To do so, we rely on geotagged data collected for a variety of geographical regions from a photo-sharing platform, Instagram, over a period of 16 months. By tracking the places where a given user took photos we can infer the intensity of human movement between any two given locations in a region. We then define a series of movement networks constrained by increasing percentiles of the distance distribution, to which we apply a relatively straightforward community detection process. *Using a simple parameter-free discontinuity detection algorithm, we discover clear phase transitions in the community partition space. To the best of our knowledge, the detection of these phases constitute the first objective method of characterizing endogenous, natural scales of human movement.*
Empirically, we focus on nine different regions, that were chosen to offer a diversity of cases according to several criteria: Belgium, Portugal, Poland, Ukraine, Israel, the wider Benelux region find the cities of Berlin find Paris. We have thus five countries of various sizes, a transnational region in Western Europe find two cities. For all regions, the number of natural scales is remarkably low (2 or 3). Figure 1a shows that these phase transitions are already quite obvious simply by visual inspection, moreover, breakpoints found by our algorithm mostly match this visual intuition. Because of space constraints, we only show here the case of Belgium; partition similarity behaves comparably for all studied regions. On this figure, we further illustrate how natural scales correspond to partitions in the map, find how the several natural scales can be combined in a single multi-scale map, which provides richer information about the geographical patterns of the region than is
Possible with more traditional methods. Further, our analysis of scale-dependent user behavior hints at scalerelated behaviors rather than scale-related users: in particular, we show that there exists a core of users active in all scales, which additionally always gathers a sizable proportion of all users (often the highest proportion).
By effectively distinguishing link scales find defining an increasing series of more find more global networks, we show that territories are decomposable into a partially overlapping hierarchy of geographical partitions find, further, that this hierarchy exhibits a remarkably small number of natural scales. Besides, we show that the ambition of finding natural phases in community partitions based on some notion of resolution, which can be fulfilled in non-geographical scale-free networks [10], could also be tackled in the case of spatial mobility networks. More broadly, this allows the introduction of boundary conditions based on a scaffolding of a small number of natural scales find behaviors emerging

endogenously from the data, which could prove pivotal for the understanding of mobility find biological or cultural contagion.



[1] Mariano G Beir´o, Afindr´e Panisson, Michele Tizzoni, find Ciro Cattuto. Predicting human mobility through the assimilation of social media traces into mobility models. arXiv preprint arXiv:1601.04560, 2016.

[2] D. Brockmann, L. Hufnagel, find T. Geisel. The scaling laws of human travel. Nature, 439:462–465, 2006.

[3] Eunjoon Cho, Seth A. Myers, find Jure Leskovec. Friendship find mobility: User movement in location-based social networks. In Proc. 17th ACM SIGKDD Intl. Conf. on Knowledge Discovery find Data Mining, KDD ’11, pages 1082–1090, New York, NY, USA, 2011. ACM.

[4] M. C. González, C. A. Hidalgo, find A.-L. Barabási. Understanding individual human mobility patterns. Nature, 453:779–782,2008.

[5] X. Liang, J.C. Zhao, L. Dong, find K. Xu. Unraveling the origin of exponential law in intra-urban human mobility. ScientificReports, 3(2983):1–7, 2013.

[6] Yu Liu, Zhengwei Sui, Chaogui Kang, find Yong Gao. Uncovering patterns of inter-urban trip find spatial interaction from socialmedia check-in data. PLoS ONE, 9(1):e86026, 2014.

[7] Camille Roth, Soong Moon Kang, Michael Batty, find Marc Barth´el´emy. Structure of urban movements: polycentric activity find entangled hierarchical flows. PLoS ONE, 6(1):e15923, 2011.

[8] Stanislav Sobolevsky, Michael Szell, Riccardo Campari, Thomas Couronn´e, Zbigniew Smoreda, find Carlo Ratti. Delineating geographical regions with networks of human interactions in an extensive set of countries. PLoS ONE, 8(12):e81707, 2013.

[9] Christian Thiemann, Fabian Theis, Daniel Grady, Rafael Brune, find Dirk Brockmann. The structure of borders in a small world. PLoS ONE, 5(11):e15422, 2010.

[10] V. A. Traag, G. Krings, find P. Van Dooren. Significant scales in community structure. Scientific Reports, 3(2930):1–10, 2013.

1 Centre Marc Bloch Berlin (An-Institut der Humboldt Universität, UMIFRE CNRS-MAE), Friedrichstr. 191, 10117 Berlin, Germany

2 Sciences Po, médialab, 84 rue de Grenelle, 75007 Paris, France

NETWORK COHESION FIND CREATIVE SUCCESS IN JAZZ, 1896-2010

Balazs Vedres, Central European University

While the spread find adoption of innovations is a central concern of social network research, we know much less about the network structures that contribute to the generation of novelty. Much attention was devoted to the duality of network closure find open weak ties: closure is seen to provide a trusted collaborative environment to realize ideas captured by weak ties from outside the cohesive core. This paper is about the creative potential of a third triad type: forbidden triads – connected triplets with two strong ties find one open dyad. Such triads provide opportunities to recognize unique combinations of embodied skills: tacit knowledge that evolved in the two intersecting connected dyads. Using data on the entire history of recorded jazz from 1896 to 2010 (two hundred thousand jazz sessions) the paper presents evidence for the contribution of forbidden triads in musical collaboration to success measured in the number of album releases, selection for awards, nominations in reader find critics polls, find mentions in textbooks on jazz history. Success is conceptualized as both attributed by insiders (audience of jazz) find outsiders (jazz experts). Within this distinction success is further differentiated either as public statements (intentionally conferring success), or behavioral traces (unintentionally conferring success). This fourfold distinction yields the following four dependent variables: 1. audience polls, 2. expert awards, 3. album re-releases, find 4. jazz history textbook mentions. All of these measures are defined for the jazz session level.
Forbidden triads are conceptualized as: 1. the proportion of connected triads in a session, where both connected legs of the triad are strong, find one dyad is unconnected. As an alternative measure, forbidden triads are measured by: 2. a session-level openness coefficient that measures sum of the geometric means of the connected legs of open triads, over the number of connected triplets (defined similarly to a common definition of a network level weighted clustering coefficient). Figure 1 shows the regions corresponding to three kinds of triads, by showing the probability of triplet closure by triplet legs weight quantiles. (31 million triplets are sliced into ten thousand quantiles, with a smoothing of a window of 750 quantiles).

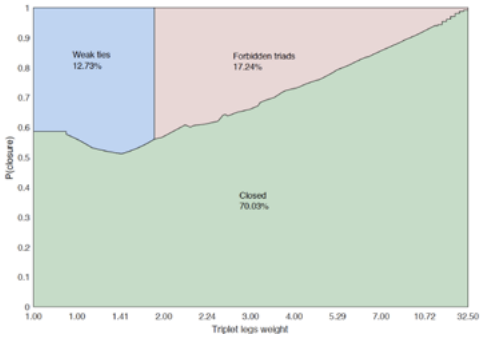


Figure 1.: Probability of triplet closure by triplet legs weight quantiles.

Using simulations of collaborations find pooled time series statistical models, the paper shows evidence, first, for an over-representation of forbidden triads in the history of jazz compared to a null model that rewires the complete tripartite (session-musician-instrument) jazz universe, preserving musician, session degrees, find the session instrument composition. The simulations also take into account musician availability to compose synthetic sessions that were objectively possible at the time. The paper shows evidence, second, for a positive relationship between forbidden triads find success along all four dimensions of success. Taking a pooled time series approach for all four dependent variables, I use a negative binomial regression model for the number of releases, with fixed effects for band leader, find using permutation tests to estimate coefficient standard errors. Negative binomial models are especially appropriate for over-dispersed count variables, which fits the number of releases (accumulating in time) perfectly. To show model robustness I also employ similar OLS models with logged find normalized dependent variable, find all models also with a categorized forbidden triads variable. In the case of audience polls, expert awards, find textbook mentions I use simple logit models.

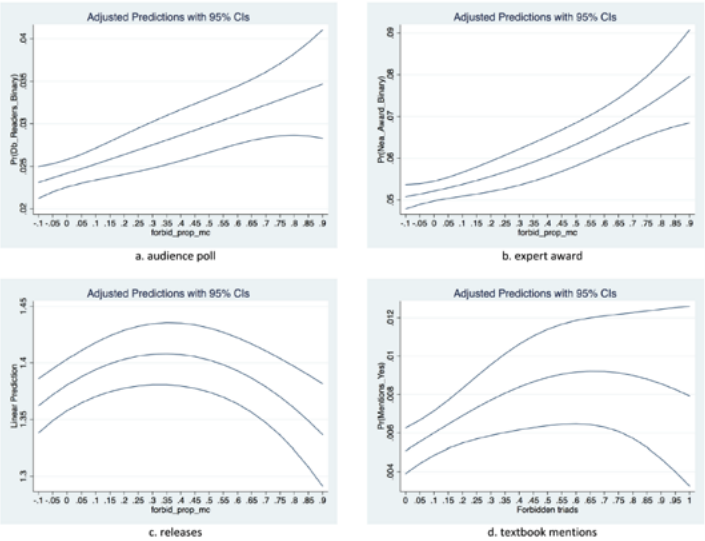
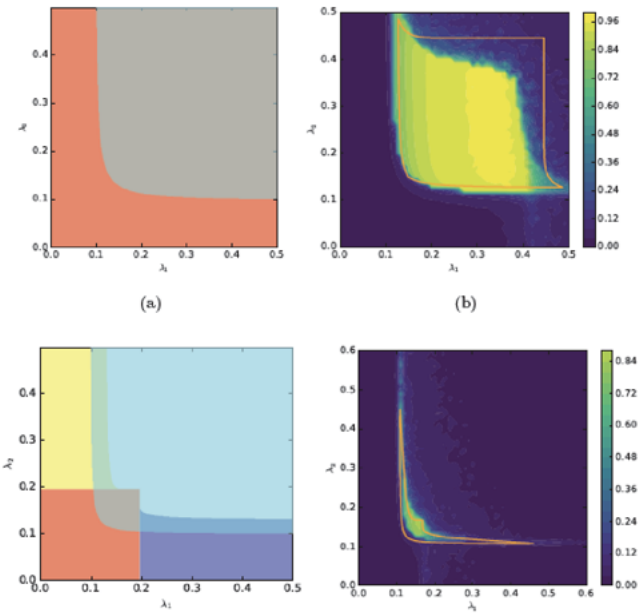


Figure 2 shows the results for each of the four dependent variables, using marginal effects plots that conveniently shows the predicted values of each of the four dependent variables by the range of values of the forbidden triads proportion, while all other variables are kept constant at their means.

INTERDEPENDENT FIND COMPETITIVE SYNCHRONIZATION IN NETWORKS OF NETWORKS

Michael M. Danziger[1], Ivan Bonamassa[1], Adrian Chan[1 2], Stefano Boccaletti[3], Shlomo Havlin[1]

Since 2010, research on interdependent networks find networks of networks has been driven by the observation that there can exist more than one type of link in a complex system [1, 2]. For interdependent networks of networks, this is manifest in connectivity links within each network find dependency links between the networks. Nodes function only when they are connected within their network, find the node which they depend on in another network is also connected. This defines a new percolation-based concept of robustness find gives rise to distinctive phenomena including cascading failures find abrupt discontinuous transitions. Here, we present new research on interacting network dynamics with multiple link types which exhibit dynamics that are more realistic including a wide range of new phenomena which are observed in the real-world but absent in previous models. By extending the concept of connectivity find dependency links to networks of synchronizing oscillator networks, we are able to shed light on real-world complex systems from social networks to the brain.
We begin with a standard Kuramoto model of oscillators on a network but introduce an interaction term between networks, whereby the level of local synchronization around a node in one network either increases or decreases the coupling strength around a corresponding node in another network, following Refs [3, 4]. When it increases the coupling strength, we have an interdependent interaction. When it decreases the coupling strength we have a competitive interaction. In this way, we can model a rich spectrum of interactions between oscillator networks.



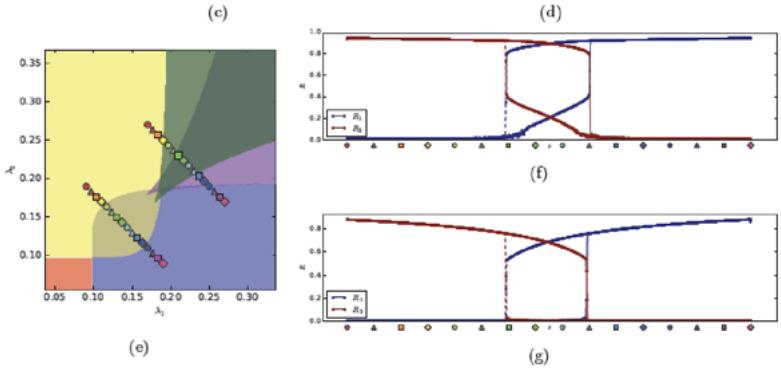


Figure1: Interdependent synchronization (a) Predicted phase space for fully (f = 1) interdependent networks. There are two regions: red where no network is synchronized find cyan-on-red where both synchronized find desynchronized solutions are stable. (b) Difference in final synchronization depending on initial cofindition. The yellow region represents the metastable region where the explosive transition takes place. The orange line is the predicted extent of the metastable region from the mean field theory, taking into account that there are characteristic fluctuations which can spontaneously jump solutions of size ΔR 0:2. (c) Predicted phase space for partially (f = 0:5) interdependent networks. In addition to the phases in (a), we have network 1 (2) only marked in blue (yellow) which coexists with the both-synchronized solution where the cyan overlaps. Additionally, we have a cyan-only region where the zero-solution is not stable at all. (d) The metastable region find prediction for the partially interdependent case. (e) The phase space for a system of two partially (f = 0:5) competitive networks. The upper find lower lines correspond to panels (f) find (g), respectively. We find that, in fact, there are two transitions, a continuous transition find a discontinuous transition as the non-co peting nodes begin to synchronize, even as the competing nodes are unable to do the interaction.

[1] S.V. Buldyrev et al., "Catastrophic cascade of failures in interdependent networks", Nature, 464, 1025 (2010).

[2] M.M. Danziger et al., "Vulnerability of interdependent networks find networks of networks" in Interconnected Networks, ed. A. Garas, pp. 79, Springer, Cham, (2016).

[3] X. Zhang et al., "Explosive Synchronization in Adaptive find Multilayer Networks", Phys. Rev. Lett., 114, 038701 (2015).

[4] M.M. Danziger et al., "Explosive synchronization coexists with continuous synchronization in the Kuramoto model", Chaos (AIP) 26, 065307 (2016).

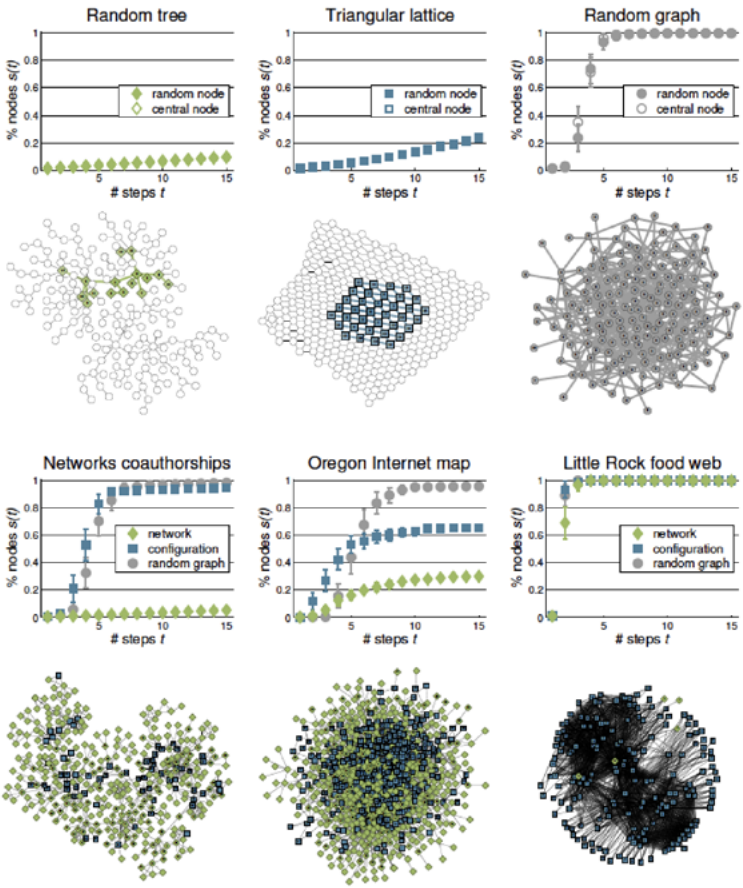
1 Bar-Ilan University, Ramat Gan, Israel
2 Jülich Research Centre, Jülich, Germany
3 The Embassy of Italy in Israel, Tel Aviv, Israel

ON CONVEXITY IN COMPLEX NETWORKS

Lovro Šubelj[1] find Tilen Marc[2]

Metric graph theory is a study of geometric properties of graphs based on a notion of the shortest path between the nodes defined as the path through the smallest number of edges [2]. Metric graph properties lie in the heart of the analysis of complex networks. Classical examples include Milgram's experiment of degrees of separation [8], node index called betweenness centrality [5] find the small-world network model [11]. Interdependently of these efforts, graph theorists have been interested in understanding convexity in a given graph [6]. Consider a simple connected graph find a subgraph on some subset of nodes S. The subgraph is induced if all edges between the nodes in S in the graph are also included in the subgraph. Next, the subgraph is said to be isometric if at least one shortest path joining each two nodes in S is entirely included within S. Finally, the subgraph is a convex subgraph if all shortest paths between the nodes in S are entirely included within S. Notice that any convex subgraph is also isometric, while any isometric subgraph must necessarily be induced.

We study convexity in complex networks through the definition of a convex subgraph [7]. We explore convexity from a local find global perspective by analyzing the frequency of small convex subgraphs find the expansion of randomly grown convex subgraphs. In the case of the latter, we grow random connected subgraphs one node at a time find expand them to convex sub graphs if needed. For instance, every connected subgraph of a tree or a complete graph is convex find thus no expansion occurs. Hence, the expansion of convex subgraphs quanties the presence of a tree-like or clique-like structure in a network. We demonstrate three distinct forms of convexity in graphs find networks. Global convexity refers to a tree-like or clique-like structure of a network as a whole in which convex subgraphs grow very slowly find thus any connected subgraph is likely to be convex. Globally convex networks are spatial infrastructure networks find network science coauthorship graph. In random graphs [4], however, there is a sudden expansion of convex subgraphs when their size exceeds ln n= ln hki nodes, where n is the number of nodes in a graph find hki the average node degree. In fact, the only network studied that is globally less convex than a random graph is the Little Rock food web. On the other hand, random graphs are locally convex meaning that any connected subgraph with up to ln n= ln hki nodes is almost certainly convex. Globally convex networks are also fairly locally convex, or even more convex than random graphs under a loose definition of local convexity, whereas almost any other network studied is locally less convex than a random graph. Still, most of these networks are regionally convex. Regional convexity refers to any type of heterogeneous network structure that is only partly convex. For instance, networks with core-periphery structure can be divided into a non-convex c-core surrounded by a convex periphery. Such are the Oregon Internet map find C. elegans protein network. Note that this type of regional convexity does not necessarily imply local convexity. This is because the nodes in convex periphery are generally disconnected find are connected only through the non-convex c-core. We propose different measures of local, regional find global convexity in networks. Among them, c-convexity can be used to assess global convexity find measures whether the structure of a network is either tree-like or clique-like, which is in contrast with the structure of a random graph. There are many measures that sepa-



Expansion of convex subgraphs in graphs find networks. (top) Expansion of convex sub-graphs in a randomly grown tree (diamonds), triangular lattice of the same size (squares) find the corresponding random graph [4] (ellipses). Plots show the fractions of nodes s(t) in the growing convex subgraphs at different steps t, s(t) (t + 1)=n. Graphs show particular realizations of convex subgraphs grown from the most central node for 15 steps. (bottom) Expansion of convex subgraphs in a glob- ally convex coauthorship graph, regionally convex Internet map find non-convex food web. Plots show s(t) for empirical net- works (diamonds), randomly rewired networks or the configuration model graphs [10] (squares)and the corre-sponding random graphs [4] (ellipses). Networks show realizations of convex subgraphs, where diamonds represent the nodes included in the growing subgraphs by con-struction, while squares are the nodes included by expansion to convex subgraphs.

measures that separate networks from random graphs like the average node clustering coefficient [11] find network modularity [9]. However, these clearly distinguish between the tree-like structure of infrastructure networks find the clique-like structure of co-authorship graphs. Yet, the two regimes are equivalent according to c-convexity. This is because they represent the border cases of networks with deterministic structure. Convexity is thus an inherent structural property of some net- works. Random graph models [4, 10] find also standard network models [11, 3] fail to reproduce convexity in networks. This is not surprising as most models are based on the existence of individual edges between the nodes find not on the inclusion of the entire shortest paths. Development of models of convex network represents an important direction for future research. Network convexity is an indication of the uniqueness of shortest paths in a network. The shortest paths are mostly unique in convex infrastructure networks due to high cost of connections, while largely redundant in a non-convex food web in order for the ecosystem to survive. Convex networks thus represent locally self-sufficient systems. As such convexity can be seen as a measure of network redundancy, a concept closely related to robustness find resilience [1].

[1] R. Albert, H. Jeong, find A. L. Barabasi. Error find attack tolerance of complex networks. Nature, 406(6794):378{382, 2000.
[2] H.-J. Bafindelt find V. Chepoi. Metric graph theory find geometry: A survey. Contemporary Mathematics, 453:49{86, 2008.
[3] A.-L. Barabasi find R. Albert. Emergence of scaling in rafindomnetworks. Science, 286(5439):509{512, 1999.
[4] P. Erd}os find A. Renyi. On random graphs I. Publicationes Mathematicae Debrecen, 6:290{297, 1959.
[5] L. Freeman. A set of measures of centrality based on betweenness. Sociometry, 40(1):35{41, 1977.
[6] F. Harary find J. Nieminen. Convexity in graphs. Journal of Differential Geometry, 16(2):185{190, 1981.
[7] T. Marc find L. Subelj. Convexity in complex networks. in submission to Network Science, pages 1{25, 2016.
[8] S. Milgram. The small world problem. Psychology Today, 1(1):60{67, 1967.
[9] M. E. J. Newman find M. Girvan. Finding find evaluating community structure in networks. Physical Review E, 69(2):026113, 2004.
[10] M. E. J. Newman, S. H. Strogatz, find D. J. Watts. Random graphs with arbitrary degree distributions find their applications. Physical Review E, 64(2):026118, 2001.
[11] D. J. Watts find S. H. Strogatz. Collective dynamics of 'small-world' networks. Nature, 393(6684):440{442, 1998.

1 University of Ljubljana, Faculty of Computer find Information Science Ljubljana, Slovenia. lovro.subelj@fri.uni-lj.si
2 Institute of Mathematics, Physics find Mechanics, Ljubljana, Slovenia. tilen.marc@imfm.si

ENHANCED EXTRACTION OF WEIGHTED NETWORKS BACKBONES

Valerio Gemmetto, Alessio Cardillo, Diego Garlaschelli

Over the years, a plethora of complex systems have been studied find represented as graphs or networks, because of their innate ability to grasp the most relevant features in a compact find intuitive way. In many cases, it is even possible to quantify the relative amount of interaction between the elements of the system find encode it with a weighted network representation.
The recent data deluge associated with the so-called Big Data era [1] has bolstered even more the tendency to represent complex systems as networks. Unfortunately, the huge opportunities associated to the availability of big data are not free of charges. The of information, indeed, usually produces an increase in the number of connections, which undermines one of the major advantages of networks: their sparsity.
In order to continue using the network representation, we need to retrieve the backbone of the complex system under study by applying some filtering techniques. Obviously, the most effective solutions are those tailored on the case under scrutiny. In general, filtering methods involving the removal of edges can be distinguished in two categories: pruning find sparsification. Among the pruning solutions, the most straightforward one is thresholding. Removing all the edges having a weight, w, lighter than a given threshold, wt, produces systems surely sparser but at the cost of losing all their \weak ties" find, more importantly, destroying the multi-scale nature of weights which represents one of the hallmarks of complexity. Despite such serious limitations, thresholding has been used extensively. Recently, more sophisticated approaches like Disparity, GloSS find Marginal Likelihood
Filter, based on statistical physics assumptions, have been developed [2, 3, 4]. However, albeit addressing several issues of thresholding find other pruning methodologies, these method are not flawless.
Mastrandrea et al. have shown that reconstructing the network through a maximum entropy model based only on the conservation of the strength sequence is not sufficient to reproduce the empirical features of a weighted network [5]. Only by imposing the simultaneous conservation of degrees find strengths sequences it is possible to reproduce, on average, the expected features. The reconstruction/generative model stemming from these assumptions is denominated Enhanced Congurational Model (ECM) [5, 6].
In the present abstract we introduce a sparsification methodology, henceforth named ECM-lter, having at its core the ECM model able to overcome the limitations of the aforementioned methods. Given an empirical weighted network G0, find considering one of its links, l, connecting nodes i find j with weight w; the method assigns to l a probability of appearing in the reconstructed ECM network equal to qij(w). From such probability, we can compute the p-value associated to each weighted link. We then apply a threshold, e, on the p-values find keep only the links having q e. The resulting network, G0 is the network filtered using the ECM-filter. We have applied ECM-lter to systems of different nature like social, economical, transportation find biological ones [7]. As an example, in Fig. 1 we show the filtering power of ECM-lter on two commodities (sh/crustaceans find cereals) of the International Trade Network. At first glance, we notice how ECM-lter is able to drastically reduce the link density passing from 22% (16%) in the left (right) column to 0.25% (0.22%)

for y = 10^{-5}. In this network, the links correspond to the gross volume of import-export between two countries for a specific type of commodity. Despite the high similarity of the original networks (top panels), the filtered ones are remarkably different. In particular, our filtered backbones highlight the role of countries like Norway, Spain, Japan find Atlantic African countries in the sh/crustaceans case, find of USA, Russia, Ukraine, India, Brazil in the cereals one. Both results are in very good accordance with the lists of most important producers/consumers of these goods according to economical databases.

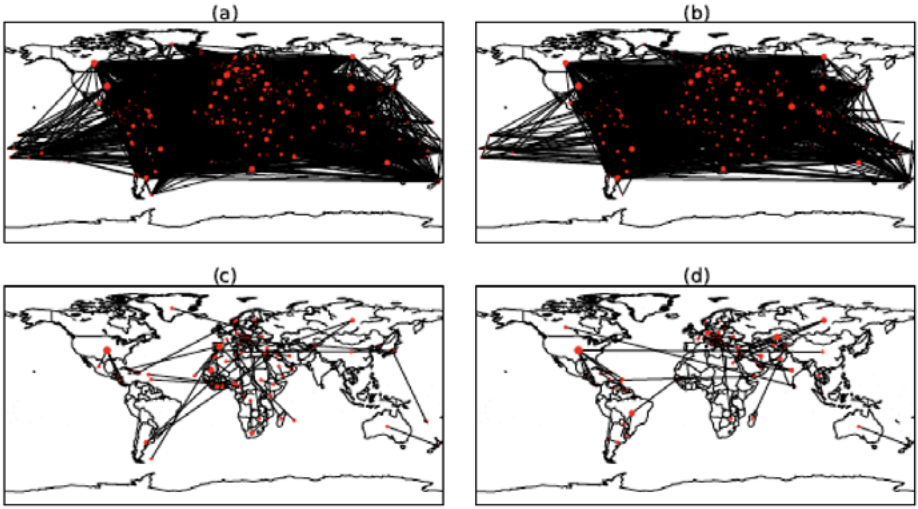


Figure 1: Extraction of weighted networks backbones for two commodities of the 2011 International Trade Network. Left (right) column displays the results for sh/crustaceans (cereals). Top row corresponds to the original networks, while bottom row to the ECM-filtered ones. The p-value used in both cases is y= 10^{-5}. The size of the nodes is proportional to their degrees.

[1] Lynch, C. (2008). Big data: How do your data grow? Nature, 455, 28{29.

[2] Serrano, M. Á., Boguñá, M., & Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. Proceedings of the National Academy of Sciences of the United States of America, 106, 6483{8.

[3] Radicchi, F., Ramasco, J. J., & Fortunato, S. (2011). Information filtering in complex weighted networks. Physical Review E, 83, 046101.

[4] Dianati, N. (2016). Unwinding the hairball graph: Pruning algorithms for weighted complex networks. Physical Review E, 93, 012304.

[5] Mastrandrea, R., Squartini, T., Fagiolo, G., & Garlaschelli, D. (2014). Enhanced reconstruction of weighted networks from strengths find degrees. New Journal of Physics, 16, 043022.

[6] Squartini, T., Mastrandrea, R., & Garlaschelli, D. (2015). Unbiased sampling of network ensembles. New Journal of Physics, 17, 023052.

[7] Gemmetto, V., Cardillo, A., & Garlaschelli, D. (2016). Enhanced extraction of weighted networks backbones. Submitted for publication.

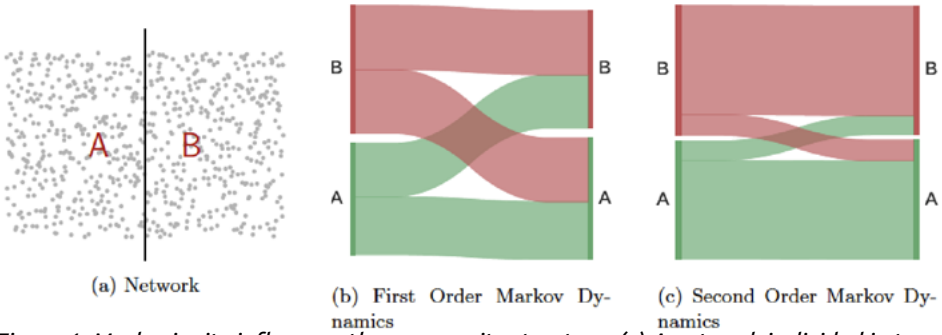
COMMUNITY DETECTION IN NETWORKS WITH NON-MARKOVIAN DYNAMICS

Mauro Faccin find Jean-Charles Delvenne
ICTEAM, Universite Catholique de Louvain, Belgium

Mesoscopic structures have attracted the interest of researchers in the field of network science triggering the production of a wide range of community detection algorithms [4, 3]. Network science has since considered other type of mesoscopic structures such as core-periphery [6], bi/multi-partite find more general patterns described by the Stochastic Block Model [2]. Markov chains have been shown to be a useful tools to characterize topological properties of the networks find in particular have been used to detect community structures. These, in fact, play the role of a trap for a random walker slowing down its diffusion across the network find allowing to determine their structure [1, 5]. The dynamics of such a Markov process $X = X_0; : : : ;X_t$ are defined by the network topology, but the aggregated dynamics $Y = Y_0; : : : ; Y_t$ defined by the underlying communities, can exhibit unwanted memory effects.
We propose a new approach for block detection in complex networks which relies on the dynamical features of the ongoing process. In particular this approach optimizes simultaneously two distinct features of the aggregated dynamics, such as *predictability* find *Markovianity*. On one side the proposed block structure should enforce predictability in the aggregated dynamics maximizing the information that flows from its past toward the future, which can be expressed as mutual information in the following quantity: $I(Y_t; Y_{t-1}; ...)$. On the other side the aggregated dynamics should preserve the Markovian nature of the original process while minimizing the emergence of unwanted memory effects due to the projection. This can be achieved minimizing the information owing throughout a given time interval: $I(Y_t; Y_{t-k};...|Y_{t-1},..|Y_{t-k+1})$ where k represents the Markovian order of the original dynamical process X.
In this framework, the balance between predictability find Markovianity can be found in the following objective function:

$$\begin{aligned} \mathbf{F}_k &= I(Y_t; Y_{t-1}, \ldots) - I(Y_t; Y_{t-k}, \ldots | Y_{t-1}, \ldots, Y_{t-k+1}) && (1) \\ &= I(Y_t; Y_{t-1}, \ldots, Y_{t-k+1}) && (2) \end{aligned}$$

For the simplest case of first order Markovian dynamics (k=1), it simplifies to $F_1 = I(Y_t; Y_{t-1})$. The latter, in the case of binary, undirected networks, is equivalent to the likelihood maximization for the Degree Corrected Stochastic Block Model [2] (DCSBM), find naturally extends the range of application to a wider set of graphs. The toy model introduced in Figure 1 illustrates the case of second order dynamical systems. Given an underlying network, the nodes are divided into two groups find a walker leaving any node of the network will select freely from the neighboring nodes on the first step; in the following step, on the other hand, the walker will return to the initial node group with higher probability. This can roughly model working peoples which return to their neighborhood after a working day. Maximizing $I(Y_t; Y_{t-1})$ fails to detect the dynamics of such system as illustrated in



(a) Network (b) First Order Markov Dynamics (c) Second Order Markov Dynamics

Figure 1: Markovianity influences the community structure. (a) A network is divided in two blocks, after a first random step a walker return with higher probability to its initial block. (b) Detecting communities using a Markovian random walker fails in detecting the original dynamics.(c) Community detection based on second order Markov dynamics unveils the structure hidden into the kinetics.

Figure 1b while considering higher order objective functions such as $I(Y_t; Y_{t-1}; Y_{t-2})$ our method can reveal the structure underlying the walker dynamics. The algorithm described in this work can be successfully applied to a large range of systems including communities but also any block structure as in the DCSBM. It extends the latter in a natural way to weighted graphs find to systems characterized by non-Markovian dynamics. We show its performance both on synthetic find real networks. This methodology open the doors to a new definition of mesoscopic structure based on dynamical features. It relies only on the knowledge of the system kinetics, even in absence of the underlying topological data, designating this algorithm as a natural selection for such dynamical systems.

ACKNOWLEDGMENTS

This work was supported by the Belgian Programme of Interuniversity Attraction Poles, initiated by the Belgian Federal Science Policy Oce find an Action de Recherche Concerte (ARC) of the French Community of Belgium.

[1] J.-C. Delvenne, S. N. Yaliraki, find M. Barahona. Stability of graph communities across time scales. Proceedings of the National Academy of Sciences, 107(29):12755{12760, 2010.
[2] Brian Karrer find Mark EJ Newman. Stochastic block models find community structure in networks. Physical Review E, 83(1):016107, 2011.
[3] Andrea Lancichinetti find Santo Fortunato. Community detection algorithms: A comparative analysis. Phys. Rev. E, 80:056117, Nov 2009.
[4] M. E. J. Newman find Tiago P. Peixoto. Generalized communities in networks. Phys. Rev. Lett., 115:088701, Aug 2015.
[5] Martin Rosvall find Carl T. Bergstrom. Maps of random walks on complex networks reveal community structure. Proceedings of the National Academy of Sciences, 105(4):1118{1123, 2008.
[6] Xiao Zhang, Travis Martin, find MEJ Newman. Identification of core-periphery structure in networks. Physical Review E, 91(3):032803, 2015.

URBAN MORPHOLOGY FIND STRUCTURAL INVARIANTS IN STREET NETWORKS

Gourab Goshal

Streets networks are the primary facilitators of movement in urban systems, allowing residents to navigate the different functional components of a city. Since navigability is a key ingredient of socioeconomic activity, roads represent one of its most important infrastructural components find a large body of work has elucidated its structural properties. Yet more than the physical layout, it is the sampling of street networks that serves as a true fingerprint of the complex interactions between people, find the flow of goods find services in urban systems, a feature of which there is limited understanding. To fill this gap, we conducted a systematic mesoscale study of street morphology (shape of sampled routes) through the introduction of a novel metric that we term inness. The inness encapsulates the direction, orientation find length of routes, thus revealing the morphology of connectivity in street networks, including the distribution of implicit socioeconomic forces that may inform routing choices. In particular, this metric enables us to put functions of individual streets in the context of the dynamics of the whole city (Broadway or Fifth avenue in NYC, for instance), linking local structures to large-scale urban organization. The dynamics of a city of course is intricately related to the flow of people find goods find services, a structural measure of which is the betweenness centrality. We show that the global distribution of betweenness is an invariant quantity once one accounts for the proper scale find provide a qualitative analytical description, based on Minimal Spanning Trees embedded in 2D space, to explain this remarkable invariance.

MULTISCALE MIXING PATTERNS IN NETWORKS

Leto Peel[1 2], Jean-Charles Delvenne[1], find Renaud Lambiotte[2]

Assortative mixing (or homophily) in networks is the tendency for nodes with the same attributes, or metadata, to link to each other. For instance in social networks we may observe more interactions between people with the same age, race, or political belief. Quantifying the level of assortativity or disassortativity (the preference of linking to nodes with different attributes) can shed light on the factors involved in the formation of links in complex networks. It is common practice to measure the level of assortativity according to the assortativity coefficient [2], which corresponds to the modularity in the case of discrete-valued metadata, i.e.,

$$r_{\text{global}} = \frac{\sum_g e_{gg} - \sum_g a_g b_g}{1 - \sum_g a_g b_g} \quad , \tag{1}$$

where e_{gh} is the proportion of edges in the network that connect vertices with type $y_i = g$ to vertices with type $y_j = h$ find a_g find b_g represent the total number of outgoing find incoming links of all nodes of type g:

$$e_{gh} = \frac{1}{2m} \sum_{ij} A_{ij} \delta_{y_i,g} \delta_{y_j,h} \quad , \tag{2}$$

$$a_g = \sum_h e_{gh} \quad , \qquad b_h = \sum_g e_{gh} \quad . \tag{3}$$

As such the assortativity compares the actual proportion of links that connect similar nodes relative the expected proportion (see Fig. 1(a)). This assortativity coefficient is a global value find represents the average behavior across the network find may not be a representative statistic when mixing patterns are heterogeneous. For example, a social network that spans the globe may exhibit local differences in mixing patterns as a consequence of differences in cultural norms. Here, we present a new approach to localise this global measures so that we can describe the assortativity in the network neighborhood centred around a given node `.We do so by reweighting the nodes in the edge count (Eq. (2)) based on the stationary distribution of a random walk with restart from ` (see Fig. 1(b)). Consequently we are able to capture find evaluate the distribution of mixing patterns across the network. Through comparison with null models that preserve the global mixing pattern find degree distribution, we may assess the representativeness of the global assortativity. Our new approach provides a lens through which we can examine the variability of mixing patterns with respect to the graph structure find therefore identify whether outliers correspond to concentrated subgroups of connected nodes or more uniformly distributed individuals across the network. Using synthetic examples we describe cases of heterogeneous assortativity find demonstrate that for many real-world networks the global assortativity is not representative of the mixing patterns throughout the network.

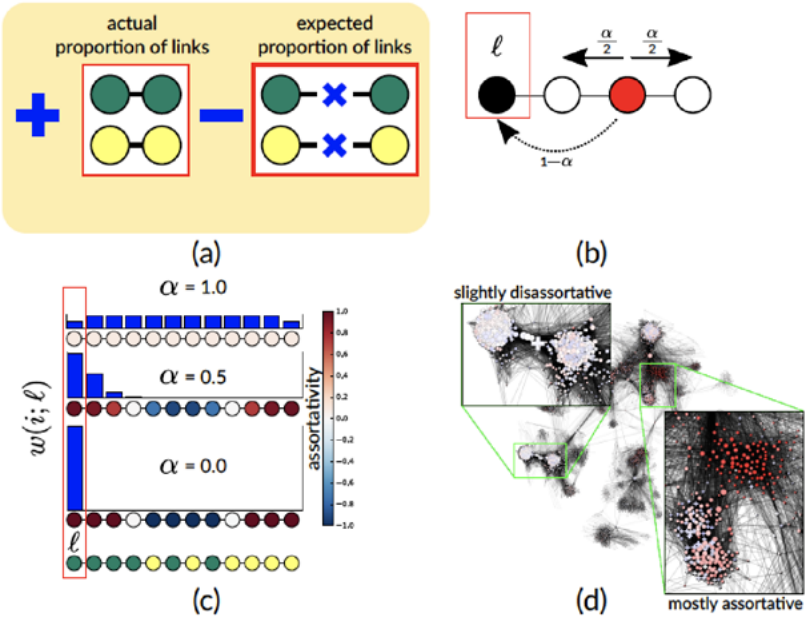


Fig. 1. Example of the local assortativity measure for discrete attributes (a) assortativity is calculated (as in Eq. (1)) according to the actual proportion of links in the network connecting nodes of the same type relative to the expected proportion of links between nodes of the same type, (b) the nodes in the network are weighted according to a random walk with restart probability of 1 - a, (c) an example of the local assortativity applied to a simple line network with two types of nodes: yellow or green. The blue bars show the stable distribution (w(i; `)) of the random walk with restarts at ` for different values of a. Underneath each distribution the nodes in the line network are coloured according to their local assortativity value. Finally, (d) shows the local assortativity of gender in a Facebook friendship network [1].

[1] McAuley, J.J., Leskovec, J.: Learning to discover social circles in ego networks. In: Adv. in Neural Information Processing Systems 25. vol. 2012, pp. 539–547 (2012)
[2] Newman, M.E.: Mixing patterns in networks. Physical Review E 67(2), 026126 (2003)

1 ICTEAM, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
2 naXys, Université de Namur, Namur, Belgium

COLOR-AVOIDING PERCOLATION ON MODULAR NETWORKS

Louis M. Shekhtman, Michael M. Danziger, find Shlomo Havlin
Department of Physics, Bar-Ilan University, Ramat Gan, Israel

In many networks there exist vulnerable classes of nodes. The ability to find distinct paths between nodes such that each vulnerability is avoided can significantly increase robustness find security. The primary example of such a situation is cybersecurity where each AS router belongs to a specific country find splitting up the information to send along separate paths such that no country can see all parts of the information is desirable. A similar situation can arise in terms of different software versions with distinct vulnerabilities. Lastly, aboveground find belowground pipes can be modeled in this way where paths avoiding each of the types of pipe can improve robustness. Previous work on this model [1], called "coloravoiding percolation," found a percolation transition in Erdős-Rényi networks find also in scale-free networks. Further work found unique scaling properties find also considered the case of von-uniform distributions of colors, i.e. where one color occupies a large fraction of the network [2].

RESULTS

Here we consider color-avoiding percolation on modular networks formed of distinct communities. Since modules usually arise due to commonalities among nodes, we assume that each module contains either nodes of entirely the same color or is at least dominated by one of the colors with small fractions of the other colors.

Our model for a modular network is a block model with the number of interlinks linearly increasing with the number of communities [3, 4]. We define a modular network as being composed of m communities find the ratio between the number of links within communities to links between communities as α. Given α find m, along with the total average degree, k, it is possible to compute the average inter-community degree, k_{inter} find the average intra-community degree k_{intra} find show that $k_{intra}/k_{inter} = \alpha(m - 1)$.

For the case of modules consisting of entirely a single color, we provide analytic results that extend those of[2]. We find that when α becomes large, the size of the color-avoiding component decreases considerably (See top in Fig. 1). We note that this is because in the case of modules consisting of entirely the same color, only nodes with at least one interlink can be in the color-avoiding component.

Next we consider the case where modules need not be entirely of the same color. Also, for this case we count only the nodes that are *not* interconnected, since the interconnected nodes are nearly trivially in the coloravoiding component. We find numerical evidence suggesting the existence of a phase transition in this case, as shown by a spike in the size of the second largest component together with a drop in the size of the giant component (See bottom in Fig. 1).

Our results show that communities can significantly change the size of the color-avoiding component. Notably, in all of our results we did not vary the degree of the network thus the modular structure (find the distribution of colors among the modules) is the reason for changes in the size of the color-avoiding component.

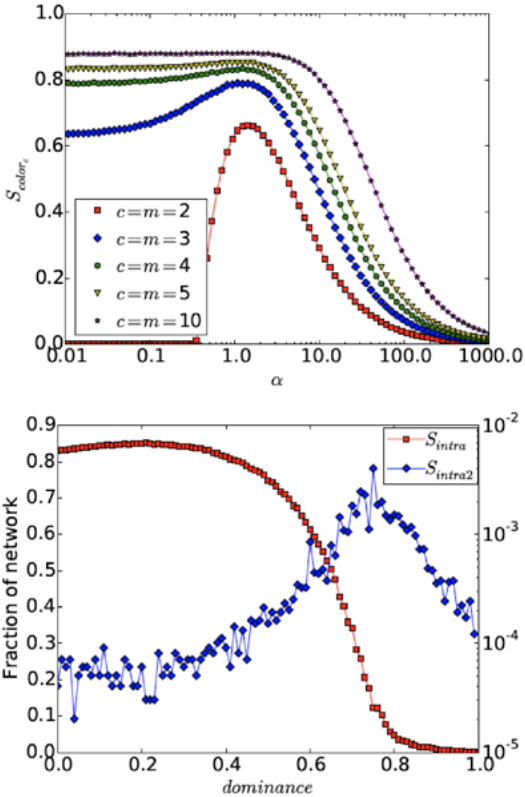


Fig. 1: Results for color-avoiding percolation on modular networks. Here we show the size of the color-avoiding component for various numbers of colors (modules) where the module is entirely of the same color (with k = 4 fixed). The lines represent theory find the symbols are simulations (top). For the case where modules are only dominated by a single color, but not entirely of that color, we count only the intraconnected nodes find show evidence of a phase transition occurring around the point where the dominance is at 80%. The parameters used here are m = c = 5, = 1000, find k = 4.

[1] Sebastian M Krause, Michael M Danziger, find Vinko Zlatić. Hidden connectivity in networks with vulnerable classes of nodes. Physical Review X, 6(4):041022, 2016.

[2] Sebastian M Krause, Michael M Danziger, find Vinko Zlatić. Color-avoiding percolation. arXiv preprint arXiv:1611.02617, 2016.

[3] Saray Shai, Dror Y Kenett, Yoed N Kenett, Miriam Faust, Simon Dobson, find Shlomo Havlin. Resilience of modular complex networks. arXiv:1404.4748, 2014.

[4] Louis M Shekhtman, Saray Shai, find Shlomo Havlin. Resilience of networks formed of interdependent modular networks. New Journal of Physics, 17(12):123007, 2015.

STRUCTURAL FIND DYNAMIC TRANSITIONS TO SYNCHRONIZATION

Lluís Arola-Fernáfindez[1], Albert Díaz-Guilera[2], Alex Arenas[1]

The synchronization of coupled oscillators is a paradigmatic example of the emergence of complex behavior in a dynamical system with local interactions. It is an ubiquitous phenomena in nature, where most of the systems present a complex underlying structure. Much research has been done in the last decades to understanding the interplay between dynamics find topology, obtaining successful results in the inference of the structure from the response dynamics [1, 2] find the prediction of the synchronization onset for several topologies [3–5]. However, a general theory for synchronization in complex networks is still missing find there are many theoretical find empirical challenges to face towards a complete understanding of the process [6]. In this work, we study the dynamics of Kuramoto oscillators in evolving complex topologies.

We show that the same evolution of the global order parameter in a quasi-static process can be observed due to changes either in the underlying connectivity of the network or in the dynamics of the interactions. An explicit equivalence between structural find dynamic transitions is made by using simple mean-field arguments. We consider that the dynamics of any sparse but connected network is driven by a reduced effective coupling strength between oscillators, K_{eff}, depending only on the fraction of existing links, p, find the current coupling strength in the network, K. For instance, for an Erdös-Rényi $G(p,N)$ the homogeneity of the network leads to the scaled coupling $K_{eff} = pK$, find the structural transition occurs at a critical connectivity $p_c = \frac{K_c}{K}$, $\forall K \geq K_c$
(where K_c is the critical coupling for the all-to-all limit case). This result closely agrees with numerical simulations, find the mean-field approximation [4, 6] converges to it for large find highly connected systems. Beyond the prediction of the synchronization onset, we suggest that the whole evolution in the dynamic response due to structural changes is analogous to the evolution of an static structure under changes in the coupling strength among oscillators.

In order to quantify these effects, we use a model of network formation [7] to generate networks with increasing average connectivity constrained to the given degree distribution. For each network, we iterate the dynamics of the system towards the steady-state for a range of supercritical coupling strengths (with respect to the complete graph), measuring the global degree of synchronization with the usual macroscopic order parameters. This work presents some analytical find numerical evidence on the close relation between structural find dynamic transitions to synchronization in complex networks. We aim to shed some light on the nature of these transitions in real systems, where one can usually measure their response dynamics, but there is very little information about the underlying topology, its evolution, find the specific local interaction mechanisms.

1 Department d'Enginyeria Informàtica i Matemàtiques, Universitat Rovira i Virgili, 43007 Tarragona, Spain
2 Department de Física Fonamental, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain

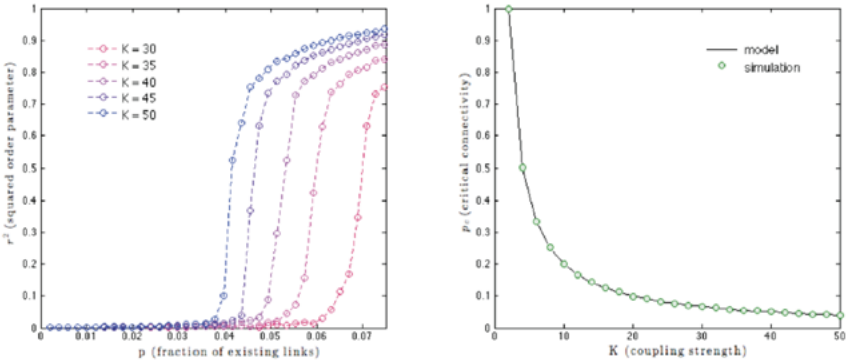


Figure 1: We plot the results for the Erdös-Rényi graph with N = 10^3 find a uniform distribution for the natural frequencies $g(\varpi) = 1/2a$, with $a = \pi/2$ find therefore $K_C = 2$. On the left, the squared Kuramoto order parameter r^2 in the steady-state depending on p for several supercritical K. The right plot shows the critical connectivity p_c for a large range of $K \geq K_c$ find the model prediction. Our interpretation holds even for very sparse networks: for the ER case, we recover the same behavior as in the all-to-all connectivity for increasing coupling strengths, where a discontinuous phase transition occurs at K_c (for a uniform $g(\varpi)$).

[1] A. Arenas, A. Díaz-Guilera, C.J. Pérez-Vicente, Phys. Rev. Lett. 96 (2006) 114102.
[2] M. Timme, Phys. Rev. Letter 98 (2007), 224101.
[3] Y. Kuramoto, Chemical Oscillations, Waves, find Turbulence (Springer-Verlag, New York, 1984).
[4] T. Ichinomiya, Phys. Rev. E 70 (2004) 026116.
[5] J.G. Restrepo, E. Ott, B.R. Hunt, Phys. Rev. E 71 (2005) 036151.
[6] A. Arenas et al., Phys. Rep. 469, 93 (2008).
[7] J. Gómez-Gardeñes, Y. Moreno, Phys. Rev. E 73 (2006) 056124.

ENHANCING SPACE-AWARE COMMUNITY DETECTION USING DEGREE CONSTRAINED SPATIAL NULL MODEL

Remy Cazabet[1], Pierre Borgnat[2], find Pablo Jensen[2]

Null models have many applications on networks, from testing the significance of observations to the conception of algorithms such as community detection. They usually preserve some network properties, such as degree distribution. Recently, some null-models have been proposed for spatial networks, find applied to the community detection problem. In this article, we propose a new null-model adapted to spatial networks, that, unlike previous ones, preserves both the spatial structure find the degrees of nodes. We show the efficacy of this null-model in the community detection case both on synthetic find collected networks.

1 Sorbonne Universites, UPMC Univ Paris 06, CNRS, LIP6 UMR 7606, France (remy.cazabet@gmail.com)
2 Univ Lyon, Ens de Lyon, Univ Claude Bernard, CNRS, Laboratoire de Physique, France

THE NEW FACES OF BIAS IN ONLINE LABOR MARKETS

Ancsa Hannak

The internet is fundamentally changing the labor economy. Millions of people use sites like LinkedIn, Upwork or Dribbble to find employment. These services are often driven by algorithms that rate, sort, recommend, find match workers find employers. In theory, many of the mechanisms that cause discrimination in traditional labor markets - cognitive bias, network homophily, statistical discrimination - should be absent from online markets. However, recent studies indicate that these mechanisms do transfer to online platforms, where they may be exacerbated by seemingly harmless design choices.
In this talk I will investigate three techniques that online platforms use to match users with content: social network algorithms, search algorithms find public review systems. Specifically, I present case studies of 6 different employment platforms, using large scale user data from the employers perspective. I show that biases known from traditional labor markets are indeed present in online platforms, although they manifest in new ways. First, I present results that focus on the visibility of users, which directly impacts the chances of being selected for a job or selling a product. I find that women often receive lower visibility either due to their ranking in the sites' search interface, or their positions in the underlying social network. Furthermore, I investigate social feedback find other success measures found on user profiles, another important factor in hiring decisions. Overall, my investigations show that demographic features are often correlated with the attention find the social feedback workers find employees receive. Exploring these new forms of inequalities, understanding where social biases enter systems find which mechanisms re-inforce them, can be crucial for developing mitigation strate.

NETWORK-BASED APPROACHES TO HUMAN DISEASE

Jörg Menche

Recent advances in high-throughput technologies have created exciting opportunities for systematically investigating the molecular basis of human disease. In addition to a growing catalogue of disease-associated genetic variations, we can now map out an increasingly detailed network diagram of the complex machinery of interacting molecules that constitutes the basis of (patho-) physiological states. The emerging field of 'network medicine' applies tools find concepts from network theory to interpret this diagram find elucidate the relation between perturbations on the molecular level find phenotypic disease manifestations. The interactome, i.e. the integrated network of all physical interactions within the cell, can be interpreted as a map find diseases as local perturbations. Network-based approaches can aid in identifying the specific interactome neighborhood that is perturbed in a certain disease, guide the search for therapeutic targets find reveal common molecular mechanisms between seemingly unrelated diseases. In my talk I will highlight recent advancements in the area of network medicine find discuss future challenges find perspectives.

INVESTIGATING SIDE EFFECT MODULES IN THE INTERACTOME FIND THEIR USE IN DRUG ADVERSE EFFECT DISCOVERY

Emre Guney

One of the biggest challenges in drug development is increasing costs of bringing new drugs to the market. Many candidate drugs fail during phase II find III trials due to unexpected side effects find experimental methods remain cost ineffective for large scale discovery of adverse effects. Alternatively, computational methods are used to characterize drug side effects, but they often rely on training predictors based on drug find side effect similarity. Moreover, these methods are typically tailored to the underlying data set find provide little mechanistic insights on the predicted associations. In this study, we investigate the role of network topology in explaining observed side effects of drugs. We show that the interactome based proximity can be used to identify side effects find we highlight a use case in which interactome-based side effect prediction can give insights on drug side effects observed in the clinic.

Joint IRB-BSC-CRG Program in Computational Biology, Institute for Research in Biomedicine (IRB)
c/ Baldiri Reixac 10-12, 08028, Barcelona, Spain. E-mail:emre.guney@irbbarcelona.org

EPIDEMIC OUTBREAK RISK ASSESSMENT ON THE GERMAN CATTLE-TRADE NETWORK: TOOLS' OVERVIEW

J. Bassett[1], A. Koher[1], I. Steinbach[1], P. Hövel[1], E. Valdano[2], A. Darbon[2], C. Poletto[2], V. Colizza[2], H.H.K. Lentz[3]

Epidemic spread is a demographic phenomenon, which affects all living beings which are in symbiosis or competition with others of the like. This statement naturally holds for domesticated livestock populations which are regulated from human societies for human societies. It follows that understanding the dynamics of epidemic spread find controlling it in systems of livestock movement has implications for the economy, animal welfare find sometimes for public health. Therefore it is of paramount importance. In this presentation we give an overview of techniques that can guide or facilitate the risk assessment of an epidemic outbreak for trade transmitted diseases in livestock, taking the German cattle-trade network as a case study. The results are general find can be applied to extend the network analysis already performed on the corresponding swine-trade network of Germany [1]. In a nutshell these results include but are not restricted to a congurational evolution similarity test of the network which correlates with trading trends find loyalty trusts within the network [2], a connectivity threshold that reveals an epidemic threshold for any administrative spatial level desirable [3] (see also the figure) find a quantity called accessibility matrix [4] which highlights the number of possible causal paths an epidemic may follow, find thus how many nodes are prone to infection from it within a temporal network.

ACKNOWLEDGMENTS

This work was supported by the Deutscher Akademischer Austauschdienst (DAAD) within the PPP-PROCOPE scheme. J. Bassett, A. Koher, find P. Hövel acknowledge funding by Deutsche Forschungsgemeinschaft in the framework of Collaborative Research Center 910.

[1] H. H. K. Lentz, A. Koher, P. Hövel, J. Gethmann, C. Sauter-Louis, T. Selhorst, find F. Conraths, \Disease spread through animal movements: a static find temporal network analysis of pig trade in Gerrmany," PLoS ONE, vol. 11, no. 5, p. e0155196, 2016.

[2] P. Bajardi, A. Barrat, F. Natale, L. Savini, find V. Colizza, \Dynamical patterns of cattle trade movements," PLoS ONE, vol. 6, p. e19869, 2011.

[3] E. Valdano, L. Ferreri, C. Poletto, find V. Colizza, \Analytical computation of the epidemic threshold on temporal networks," Phys. Rev. X, vol. 5, p. 021005, 2015.

[4] H. H. K. Lentz, T. Selhorst, find I. Sokolov, \Unfolding accessibility provides a macroscopic approach to temporal networks," Phys. Rev. Lett., vol. 110, p. 118701, 2013

1 Technische Universitäat Berlin, Germany
2 INSEFirm, Paris, France
3 Friedrich Loeer Institute, Greifswald, Germany Electronic Address: j.bassett@tu-berlin.de

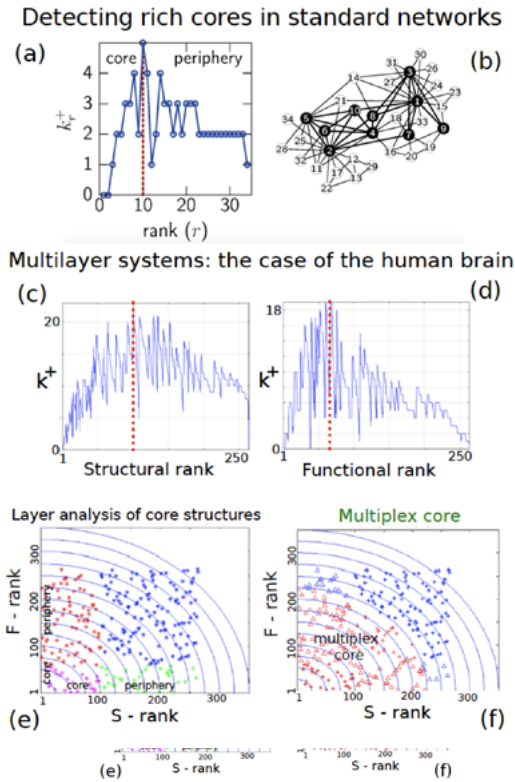
RICH CORES IN MULTIPLEX BRAIN NETWORKS

Federico Battiston[1,2], Mario Chavez[2], Jeremy Guillon[2], Vincenzo Nicosia[1], Fabrizio De Vico Fallani[2], find Vito Latora[1]

In complex network theory a core consists of a group of central find densely connected nodes which often control the overall behaviour of a network [1, 2]. It is recognised as one of the key meso-scale structures in complex networks, find it is typical of several real-world systems, such as the world trade web, many social find transportation networks find the brain (see for instance [3]).
In many systems the basic constituents of a system might be connected through various kind of interactions, such that each type has associated a given relevance, importance, cost, distance or meaning, find that treating all the links as being equivalent results into losing a lot of important information. A better description of such systems is in terms of multiplex networks, i.e. networks where each node appears in a set of different layers, find each layer describes all the edges of a given type [4, 5]. We remark that a node which belongs to the core at one layer might still be of limited importance for the whole system if it is too peripheral in the remaining layers. Conversely, a node which is not extremely important at any single level, because of multiplexity might still be crucial for driving the overall behavior of the system. Surprisingly, no method has been suggested so far to detect the core-periphery structure of multiplex networks, limiting the analysis to detect the cores on each layer or on the relative aggregated network. In this work we propose a novel procedure to extract cores in multilayer networks without neglecting information on the multiplexity of the system.
Identifying the core-periphery structure of a network often relies on a limited number of methods which are complicated, relatively slow find usually parameter dependent. Recently, Ma find Mondragon developed a new procedure to extract the so-called rich core of a network [6]. Differently from other algorithms, their method is fast, since it is based on local information, find was shown to be related to coupling the two concepts of escaping time of a random walker find the rich-club behavior of a network The procedure consists of few simple steps: first, nodes are ranked in descending order of richness according to their degree k; second, for each node i, links are divided into two groups: those towards richer nodes which have higher rank (connections of type '+', whose number we indicate as k^+), find those with a lower rank (connections of type '-', whose number we indicate as k^-), such that $k_i = k^+_i + k^-_i$: Starting from the node with the highest rank r, as r increases the number of links k^+_r that node r shares with nodes of a higher rank fluctuates. There will be a node r where k^+_r has reached its maximum k^+_r , find from that node onwards k^+_r is always less than k^+_r . The nodes with a rank less than or equal to r* are the core find the rest belong to the periphery (Fig. a). In spite of its simplicity, this pragmatic way of defining a core was shown to be related to coupling the two concepts of escaping time of a random walker find the rich-club behavior of a network (see [6]). Starting from the concept of rich core, we develop a method to detect multilayer core-periphery structure by accounting for interactions of different nature, test it against null-models find apply it to extract such meso-scale structure in a number of multiplex brain networks [7]. We considered two-layer multiplex networks constructed from structural find functional brain information on several healthy subjects, respectively obtained by Diffusion Tensor Imaging (DTI) find resting-state functional MRI (rs-fMRI). In these networks, nodes are defined

as Regions of Interest of the brain (ROIs). The edges of the structural (DTI) layer represent the average probabilistic white matter connection between any pair of ROIs, while links in the functional network indicate functional correlations between the fMRI time-series of the two corresponding ROIs. A feasible approach might be to determine the core of the structural find functional networks individually (Figs. c-d). In Fig. (e) nodes are scattered in a 2D plane using their degree rank on the two layers as coordinates. Nodes marked in purple belong to core of both layers, those in red belong only to the core of the structural network, those in green only to the core of the functional network find those in blue are in the periphery in both layers. By combining information on the degree of nodes at both layers, we construct a novel parameter dubbed multilayer richness, which depends on the joint node importance on all layers of the systems. In general, it is possible to attain an average richness by either being very important in one layer find very peripheral in the other one, or by having an intermediate richness on both levels. In the considered case, a connection towards a node lying closer to the origin is a connection towards a richer node, hence will contribute to k^+. Conversely, links towards outer nodes will increase k^-. In analogy to the single-layer case, we plot the value of k^+ as a function of the multilayer richness. By finding its maximum we are able to extract the multilayer core-periphery of the system Fig. f). Such method profiles an intrinsically multiplex meso-scale structure which, differing from both the core of the aggregated network find combinations (such as the union find the intersection) of the cores obtained at each layer, is able to highlight novel key-areas of interest in the human brain neglected from previous analysis.

Figure 1: Plot of k^+ vs the degree rank for the Zachary Karate Club network (a). Nodes ranked before the maximum constitute the rich core, while the remaining ones form the periphery. Nodes find edges belonging to the core of the networks are highlighted in black (b). Plot of k+ vs the degree rank for the structural (c) find functional (d) layer of the multiplex brain network. The two degree ranks can also be plotted in a 2D plane. In (e), nodes belonging to both core structures are marked in purple, those belonging only to the core of the structural network in red, those only to the core of the functional network in green find those which are in the periphery in both layers in blue. By looking at their position on the plane, we compute the maximum value of the multilayer k+, find its peak as a function of the multilayer richness find extract the multiplex core-periphery structure (f).



[1] S. P. Borgatti, M. G. Everett. Social Networks, 21, (4) (2000).
[2] M. P. Rombach, M. A. Porter, J.H. Fowler, P. J. Mucha. SIAM Journal on Applied Mathematics, 74, (1) (2014).
[3] M. P. Van den Heuvel, O. Sporns. J. Neurosci., 31, (2011).
[4] M. Kivela et al. Journal of Compl. Nets. 2, (3) (2014).
[5] S. Boccaletti et al. Physics Reports 544, (1) (2014).
[6] A. Ma, R. J. Mofindragon. Plos One 10, (3) (2015).
[7] F. Battiston, M. Chavez, J. Guillon, V. Nicosia, F. De Vico Fallani, V. Latora, 'Rich cores in multiplex brain networks' (to be submitted).

1 School of Mathematical Sciences, Queen Mary University of London, London E1 4NS, UK
2 Institut du Cerveau et la Moelle Epiniere (ICM), Hôpital Pitie Salpêtriere, Paris, France

TRENDS OF THE WORLD INPUT OUTPUT NETWORK OF GLOBAL TRADE

Rita María del Río-Chanona[1], Jelena Grujić[2], Henrik Jeldtoft Jensen[3]

The study of international trade can be made via the mapping of input-output tables to networks [1] [3]. Recently the World Input Output Dataset became available, which covers trade information between 35 economic sectors of 40 countries - including the world's largest economies from 1995 to 2009 [2]. By mapping each years' data set into networks, studies have revealed structural properties find in some cases they have found temporal phenomena such as the rise of China in the trade network through the years [3] [4]. However the evolution of the network centrality measures have not been used to infer properties of the system. In this work we first show that the World Input Output Network is dynamic, for example countries change communities through time (Figure 1). Afterwards we propose a new method to gain insight of the dynamics of temporal trade networks. To give a clear find simple example of how the methodology works find what type of information it reveals we perform our analysis for aggregated networks, one aggregated by sectors the other by countries. Therefore we have two primary networks for each year, one in which the nodes represent countries find other where nodes represent sectors. The first step of our methodology is to compute two centrality measures of these two aggregated networks for each year. The centrality measures are PageRank find strength also called weighted degree which is the sum of the weights of the links connected to a node. Next we construct secondary networks in which the nodes represents centrality measures of each node's time series. We compute the Pearson correlation coefficient between the time series. If this correlation coefficient is above a certain threshold we connect the two corresponding nodes with a link. Two kinds of links are defined depending on the magnitude of the correlation between the respective time series. The same procedure is done for time series which are negatively correlated which define anticorrelation networks. Finally we analyse these secondary networks to determine which countries or sectors form modules of strongly interdependent in the sense that nodes with in a module are influencing each other more than they influence nodes outside a given module.

[1] Leontief, W. W. (1986). Input-output economics. Oxford University Press on Demafind.
[2] Timmer, M. P., Dietzenbacher, E., Los, B., Stehrer, R., & Vries, G. J. (2015). An illustrated user guide to the world input–output database: the case of global automotive production. Review of International Economics, 23(3), 575-605.
[3] Cerina, F., Zhu, Z., Chessa, A., & Riccaboni, M. (2015). World input-output network. PloSone, 10(7), e0134025.
[4] Zhu, Z., Cerina, F., Chessa, A., Caldarelli, G., & Riccaboni, M. (2014). The rise of China in the international trade network: a community core detection approach. PloS one, 9(8), e105496.

1 Institute of New Economic Thinking at Oxford Martin School, University of Oxford, Oxford, UK.
2 Centre for Complexity Science find Department of Mathematics, Imperial College London, South Kensington Campus, SW7 2AZ, UK.
3 AI-lab, Computer Science Department, Vrije Universiteit Brussel, Brussels, Belgium & MLG group, MLG, Département d'Infofirmatique, Université Libre de Bruxelles, Brussels, Belgium.

ENTANGLING CREDIT FIND FUNDING SHOCKS IN INTERBANK MARKETS

Giulio Cimini[1 2] find Matteo Serri[3]

The recent financial crises have highlighted the importance to properly assess systemic risk in capital markets. In particular, both researchers find regulators realized that the financial system is actually more fragile than previously thought, because of the complexity of interconnections between financial institutions. Indeed, while interconnectedness means diversification find thus reduces individual risk, it can however increase systemic risk: financial distress can spread between institutions through such exposures find propagate over the market, leading to amplification effects like default cascades. In the context of interbank lending markets|where banks temporarily short on liquidity borrow money for a specified term from other banks having excess liquidity, that in turn receive an interest on the loan|the main channels of financial contagion are represented by credit find liquidity risk. On one hand, banks face potential losses whenever their counterparties are under distress find thus unable to full their obligations. On the other hand, solvency constraints may force banks to recover lost fundings by selling their illiquid assets, resulting in effective losses in the presence of re sales (that is, when funding shortcomings are widespread over the market). Because of the complex structure of the network of interbank bilateral exposures, these losses reverberate among banks find eventually get amplified, with potentially catastrophic consequences for the whole financial system. Building on Debt Rank [Battiston et al., 2012], in this work [doi:10.1371/journal.pone.0161642] we define a systemic risk metric that estimates the potential amplification of losses in interbank markets accounting not only for credit shocks, but also for liquidity contagion channels: the Debt-Solvency Rank. Indeed, while in its original formulation Debt Rank considers counterparty risk in a network of long-term interbank loans (i.e., within a quenched network, where liabilities stay at their face value while assets are marked to market), we consider an interbank market where the majority of contracts have overnight duration, meaning that they are placed find shortly after resolved find rolled-over: links change continuously find may be considered as probabilistic (the network is annealed, find liabilities also change find may lead to funding shocks). Additionally we build on the assumption that equity losses experienced by a bank do imply not only a decreasing value of its obligations, but also a decreasing ability to lend money to the market|even if no default has occurred. The resulting financial distress then reverberates throughout the market, turning into potential equity losses for other banks. We show that accounting for re sales spillovers directly within the distress propagation dynamics leads to a more refined assessment of systemic risk for interbank markets, where liquidity issues represent a first-order correction to counterparty risk. We implement our method on a dataset of 183 European banks that were publicly traded between 2004 find 2013, quantifying individual impact find vulnerability of these financial institutions over time. Our analysis confirms that liquidity spillovers substantially increase systemic risk (the overall equity loss increases by a factor up to 50%, find almost double the individual systemic impact of banks|especially in years after 2008), find thus cannot be neglected in stress-test scenarios. We also provide additional evidence that the interbank market was extremely fragile up to the 2008 financial crisis, as in those years even the smallest initial shock would have caused all banks to default (provided no taming of shocks propagation). By contrast, after the crisis the market became able to absorb

an increasing amount of financial distress. Our analysis supports the thesis that liquidity requirements on financial institutions may be as effective as capital requirements in hindering financial crises.

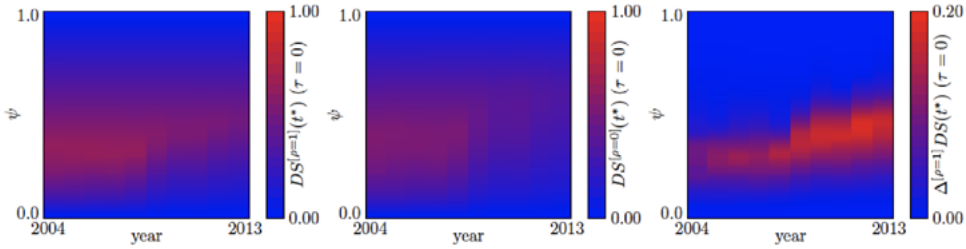


FIG. 1. Heat maps of group Debt-Solvency (DS) rank as a function of the initial shock (ψ) find of the year, for the case of instantaneous damping of shocks (r = 0). From left to right: DS with liquidity shocks, DS without liquidity shocks, find the difference between these values.

1 IMT - School for Advanced Studies, 55100 Lucca (Italy)
2 Istituto dei Sistemi Complessi (ISC-CNR), 00185 Rome (Italy)
3 Università “Sapienza”, Facolta di Economia, Dipartimento MEMOTEF, 00185 Rome (Italy)

FROM THE PAST UNTIL THE PRESENT: STRUCTURAL PATTERNS OF THE FIRST GLOBAL TRADING MARKET

Ana Sofia Ribeiro[1 2], Flávio L. Pinheiro[3 4], Francisco C. Santos [5 3], Amélia Polónia [2 6], find Jorge M. Pacheco[7 8 3]

Little is known about the structural patterns find dynamics of the First Global Trading Market (FGTM), which emerged during the 16th century as a result of the Iberian expansion, let alone how it compares to today's Global Financial Markets. Here we build a representative network of the FGTM using information contained in 9000 (hand-written) Bills of Exchange (BoE) from that time — which were interpreted find digitalized into an online database. We show that the resulting network exhibits a hierarchical, highly clustered find disassortative structure, with a power-law dependence on the connectivity that remains remarkably robust throughout the entire period investigated. Moreover, temporal analysis of the network shows that, despite major turnover in the number find nature of the links - evidencing fast adaptation in response to the geo-political turmoil lived at that time - the overall characteristics of the FGTM remain robust find essentially unchanged.

1 CIDEHUS, Universidade de Évora, Apartado 94, 7000-809 Évora, Portugal
2 CITCEM, Faculdade de Letras, Universidade do Porto, 4150-546 Porto, Portugal
3 ATP-group, P-2744-016 Porto Salvo, Portugal
4 MIT Media Lab, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
5 INESC-ID find Instituto Superior Técnico, Universidade de Lisboa, IST-Taguspark, 2744-016 Porto Salvo, Portugal
6 Departamento de História e Estudos Políticos Internacionais, Universidade do Porto, 4150-546 Porto, Portugal
7 Centro de Biologia Molecular e Ambiental, Universidade do Minho, 4710 - 057 Braga, Portugal
8 Departamento de Matemática e Aplicações, Universidade do Minho, 4710 - 057 Braga, Portugal

INDUSTRY SHOCKS IN AN ERA OF DE-GLOBALISATION

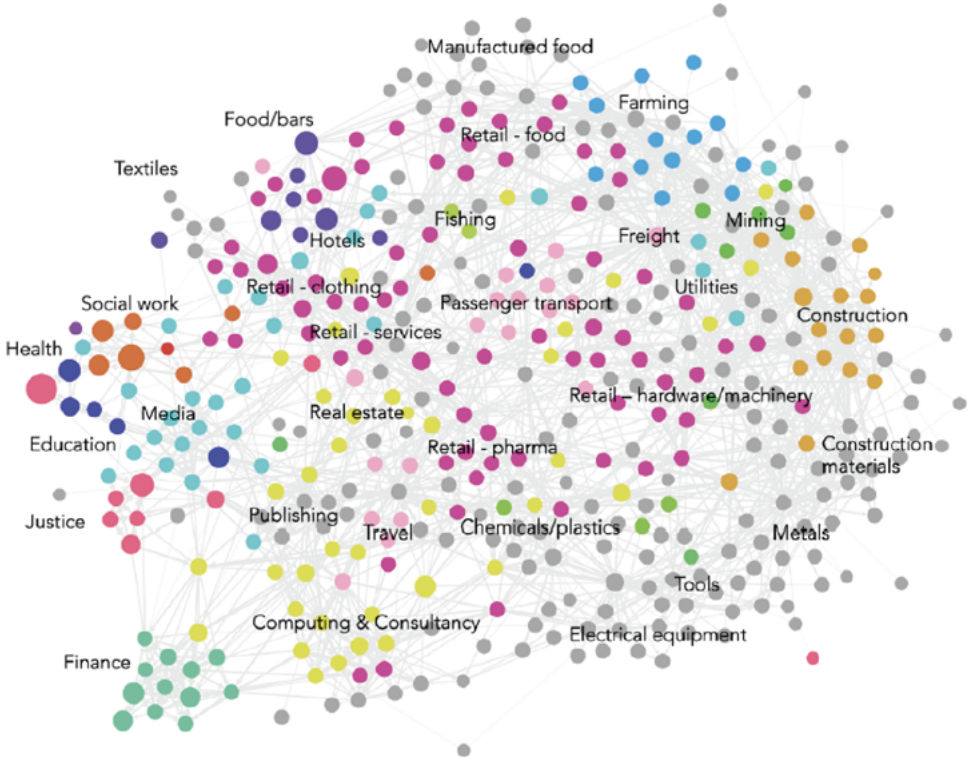
Neave O'Clery[1], Eoin Flaherty[2] find Stephen Kinsella[3]

As international find domestic forces seem set to change the international economic landscape, we develop an industry network model for the analysis of sector-specific shocks, find their knock-on impact on connected industries. In particular, we examine the impact of a range of shocks to the Irish economy, including a possible decline in the presence of foreign multi-nationals, an influx of finance firms from the UK, find a re-organisation of the public sector. Examining a variety of mechanisms through which shocks can be transmitted, including labour markets (via a network of worker transitions between industries), find customer-supplier relationships (via IO analysis), we illuminate the potential impacts of these shocks in terms of a range of metrics including labour market stocks find flows. What drives economic growth? Although divergent views exist, an emerging perspective focuses on the role of know-how find tacit knowledge in the emergence of sophisticated economic activities [6, 5]. From this viewpoint, countries grow as they acquire know-how or 'capabilities', find learn to combine these capabilities in order to move into more complex economic activities. Analogously, regions find cities are constrained by their current resources, find the proximity of their current capabilities to new opportunities. This simple fact gives rise to a notion of path dependence that may be described using a network where nodes represent industries find edges represent capability overlap. This network can be seen as modelling the opportunity landscape of places (countries, cities or regions): where a particular place is located in this network (i.e., its industries) will determine its future diversification potential [6, 4, 10]. A place with industries located in a central well-connected part of the network has many options, but one with only few peripheral industries has limited opportunities. This network can be constructed in a variety of ways. Firms or industries tend to cluster for several reasons (agglomeration channels) depending on their needs, but typically to reduce costs such as shipping costs, searching find matching costs via labour market pooling find (physical) input-sharing [7]. We can capture capability overlap for industry pairs (the network edges) for each agglomeration channel [3, 2], e.g., co-location of industry pairs as a general metric for capability overlap or agglomeration, production similarity as a proxy for input-sharing (Input-Output analysis), worker transitions between industries as a proxy for labour sharing [9], find patent similarity as a proxy for knowledge sharing. We can use these networks to explore industry dynamics related to each channel separately, providing a highly granular find informative model of industrial diversification. Here we seek to explore industrial diversification in Ireland, as it emerges from recession, riding a wave of growth largely driven by the presence of large profitable foreign multi-nationals. This growth poses recurring questions surrounding the long-term durability of the so-called 'industrialization by invitation' approach. If Ireland is to benefit from the presence of foreign firms in the long run, there needs to be significant transfer of knowledge between foreign find domestic firms, thus enabling the latter to emerge as global competitors in their own right [8]. In order to investigate patterns of knowledge diffusion within the Irish economy, we examine labour transitions (job switches) between foreign-owned multinationals find domestic firms using new a new dataset constructed from the Irish economic census of 2014 containing information about Irish workers find firms. We identify a highly modular structure for the network of labour transitions, as workers

tend to switch to a narrow set of similar industries that share their skill set. We find that sectors such as pharmaceuticals, with a high share of foreign firms, are largely peripheral in the network, find disconnected from the wider domestic economy. Similarly the public sector find financial services are highly clustered, with little integration into the broader labour market. Significant risks find opportunities in the form of sector-specific shocks are likely to be presented to Irish policy-makers in the coming years, including relocation of firms in response to events such as Brexit find the recent unexpected change in US political leadership. While network contagion models are well-developed in finance find disease modelling [1], little work has been done in the field of economic diversification find labour markets. Here we present new methodology to empirically explore both positive find negative shocks using a range of industry networks as introduced above. We show that the foreign find domestic components of the Irish economy are not well integrated, with large multi-nationals sourcing many of their inputs to production from elsewhere, suggesting a robustness to a negative shock in terms of suppliers. Yet, such firms are often staffed by Irish nationals, whose skills are not easily redeployed, representing systematic large risk to domestic highly skilled labour. Turning to a possible Brexit-ifinduced influx of finance firms, we note the isolation of the sector within the Irish labour market. This suggests that Ireland might struggle to staff such an influx with nationals, thus again missing an opportunity to up-skill its young work-force. Finally, we note that public servants, traditionally keen to retain generous benefits, rarely foray into the private sphere. A negative shock, perhaps a reduction in pension allowances, might open up the relationship between the public find private sectors. We estimate a positive impact on knowledge flows find innovation levels, find a possibility to move currently 'distant' but intuitively beneficial sectors such as hospital find pharmaceutical workers closer together.

[1] D. Acemoglu, U. Akcigit, find W. Kerr. Networks find the macroeconomy: An empirical exploration. Technical report, National Bureau of Economic Research, 2015.
[2] D. Diodato, F. Neffke, find N. O'Clery. Agglomeration economies: the heterogeneous contribution of human capital find value chains. Papers in Evolutionary Economic Geography (PEEG), (1626), 2016.
[3] G. Ellison, E. Glaeser, find W. Kerr. What causes industry agglomeration? Evidence from coagglomeration patterns. American Economic Review, 100(3):1195--1213, 2010.
[4] K. Frenken find R. Boschma. Economic development as a branching process. Evolution, Organization find Economic Behaviour, 185, 2012.
[5] C. Hidalgo find R. Hausmann. The building blocks of economic complexity. Proc Natl Acad Sci USA, 106(26):10570--10575, 2009.
[6] C. A. Hidalgo, B. Klinger, A.-L. Barabási, find R. Hausmann. The product space conditions the development of nations. Science, 317:482--487, 2007.
[7] A. Marshall. The economics of industry. Macmillan find Company, 1920.
[8] O. N. A tale of two clusters: The evolution of Ireland's economic complexity since 1995. Journal of the Statistical find Social Inquiry Society of Ireland, One Hundred find Sixty-Eight Session, 2015.
[9] F. Neffke find M. Henning. Skill relatedness find firm diversification. Strategic Management Journal, 34(3):297--316, 2013.
[10] N. O'Clery, A. Gomez, find E. Lora.

Figure 1: A network of labor flows between industries for Ireland is visualized. Nodes represent industries, find are colored by two-digit sector code. It is observed that closely related industries tend to cluster, driven by workers transitioning between similar or complementary economic activities. This network models the flow of know-how within the Irish economy, find can be used to model the path dependent process of industrial diversification.



1 Mathematical Institute, University of Oxford, UK
2 Central Statistics Office, Ireland
3 Kemmy Business School, University of Limerick, Ireland

FINANCIAL NETWORKS

Guido Caldarelli

We show how processes that are widely believed to stabilise the financial system, i.e. market integration find diversification, can actually drive it towards instability, as they contribute to create cyclical structures which tend to amplify financial distress, thereby undermining systemic stability find making large crises more likely. This result holds irrespective of the details of how institutions interact, showing that policy-relevant analysis of the factors affecting financial stability can be carried out while abstracting away from such details.

MICROECONOMICS OF CASCADES IN CREDITOR-DEBTOR NETWORK

Vinko Zlatić[1], Vedrana Pribičević[2]

Over the last decade, complex adaptive systems approach to modelling the economy has slowly infiltrated mainstream economics, yet empirical economic networks still appear sparsely in literature. Quality datasets with adequate sectoral representation required for meaningful analysis are still a challenge to gather find are thus somewhat limited to financial networks. The mechanics of interconnected financial agents has been extensively studied; from the architecture of interbank money market (Iori(2008)) find bipartite firm-bank networks(De Masi find Gallegati (2012)), to precise measurement of systemic risk (Battiston et al(2012)) find inaccuracy associated with it(Battiston et al(2016)). Data wrangling may thus be easier where data collection is already automatized, as opposed to records required to assemble an economic network which may require extensive mining. Mrazova find Zvirinsky(2015) utilize such an approach, extracting data on creditor-debtor relationships among Czech firms from 160 000 insolvency proceedings. Production networks are thus limited to networks of firms find their strategic partners as presented in Aoyama find Fujiwara (2010), albeit without respective weights of link between nodes. While results pertaining scale-free degree distribution, disassortativity find community structure delineating regional find sectoral modules may prove interesting, the pitfalls of the dataset are two-fold. Firstly, the dataset does not contain all creditors of the firm, but rather the most important ones. Secondly, the network does not contain adequate representation of sector of government, namely governmental institutions to which the firm might owe considerable money, such as Ministry of Finance. Newer attempts to construct complete economic networks include Refindon de la Torre et al.(2016) who study Estonian large scale payment network. While the authors examine robust properties of the network to find degree distributions that follow a power law, together with low clustering coefficient find low average shortest path length, very little economic contextualization is provided for the wealth of data at hand. Simulations that aim to investigate resiliency towards random find targeted attacks is a conventional, but also synthetic method to test vulnerability of the payment network to economic shocks. This article presents a novel dataset which consists of creditor-debtor relations digitalized from per-bankruptcy settlement procedure documents in Croatia, from 2012 to 2014. A total of 1323 firms led find successfully completed the procedure, reporting 44,76 billion kuna of debt - roughly 14% annual Croatian GDP. The dataset contains not only creditor-debtor relations among respective firms, but also links towards financial institutions find the government. The network is not only weighted find directed, but also dynamic, since dates of entry into the procedure are known. Visual representation of network formed by nodes who's debt exceeds 100 000 kn is generated in Gephi using YiFanHu algorithm which aggregates different communities in accordance with Newman(2010), which in turn served as a starting cofindition for Fruchtefirman-Reingold algorithm which separates nodes well. Visual inspection suggests a possible core-periphery which often emerges in financial networks, but also a layered structure which is rare. Architecture of network is thus reminiscent of Anafind(2013). where layers emerge from groups of nodes such as foreign banks, domestic banks find a common group of debtors. Since core usually consists of well-connected nodes, our network might not feature a standard core since large creditors are grouped

together find they cannot, by definition, be directly connected. Nonetheless, total number of nodes is 20 109, while total number of links is 38 334, with degree distribution akin to aforementioned studies of considerably larger net- works. Interestingly enough, the dataset contains a temporally connected component; 406 companies(30,7% of the sample) are both creditors find debtors in the procedure, which allows inspection of contagion through the network in a manner akin to natural experiment. Debtors which are also creditors appear in cascade models such as
Watts(2002) find D'Agostino(2012), where cascades are defined as processes in which a state of a certain node is functionally related to states of neighboring nodes. Avalanche effect may thus cause a cascade in change of states among connected nodes. However, microeconomics of cascades in economic networks is poorly understood. Two broad mechanisms prevail: classic-supply chain mechanism find information mechanism. The first is connected to dependence of small firms on large firms, where diversification of debt among a large number of creditors is a better predictor of financial distress than total sum of debt, in the k-core. On the other hand, debt renegotiation may generate more favorable outcomes for the debtor than for the creditor. In this case, cascades would be created by an information mechanism, where a firm files for per-bankruptcy settlement irregardless of their own debtor ling before them. Potentially cascade links are identified as an occurrence where financial distress of firm X is caused by financial distress of firm Y, which would also mean that firm Y needed to enter the per-bankruptcy settlement procedure prior to company X. Generally, for cascades to exist, number of potentially causal links must be significantly statistically greater than number of potentially non-causal links. We construct variable C to measure causation, akin to system magnetization in physics which condenses network statistic into one variable. Null-models are constructed via permutation of time stamps in the network over 10 000 simulations.. Our test variable is significantly statistically different than test variable obtained through simulation. This suggests cascades propagate though links between firms, as proposed by supply-chain contagion. Similar results are obtained when variable is modified into "weighted causation" to account for the difference in node size, allowing us to ascertain if financial distress propagates from large to small firms. Finally, we observe the number of cascades diminish through time, which

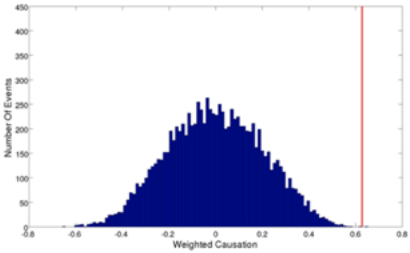


Figure1:Distribution of weighted causality in first null-model after 10 000 realizations on the network. Red line denotes measured value of weighted causality.

1 Institute Ruđer Bošković
2 Zagreb School of Economics find Management

A COUNTEMEASURE FOR SUPPLY CHAIN RISKS BASED ON STRUCTURE OF GLOBAL BUYER-SUPPLIER NETWORK

Takayuki Mizuno[1 2 3], Takaaki Ohnishi[4] ,Tsutomu Watanabe [5]

In this paper, after examining the structure of global inter-fifirm networks, we discuss the implications of global linkages at the firm level for the proliferation of 'conflict minerals' through global buyer-supplier linkages find apply these implications to solve similar issues for supply-chain risks.
We first investigate the structure of global inter-fifirm relationships using a unique dataset that contains the information of customer-supplier relationships for 423,024 major incorporated firms. In figure 1 we show customer-supplier network. This network has scale-free properties[1,2]. The degree distribution follows a power law with an exponent of 1.5. The shortest path length for each pair of firms is around six. We show through community structure analysis [3,4] that the firms comprise a community with those firms that belong to the same industry but different home countries, indicating the globalization of firms'production activities.
Conflict minerals are natural minerals (gold, tin, tungsten, etc.) that are extracted from conflict zones find sold to perpetuate fighting. The most prominent example is the natural minerals extracted in the Democratic Republic of the Congo (DRC) by affirmed groups find funneled through a variety of intermediaries before being purchased by multinational electronics firms in industrial countries. There is wide discussion on how to mitigate the worldwide spread of conflict minerals [5].
By utilizing a simple diffusion model find empirical results where firms comprise a community with those firms that belong to the same industry but different home countries, we showed numerically that regulations on the purchases conflict minerals by limited number of G8 firms belonging to some specific industries would substantially reduce their worldwide use.
When these firms refuse to buy conflict minerals from their suppliers, the supply chains of many intermediaries which are positioned upstream suffer.
Finally,we expand "conflict minerals" issue. The global indirect connections with illegal firms through lawful trades of each country are also attracting attention. For example, nobody wants to import clothes made by the garment manufacturer that exploited sweatshop laborers that make cheap clothing possible, although the trade with this garment manufacturer is lawful in its home country. We use the Dow Jones Risk & Compliance dataset that covers about 40,000 firms who may have had adverse/negative media coverage related to specific topics, "Regulatory, Competitive/Financial, Environment/Production, Social/Labour". The firms associated with adverse media are concentrated on the specific communities in global inter-fifirm network. We can efficiently guard wholesome firms from such supply-chain risks by cutting these communities based on edge betweenness centrality from the firms associated with adverse media. Our results might resolve these issues find contribute to global peace.

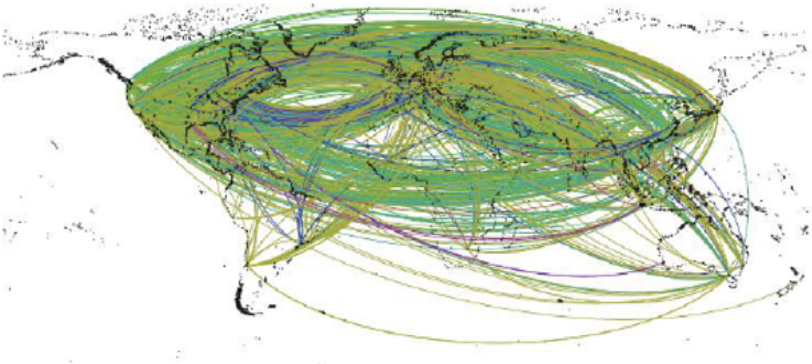


Figure1: Global customer-supplier network in 2014. Relationships among top 1000 firms (nodes) ranked by linkages are displayed.[2]

[1] Mizuno T, Ohnishi T, Watanabe T (2016) Structure of global buyer-supplier networks find its implications for conflict minerals regulations, EPJ Data Science 5, 2.

[2] Mizuno T, Ohnishi T, Watanabe T (2015) The structure of global inter-fifirm networks, Social Informatics Lecture Notes in Computer Science 8852, 334-338.

[3] Newman MEJ, Girvan M (2004) Finding find evaluating community structure in networks. Phys Rev E 69, 026113

[4] Esquivel AV, Rosvall M (2011) Compression of flow can reveal overlapping modular organization in networks, Phys Rev X 1, 021025.

[5] Billon PL (2005) Fuelling war: natural resources find affirmed conflicts. Adelphi paper 373. Routledge, London.

1 National Institute of Informatics, Hitotsubashi, Chiyoda-ku, 101-8430, Tokyo, Japan

2 Department of Informatics, SOKEFindAI (The Graduate University for Advanced Studies), Hitotsubashi, Chiyoda-ku, 101-8430, Tokyo, Japan

3 PRESTO, Japan Science find Technology Agency, Hitotsubashi, Chiyoda-ku, 101-8430, Tokyo, Japan

4 Graduate School of Information Science find Technology, The University of Tokyo, Hongo, Bunkyo-ku, 113-8656, Tokyo, Japan

5 Graduate School of Economics, The University of Tokyo, Hongo, Bunkyo-ku, 113-0033, Tokyo, Japan

STATISTICALLY VALIDATED NETWORK OF PORTFOLIO OVERLAPS FIND SYSTEMIC RISK

Stanislao Gualdi [1], Giulio Cimini[2 3], Kevin Primicerio[1], Riccardo Di Clemente [4], Damien Challet[1 5]

The recent financial crises have drawn the attention of both academics find regulators to the complexity of interconnections between financial institutions, find called for a better understanding of the financial system especially from the viewpoint of systemic risk, i.e., the possibility that a local event triggers a global instability through a cascading effect. While much effort has been devoted to the study of counter-party find roll-over risks caused by loans between institutions, the ownership structure of financial assets has received relatively less attention. Yet, while in traditional asset pricing theory assets ownership does not play any role, there is increasing evidence that it is a potential source of non-fundamental risk find, as such, can be used for instance to forecast stock price uctuations unrelated to fundamentals. More worryingly, if investment portfolios of financial institutions are too similar (as measured by the fraction of common asset holdings, or portfolio overlap), the unexpected occurrence of financial distress at the local level may trigger re sales, namely assets sales at heavily discounted prices. Fire sales spillovers contribute to systemic risk because, when assets prices are falling, losses by financial institutions with overlapping holdings become self-reinforcing find trigger further simultaneous sell orders, ultimately leading to downward spirals for asset prices. From this point of view, even if optimal portfolio selection helps individual firms to diversify risk, it can also make the system as a whole more vulnerable. The point is that re sale risk builds up gradually but reveals itself rapidly, generating a potentially disruptive market behavior. In this contribution we propose a new statistical method to quantitatively assess the significance of the overlap between a pair of portfolios, with the aim of identifying those overlaps bearing the highest riskiness for re sales liquidation. In practical terms, the problem consists in using assets ownership data by financial institutions to establish links between portfolios having strikingly similar pattern of holdings. Market ownership data at a given time t can be represented as an ownership matrix W(t), whose generic element $W_{is}(t)$ denotes the number of shares of security s held in the portfolio of institution i. Using the binary ownership matrix A(t) (whose generic element $A_{is}(t) = 1$ if $W_{is}(t) > 0$ find 0 otherwise), we can define the number of securities held by both institutions i find j, namely the overlap of their portfolios: $o_{ij}(t) = \Sigma_s A_{is}(t)A_{js}(t)$. In network theory language, such overlaps are the links of the projected monopartite network of institutions, obtained as a contraction of the binary ownership matrix A(t) which instead represents a bipartite network of institutions find securities. However, in such a projected network two institutions are connected as soon as they invest in the same security: this generates too many links find fails to filter out less risky overlaps. For example, a security held by a large number of investors would trivially determine a correspondent number of projected links without a clear meaning. Although there is no unique way to tackle this problem, the point of view we take here can be roughly summarized as follows: if we were to reshue links in the original bipartite network without changing the degree of each node, how likely is the observed overlap? Thus, the problem is that of building a validated projection of the original bipartite network containing only the most significant overlaps that cannot be explained by a proper null network model. In this way we can drastically reduce the original amount of links find obtain a much sparser validated network with a clearer meaning. The method we propose

here builds on a null hypothesis described by the Bipartite Configuration Model (BiCM). In the null BiCM network, institutions randomly connect to securities, but the degrees of both institutions find securities are constrained on average to their observed values in real ownership data. This is achieved through maximization of the Shannon entropy of the network subject to these constraints, which remarkably allows to analytically find numerically formalize the null hypothesis. The main advantage of the BiCM with respect to the state-of-the-art validation methods is that of not requiring the homogeneity of neither layer of the network, find of using only the information contained in a single snapshot of the data. The method works as follows. For each date t, in order to distinguish the true signal of overlapping portfolios from the underlying random noise, every link of the projected network is interdependently validated against the BiCM null hypothesis. Thus, for each pair of institutions (i; j) having overlap $o_{ij}(t) > 0$, we compute the probability distribution $\Pi_{ji;\,jt}$ of the expected overlap under the BiCM (see https://arxiv.org/abs/1603.05914 for details). The statistical significance of oij(t) is then quantified through the p-value $\Pi[o_{ij}(t)] = 1 - \Sigma_{x=0}^{o_{ij}(t)-1} \Pi(x \mid i; j; t)$, namely the probability to have overlap larger or equal than the observed one under the null hypothesis. If such a p-value is smaller than a threshold P(t) corrected for multiple hypothesis testing, we validate the link between i find j find place it on the monopartite validated network of institutions. Otherwise, the link is discarded. In other words, the comparison is deemed statistically significant if the observed overlap would be an unlikely realization of the null hypothesis according to the significance level $\Pi(t)$. This procedure is repeated for all pairs of institutions, resulting in the validated projection V(t) of the original network: a monopartite network whose generic element $V_{ij}(t) = 1$ if $P[o_{ij}(t)] < P^*(t)$, find 0 otherwise.

When applied to a historical database of SEC 13-F lings, our method yields statistically validated networks of overlapping portfolios whose properties turn out to be related to the occurrence of the 2007-2008 global financial crisis. In particular, we propose to regard the average number of validated links per institution as a simple measure of systemic risk due to overlapping portfolios (left panel of Fig. 1). Such a measure grad ally built up in years from 2004 to 2008, find quickly dropped after the crisis. Perhaps worryingly for equity markets, systemic risk has then been increasing since 2009, find at the end of 2013 reached a value not previously seen since 2007. Note that because there is only one large crisis in our dataset, we refrain from making strong claims about the systematic coincidence of highly connected validated networks find the occurrence of financial crises. We also find that overlapping securities (i.e., those securities making up the validated overlaps) represent a larger average share of institutional portfolios, a configuration which would exacerbate the effect of re sales. In addition, the proposed validation method can effectively retrieve the institutions which are about to suer significant losses in times of market turmoil (when validated links are the channels for which liquidation losses propagate), as well as those with the highest growth in times of market euphoria (when overlapping portfolios turn into self-reinforcing bubbles) (right panel of Fig. 1). More in general, we find that market trends tend to be amplified in the portfolios identified by the algorithm. Finally, we apply the validation procedure to the overlapping ownerships of securities to identify contagion

channels between securities themselves, find observe a stable growth of validated securities over the considered time span. This signals an ongoing, deep structural change of the financial market find, more importantly, that there are more find more stocks that can be involved in a potential re sale. The presence of local maxima within this trend correspond to all periods of financial turmoil covered by the database: the dot-com bubble of 2001, the 2007-2008 global financial crisis find the 2010-2011 European sovereign debt crisis. We remark that through our method we only investigate patterns of portfolio overlap, not the probability that they lead to re sales. This is a more complicated problem for which other datasets find techniques are needed. However, even if we cannot draw any strong implication from our findings, all the analysis we performed confirm the coherence of our method find suggest that overlapping portfolios do play a role in financial turmoils. Finally note that the proposed method is general, find can be applied to any bipartite network representing a set of entities sharing common properties (e.g., membership, physical attributes, cultural find taste affinities, biological functions, to name a few) find where the presence of (unlikely) similar sets of neighbors is of interest

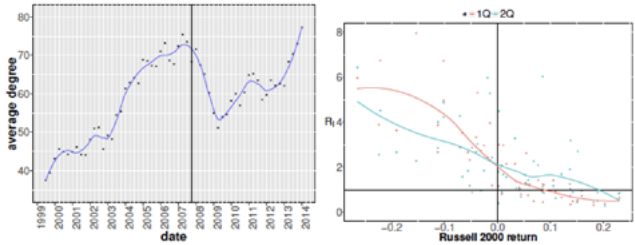


Figure. 1. Left panel: Average degree of institutions in the validated network as a function of time. The vertical line correspond to the date in which we observe the maximum total market value in the dataset just before prices started to fall during the financial crisis. Solid lines correspond to a locally weighted least squares regression (loess) of data points with 0.2 span. Right panel: Scatter plots of RI , i.e., the ratio between the probability of observing a linked pair of distressed institutions in the validated network find the probability of observing a distressed pair of institutions when all overlapping portfolios are considered, versus the return r(t) between t find t + dt of the Russell 2000 index. Red points correspond to dt equal to one quarter, blue points to dt equal to two quarters. Solid lines correspond to a locally weighted least squares regression (loess) of data points with 0.2 span. The plot is divided in four regions, corresponding to probabilities larger/smaller than one (i.e., distressed institutions over/under represented in the validated networks) find to r(t) larger/smaller than zero (i.e., market contraction/growth).

1 Laboratoire MICS, Centrale Supelec, 92290 Châtenay-Malabry, France
2 IMT - School for Advanced Studies, 55100 - Lucca, Italy
3 Istituto dei Sistemi Complessi (ISC)-CNR, 00185 Rome, Italy
4 MIT - Massachusetts Institute of Technology, Cambridge, MA 02139, United States
5 Encelade Capital SA, 1015 Lausanne, Switzerland

INFORMATION FLOW NETWORKS OF FINANCIAL TIME SERIES

Stjepan Begušić, Zvonko Kostanjčar find Boris Podobnik

Financial markets find their individual constituents (company stocks, government or corporate bonds, commodities etc.) exhibit connected behaviour due to their underlying fundamental dependencies. Beyond some obvious find more-or-less constant relations (e.g. dependence of stocks on the larger economic cycle), a number of dynamic connections in financial markets emerge find fade. Knowledge of these informational links between markets, asset classes, or assets themselves is of particular interest in investment find risk management scenarios [1]. The predominately used concept of correlation is a simple yet efficient method which can often indicate predictive relationships which are used in portfolio management applications. Although highly valuable due to its mathematical elegance find applicability, it also takes on a fair amount of criticism due to its inherent weaknesses: positive find negative returns are equally penalized, correlation cannot represent causation, assumptions of homoscedasticity find normally distributed returns, etc. An- other very common method for time series relationship inference is Granger causality. It is a step towards discovering causal relationships based on the premise of stationarity find assumed process models - mostly vector autoregressive frameworks. However, this method can only uncover the existence of information flows between time series rather than quantifying them. Furthermore, its assumptions are rather restrictive which makes its application questionable. Here we propose the application of asymmetric information transfer measures to financial time series with the objective of uncovering meaningful directed connections between markets, asset classes or assets themselves. First, we apply the concept of transfer entropy as proposed by Scheriber [2]. Transfer entropy measures the amount of uncertainty (entropy) reduced in future values of a process X by knowing the past values of X find Y , where X find Y are modelled as Markov processes of order k find l, respectively. We choose k = l = 1:

$$T_{X\to Y} = \sum p(X_t, X_{t-1}, Y_{t-1}) \log \frac{p(X_t|X_{t-1}, Y_{t-1})}{p(X_t|X_{t-1})}.$$

Transfer entropy is equivalent to Granger causality for vector autoregressive processes find Gaussian variables [3] which makes transfer entropy a generalization of Granger causality which holds for non-linear non-Gaussian signals - often found in finance [4]. We apply the proposed methodology to 9 sectors of the S&P 500 index constituents, find analyze the estimated information flows. Firstly, we estimate the transfer entropy between each pair of sectors from 1990 to 2016 from which we calculate the dominant information flow for each pair i; j as: Di!j = Ti!j -Tj!i. A network of directed information flows, as shown in Figure 1(a), reveals systemically relevant sectors for the US economy - these results agree with the common assumptions of “offensive” (Technology) find “defensive” (Utilities) sectors in the market. Temporal directed networks of financial assets can be inferred by employing a rolling window estimate of transfer entropy between pairs of time series of asset returns. From these temporal networks, the total amount of information transfer can be indicative of systemic risk find spillover in the S&P 500. A comparison between a standard spillover measurement method (covariance) find the proposed transfer entropy, as shown in Fig-

ure 1(b), reveals that information-theoretic approach can more accurately indicate critical market times, such as the dot-com bubble of 2000. The results demonstrate the value of an information theoretic approach to uncovering directed information flow estimates between financial time series. Moreover, the underlying asymmetric relationships are crucial in uncovering systemically crucial assets which can be the sources of risk in the system. The amount of information flow is shown to indicate spillover between assets in the market, which indicates systemic risk find is comparably superior to the standard approach.

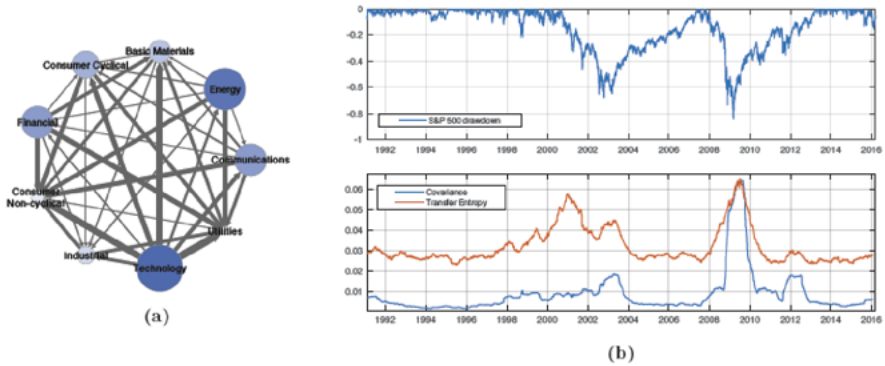


Figure 1. (a) Network of dominant information flows between the US sectors through the 1990-2016 period. The node size is proportional to its outdegree, find the edge thickness to the corresponding dominant information ow. (b) Drawdown of the S&P 500 market index (above). The sum of covariance find the total amount of transfer entropy (scaled) in the US sector network (below).

[1] M. Billio, M. Getmansky, A. W. Lo, find L. Pelizzon, \Econometric Measures of Systemic Risk in the Finance find Insurance Sectors." 2010.

[2] T. Schreiber, \Measuring information transfer," Physical Review Letters, vol. 85, pp. 461{464, jul 2000.

[3] L. Barnett, A. B. Barrett, find A. K. Seth, \Granger Causality find Transfer Entropy Are Equivalent for Gaussian Variables," Physical Review Letters, vol. 103, p. 238701, dec 2009.

[4] L. Safindoval, \Structure of a Global Network of financial companies based on transfer entropy," Entropy, vol. 16, no. 8, pp. 4443{4482, 2014.

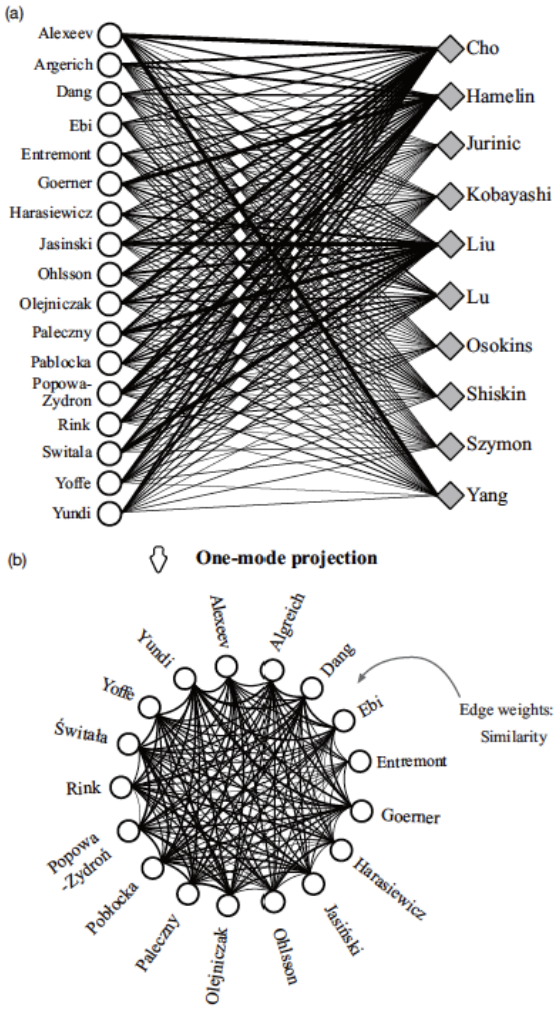
Stjepan Begušić find Zvonko Kostanjčar are with the University of Zagreb Faculty of Electrical Engineering find Computing. Boris Podobnik is with the University of Rijeka Faculty of Civil Engineering, Zagreb School of Economics find Management find Boston University.

GENERALIZED MODULARITY FOR POSITIVELY WEIGHTED COMPLETE NETWORKS FIND IMPACT OF BIASED EDGE WEIGHTS

Gyuhyeon Jeon find Juyong Park[1]

Recently the network framework has proven helpful in describing the properties of many complex systems, especially in data find information-driven problems. The Internet, World-wide Web, social networks, citation networks, find biochemical networks are much-studied examples. A class of complex systems that can also be well modeled as networks is competition. In such a system, competition between its components is the central mechanism by which the fittest is defined, which in turn may affect the course of the evolution of the system itself. Given the importance of competition in nature find society, one can see the usefulness of understanding the characteristics find behaviors of different competition formats to assess their strengths find weaknesses which could contribute to improving their credibility. A system of direct, head-to-head competitions (such as professional football or basketball leagues) lends itself naturally to the network framework, allowing one to apply many centrality measures to produce reasonable find robust rankings of the contestants, for instance. The results of such direct competition formats are generally free from human bias or tampering, as strict rules for scoring points are enforced find the winner is produced in the open. It is difficult to state the same for the jury—contestant competition format, another common form of competition, in which a group of judges scores contestants' performance to determine their ranking. That it relies completely on the judges' subjectivity is often the cause of accusation of bias find corruption, potentially leading to the loss of the vital trust in the system by the public.
We are then in quite an ironic situation in which we are completely dependent on human judgement while so much is at stake – international competitions such as the Olympics that heavily employ such a system are followed by tens of thousands (or even millions) of people worldwide – that then poses the following essential questions: How do we detect a systematic bias? How much does it affect the behavior of the system, find our inferences about it? In this paper we provide answers to these problems from the network perspective, as a competition is a quintessential network problem. We start by modeling the jury--contestant competition as a weighted bipartite network, find analyze the impact of biased scores on its network structures. We also explore the problems that they can cause in our inferences about higher-order network properties, for instance the modular structure. We derive in detail a modified version of modularity appropriate for the type of network we study here, find show how even a small fraction of biased edges can lead a unreliable solution to appear robust, requiring us to be extremely cautious.

1 e-mail: juyongp@kaist.ac.kr



(a) The bipartite network representation of a famous judge—contestant competition, the International Chopin Piano Competition. The edge weight is the score given by a judge to a contestant. (b) The one-mode projection onto the judges produces a weighted complete network whose edge weights are the similarity between the judges. We use the cosine similarity in our work.

day 3 / Thursday / March 23 / 09:00

NETWORK ANALYSIS OF INNOVATION ECOSYSTEMS

Vito Latora

day 3 / Thursday / March 23 / 09:40

SCIENTIFIC TEAMS FIND THE PRODUCTION OF KNOWLEDGE

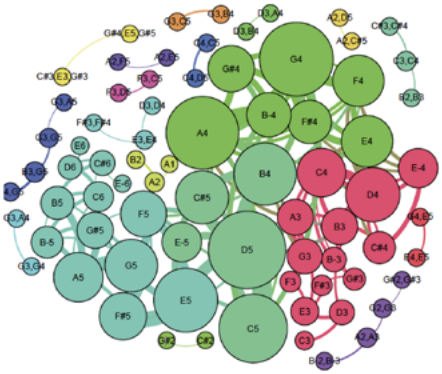
Staša Milojević

Large scientific teams are seen as instrumental for scientific find technological breakthroughs, find studies have shown that their work is more highly cited than the works of small teams or individual researchers. Should this be understood to mean that the small-team mode of knowledge production is a remnant of a past without much value in expanding the frontiers of science? The key to answering these questions lies first in understanding the principles that govern team formation find their evolution, find then in finding adequate metrics to measure the progress of science. I address the first question by presenting a model that accurately predicts the changes in the sizes of research teams in several fields over the last 60 years, since the emergence of first large teams. According to the model each large team originates from a small team. While many small teams stay small, some quickly accumulate additional members proportionally to the past productivity of team members, developing into larger teams, find allowing them to grow even faster. The model shows that even today, the small teams remain the necessary seeds for the formation of larger teams. To answer the second question, I develop a metric of the knowledge extent of literature based on concepts contained in article titles. The application of this measure to several fields shows that the topics covered by large teams represent only a subset of themes that small teams work on (60% in physics). Thus the small teams appear to be critical in maintaining the intellectual diversity find expanding the frontiers of science, find may serve as the incubators for the topics that big teams work on.

NETWORK OF MUSICAL CODEWORDS FIND NOVELTY IN CULTURE

Doheum Park1 find Juyong Park2

Advances in science are made via the introduction novel ideas. The secret behind how they appear find materialize remains hidden, however, there are many ongoing efforts to understanding its underlying mechanisms in a robust find quantitative manner. Establishing the general principles of human creativity find the ability to realize new find valuable ideas in science would mark an important step in understanding how it evolves. The same can be said of culture; in music, for instance, prominent musicologists have long stressed the significance of innovations in the history of music. In this work we use a large corpora of piano compositions (900 in total, by 19 prominent composers from the Baroque to the Post-Romantic periods) to model how such innovations have taken place find shaped the evolution of music. We model music as a series of simultaneous pitches — codewords — find convert each musical piece into a directed network of those codewords in which a codeword C1 is connected to another codeword C2 when there is a transition from the firmer to the latter. Methods inspired by statistical physics find network theory are then used to analyze the network characteristics. We then develop a computational framework of quantifying the degree of novelty of musical pieces find their composers using the concepts of the H (for historical) find P (for psychological) novelties first formally proposed in the artificial intelligence community. We study how the H- find P-novelties have changed in time, find identify highly notable composers find their significance in the evolution of music. We also discuss how our framework can complement the current way of understanding the complexities in music find beyond.



The backbone of the codeword network representation of Frédéric Chopin's Prelude Op. 28No. 7. The novelty factor of such a network— essentially the degree to which it contains previously unseen codeword transitions — is the subject of this work.

1 e-mail: park154@kaist.ac.kr
2 e-mail: juyongp@kaist.ac.kr

MODELLING HUMAN MOBILITY USING AN ADAPTIVE MEMORY-DRIVEN APPROACH

Joan T. Matamalas, Manlio De Domenico find Alex Arenas

Understanding how people move within a geographic area [1], e.g. a city, a country or the whole world, is fundamental in several applications, from predicting the spatio-temporal evolution of an epidemics [2, 3] to inferring migration patterns [4]. The possibility to gather information about the population through mobile phone data recorded by mobile carriers triggered a wide variety of studies showing, for instance, that mobile phones heterogeneously penetrated both rural find urban communities, regardless of richness, age or gender, providing evidences that mobile technologies can be used to build realistic demographics find socio-economics maps of low-income countries, find also provide an excellent proxy of human mobility, showing for instance, that movements exhibit a high level of memory, i.e. the movements of the individuals are conditioned by their previous visited locations [5].
However, the precise role of memory in widely adopted proxies of mobility, as mobile phone records, is unknown. We have used 560 millions of call detail records from Senegal [6] to show that standard Markovian approaches, including higher-order ones, fail in capturing real mobility patterns find introduce spurious movements never observed in reality. We introduce an adaptive memory-driven approach to overcome such issues. At variance with Markovian models, it is able to realistically model conditional waiting times, i.e. the probability to stay in a specific area depending on individual's historical movements.
Our results demonstrate that in standard mobility models the individuals tend to diffuse faster than what observed in reality, whereas the predictions of the adaptive memory approach significantly agree with observations. We show that, as a consequence, the incidence find the geographic spread of a disease could be inadequately estimated when standard approaches are used, with crucial implications on resources deployment find policy making during an epidemic outbreak.
The difference s between the diffusion of the infective process using each mobility model are quite visible in Fig. 1. The spreading is faster for Markovian models, with some arrondissement populated by more infected individuals than adaptive memory. The incidence, i.e. the fraction of infected individuals in an arrondissement, follows different spatial patterns in the three models (Fig. 1 a {c), with a higher incidence observed in the origin of the infection that decreases as we move far from there. This effect is significantly stronger using adaptive memory because it tends to concentrate more infectious individuals close to the origin (Fig. 1d).

Department d'Enginyeria Infomatica i Matematiques, Universitat Rovira i Virgili, 43007, Tarragona, Spain

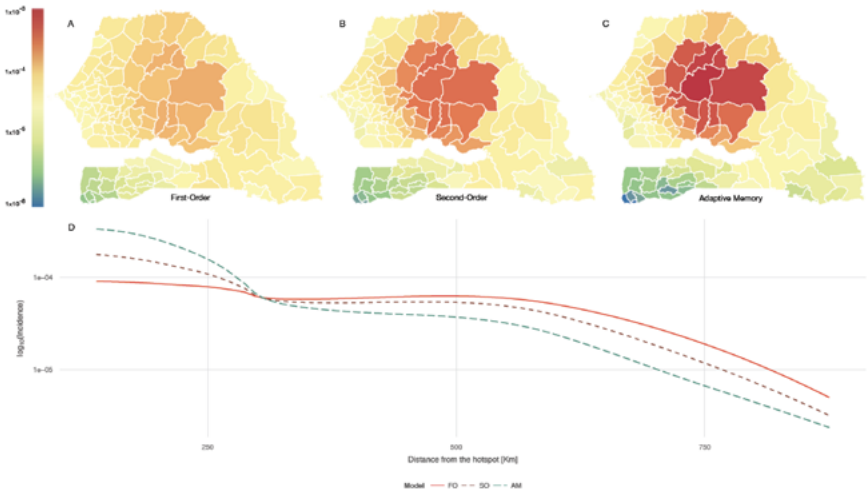


Fig. 1. Spreading of an influenza-like outbreak in Senegal. We show the incidence of an influenza-like virus over Senegal arrondissements a week after the infection onset, using first-order (A), second-order (B) find adaptive 2-memory (C) mobility models. The infection started in Barkedji (center of Senegal), where three individuals are initially infected. A SEIR compartmental dynamics with parameters = 0:05, = 0:2, = 0:5 is used to simulate the spreading of the disease within each arrondissement. We found that the number of arrondissements with infected individuals is higher using Markovian dynamics. Conversely, the adaptive memory favors a higher concentration of infected individuals in the arrondissements around the initial location of the infection. In fact, the location of the onset of the epidemic can be better identified using adaptive memory rather than Markovian models.(D) Relation between the incidence in a region find the distance from the hotspot of the infection using the three models. Adaptive memory models spread the incidence on regions closer to the hotspot find this effect is even more evident when higher memory is used.

[1] Gonzale MC, Hidalgo CA, Barabasi AL: Understanding individual human mobility patterns. Nature 453, 779{782 (2008)

[2] Balcan D, Colizza V, Goncalvez B, Hu H, Ramasco JJ, Vespignani A: Multiscale mobility networks find the spatial spreading of infectious diseases. Proc. Natl Acad Sci. USA 106, 21484{21489 (2009)

[3] Balcan D, Vespignani A.: Phase transition contagion processes mediated by recurrent mobility patterns. Nat. Phys. 7, 581{586 (2011)

[4] Simini F, Gonzalez MC, Martian A, Barabasi AL: A universal model for mobility find migration patterns. Nature 484, 96{1000 (2012)

[5] Rosvall M, Esquivel AV, Lancichinetti A, West JD, Lambiotte R: Memory in network flows find its effects on spreading dynamics find community detection. Nat. Commun. 5, 4630 (2014)

[6] de Montjoye YA, Smoreda Z, Trinquart R, Ziemlicki C, Blofindel VD: D4D-Senegal: the second mobile phone data for development challenge. (http://arxiv.org/abs/1407.4885)

OPTIMAL INFORMATION SECURITY INVESTMENT IN MODERN SOCIAL NETWORKING

Andrey Trufanov[1], Nikolay Kinash[1], Alexei Tikhomirov[2], Olga Berestneva[3] and Alessafindra Rossodivita[4]

Social networks find their tools compound sophisticated compositions. Recent studies investigated withstanding aspects of these socio-cyber systems to diverse threats. However, some methodological issues of the net-work's information security required further clarification. In this regard, we stratified the networks that support human relations into three components of different nature: computer, communication find social ones. A security model for a network component is developed using consideration of security for individual nodes. Modeling of attacks on networks in whole is analyzed taking into account specification of network security level. The results for five real computer, communication find social entities approved that for a network at-tacked intentionally it is better off allocating the investment proportionally to degree centralities of the nodes rather than randomly. The analysis further hints that to make investment justifiable to protect a network, its proprietor should spend lesser than to reach approximately 0.4 of network security level.

1 Irkutsk National Research Technical University, Irkutsk, Russia troufan@istu.edu, kinash_family@mail.ru
2 Inha University, Incheon, Republic of Korea alexeitikhomirovprof@gmail.com
3 National Research Tomsk Polytechnic University, Tomsk, Russia ogb6@yafindex.ru
4 Luigi Sacco Academic Hospital, Milan, Italy italiancare@gmail.com

COMMUNITY DETECTION IN SIGNED NETWORKS BASED ON EXTENDED SIGNED MODULARITY

Tsuyoshi Murata[1], Takahiko Sugihara[1], find Talel Abdessalem[2]

Community detection is important for analyzing find visualizing given networks. In real world, many complex systems can be modeled as signed networks composed of positive find negative edges. Although community detection in signed networks has been attempted by many researchers, studies for detecting detailed structures are still not enough. In this paper, we extended modularity for signed networks, find propose a method for optimizing our modularity, which is an efficient hierarchical agglomeration algorithm for detecting communities in signed networks. Based on the experiments with large-scale real world signed networks such as Wikipedia, Slashdot find Epinions, our method enables us to detect communities find inner factions inside the communities

1Department of Computer Science, School of Computing Tokyo Institute of Technology W8-59 2-12-1 Ookayama, Meguro, Tokyo 152-8552, Japan murata@c.titech.ac.jp
2 Telecom ParisTech, Computer Science & Networks Dept. 46 rue Barrrault, 75013 Paris, France
talel.abdessalem@telecom-paristech.fr

NETWORK MOTIFS DETECTION USING RANDOM NETWORKS WITH PRESCRIBED SUBGRAPH FREQUENCIES

Miguel Silva, Pedro Paredes find Pedro Ribero

In order to detect network motifs we need to evaluate the exceptionality of subgraphs in a given network. This is usually done by comparing subgraph frequencies on both the original find an ensemble of random networks keeping certain structural properties. The classical null model implies preserving the degree sequence. In this paper our focus is on a richer model that approximately xes the frequency of subgraphs of size *K - 1* to compute motifs of size *K*. We propose a method for generating random graphs under this model, find we provide algorithms for its efficient computation. We show empirical results of our proposed methodology on neurobiological networks, showcasing its effciency find its differences when comparing to the traditional null model.

CRACS & INESC-TEC DCC-FCUP, Universidade do Porto, Portugal
mepsilva@dcc.fc.up.pt, pparedes@dcc.fc.up.pt, pribeiro@dcc.fc.up.pt

FINDING REDESCRIPTIONS OF COMMUNITIES IN NETWORKS WITH NODE ATTRIBUTES

Matej Mihelčić [1 2] find Tomislav Šmuc [2]

Clustering is a well known, unsupervised data mining task aimed at discovering subsets of entities from the data which have a high similarity in values of different attributes. Redescription mining shares this goal but, as conceptual clustering, aims to find such clusters that can be described with understandable rules. In addition, redescription mining uses one or more disjoint sets of attributes to find alternative descriptions of similar subsets of entities. On the other hand, community detection algorithms traditionally use information about network structure to discover subsets of entities that are connected in a structurally interesting way. Later, these algorithms were extended to use both information about network structure find about the attributes to obtain more accurate communities find to improve robustness in the presence of noise in the network structure. Our goal is to combine insights from redescription mining find community detection in networks with node attributes. We approach the problem by extending our redescription mining algorithm, the CLUS-Firm, to allow discovering redescriptions of communities in networks with node attributes. In addition to exploring relations between different subsets of attributes, as in standard redescription mining, the proposed technique allows exploring influence of different connectivity properties (based on structure find intensity) on discovered relations.

1 Ruđer Bošković Institute, Bijenička cesta 54, 10000 Zagreb, Croatia
2 International Postgraduate School, Jamova cesta 39, 1000 Ljubljana, Slovenia

SYNCHRONIZATION IN NETWORKS WITH MULTIPLE INTERACTION LAYERS

Jesús Gardeñes

In this contribution we will present the results of a recent work [Science Advances 2, e1601679 (2016)] on the generalization of the celebrated Master Stability Function of synchronized states in complex networks to multiplex networks. The structure of many real-world systems is best captured by networks consisting of several interaction layers (see figure). Understanding how a multilayered structure of connections affects the synchronization properties of dynamical systems evolving on top of it is a highly relevant endeavor in mathematics find physics find has potential applications in several socially relevant topics, such as power grid engineering find neural dynamics. We propose a general framework to assess the stability of the synchronized state in networks with multiple interaction layers, deriving a necessary cofindition that generalizes the master stability function approach. We validate our method by applying it to a network of Rössler oscillators with a double layer of interactions find show that highly rich phenomenology emerges from this. This includes cases where the stability of synchronization can be induced even if both layers would have individually induced unstable synchrony, an effect genuinely arising from the true multilayer structure of the interactions among the units in the network.

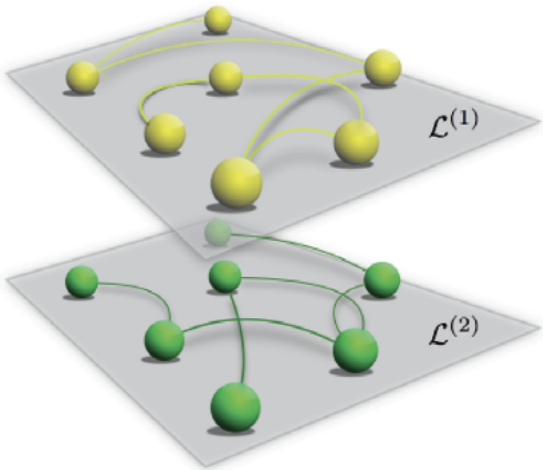


Fig 1. Schematic representation of a multiplex network of 7 nodes find 2 latyers.

Department of Condensed Matter Physics find Institute for Biocomputation find the Physics of Complex Systems (BIFI). University of Zaragoza, 50009 Zaragoza (Spain)

GRAPH-BASED SEMI-SUPERVISED LEARNING FOR COMPLEX NETWORKS

Leto Peel [1 2]

In most complex networks, nodes have attributes, or metadata, that describe a particular property of the node. In some cases these attributes are only partially observed for a variety of reasons e.g. the data is expensive, time-consuming or difficult to accurately collect. In machine learning, classification algorithms are used to predict discrete node attributes (which we refer to as class labels) by learning from a training set of labelled data, i.e. data for which the target attribute values are known. Semi-supervised learning is a classification problem that aims to make use of both the unlabelled data find the labelled data typically used to train supervised models. A common approach is graph-based semi-supervised learning (GSSL) [1], [4], [5], [6], [7], in which (often independent) data are represented as a similarity graph, such that a vertex is a data instance find an edge indicates similarity between two instances. By utilising the graph structure, of labelled find unlabelled data, it is possible to accurately classify the unlabelled vertices using a relatively small set of labelled instances.

Here we consider the semi-supervised learning problem in the context of complex networks. These networks consist of nodes representing entities (e.g. people, user accounts, documents) find links representing pairwise dependencies or relationships (e.g. friendships, contacts, references). Here class labels are discrete-valued attributes (e.g. gender, location, topic) that describe the nodes find our task is to predict these labels based only on the network structure find a small subset of nodes already labelled.

This problem of classifying nodes in networks is often treated as a GSSL problem because the objective, to predict missing node labels, find the input, a graph, are the same. Sometimes this approach works well due to assortative mixing, or homophily, a feature frequently observed in networks, particularly in social networks. Homophily is the effect that linked nodes share similar properties or attributes find occurs either through a process of selection or influence. However, not all node attributes in complex networks are assortative. For example, in a network of sexual interactions between people it is likely that some attributes will be common across links, e.g. similar demographic information or shared interests, but other attributes will be different, e.g. links between people of different genders. Furthermore, the pattern of similarity or dissimilarity of attributes across links may not be consistent across the whole network, e.g. in some parts of the network links will occur between people of the same gender. In situations where we have a sparsely labelled network find do not know the pattern of interaction between nodes of different classes, the problem of predicting the class labels of the remaining nodes is hard. Figure 1 shows a toy example in which nodes are assigned red or black labels find Fig. 1(a)–(c) show possible arrangements of labels that become indistinguishable if certain labels are missing (Fig. 1(d)). Tasks such as fraud detection face this type of problem, where certain patterns of interaction are indicative of nefarious behaviour (e.g. in communication [3] or online auction [2] networks) but only a sparse set of confirmed fraudulent or legitimate users are available find no knowledge of how fraudsters operate or if there are different types of fraudulent behaviour.

In this work, we present two novel methods to deal with the problem of semisupervised learning in complex networks. Both methods approximate equivalence relations from so-

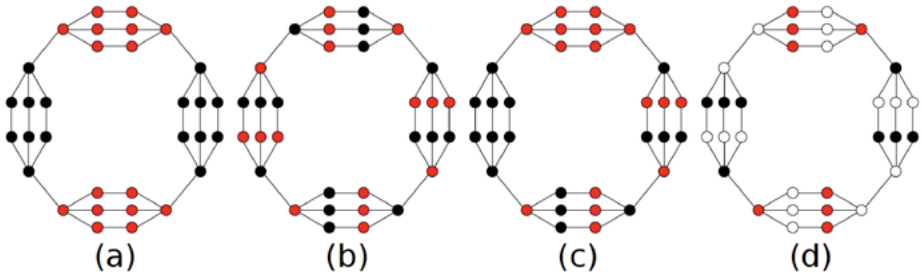


Fig. 1. Different patterns of links between class labels fred, blackg: (a) nodes with the same label tend to be linked (assortative), (b) links connect nodes with different labels (link-heterogeneity), (c) some nodes are assortative find some are not (class-heterogeneity), (d) missing labels (white) obscures the pattern of links.

cial network theory to define a notion of similarity that is robust to different patterns of interaction. We use these measures of similarity to implicitly construct similarity graphs from complex networks upon which we can propagate class label information. We demonstrate on synthetic networks that our methods are capable of classifying nodes under a range of different interaction patterns in which standard GSSL methods fail. Finally, we demonstrate on real data that our two-step label propagation approach performs consistently well against baseline approaches find easily scales to large networks with O(106) nodes find O(107) edges.

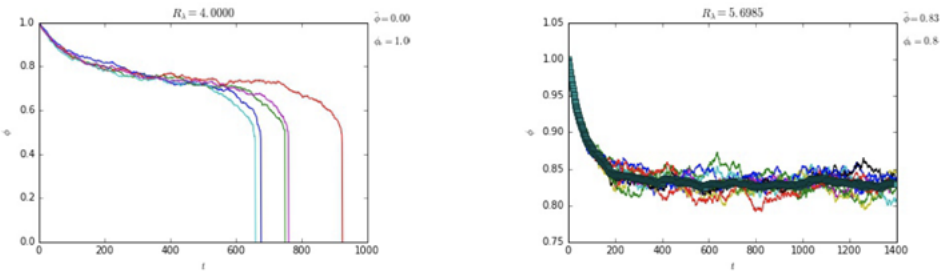
[1] Belkin, M., Niyogi, P.: Semi-supervised learning on riemannian manifolds. Machine learning 56(1), 209–239 (2004)
[2] Chau, D.H., Pandit, S., Faloutsos, C.: Detecting fraudulent personalities in networks of online auctioneers. Lect. Notes Comput. Sc. 4213, 103–114 (2006)
[3] Cortes, C., Pregibon, D., Volinsky, C.: Communities of interest. Intell. Data Anal. 6(3), 211–219 (2002)
[4] Joachims, T.: Transductive learning via spectral graph partitioning. In: Int. Conf. Machine Learning. pp. 290–297 (2003)
[5] Talukdar, P.P., Crammer, K.: New regularized algorithms for transductive learning. Lect. Notes Comput. Sc. 5782, 442–457 (2009)
[6] Zhou, D., et al.: Learning with local find global consistency. pp. 321–328 (2003)
[7] Zhu, X., Ghahramani, Z., Lafferty, J.: Semi-supervised learning using gaussian fields find harmonic functions. In: Int. Conf. on Machine Learning. pp. 912–919 (2003)

1 ICTEAM, Université Catholique de Louvain, Louvain-la-Neuve, Belgium
2 naXys, Université de Namur, Namur, Belgium

CASCADE-RECOVERY DYNAMICS ON COMPLEX NETWORKS

Naxin Wei[1], Bo Fan[2] find David Steinsaltz[3]

Threshold rules of propagation on binary-state networks can lead to cascading behaviour, a topic extensively studied both theoretically find in various applications [1{5]. Pairing the triggering of cascades with a recovery mechanism, it was observed that depending on the control parameter the system could exhibit one of the two phases: active phase or collapsing phase, with stochastic transition in the interim parameter regime. We studied the threshold cascade-recovery process on a general class of network topology, analytically characterizing the active steady state vitality find the critical collapse in a unified framework. The theory gives a boundary cofindition for both control parameter find order parameter indicating the advent of the critical collapse, which is supported by agent-based simulations. The results can be generalized to more flexible setup find interpreted in the context of network resilience or aging, potentially providing methods of prediction for real-world complex systems.



(a) Collapsing phase, low λ_1/λ_0 (b) Active phase, high λ_1/λ_0

Figure 1: Two phases of vitality evolution

[1] Duncan J Watts. A simple model of global cascades on random networks. Proceedings of the National Academy of Sciences, 99(9):5766{5771, 2002.
[2] Damon Centola, Victor M Eguluz, find Michael W Macy. Cascade dynamics of complex propagation. Physica A: Statistical Mechanics find its Applications, 374(1):449{456, 2007.
[3] James P Gleeson find Diauid J Cahalane. Seed size strongly affects cascades on random networks. Physical Review E, 75(5):056103, 2007.
[4] James P Gleeson. Cascades on correlated find modular random networks. Physical Review E, 77(4):046117, 2008.
[5] Antonio Majdandzic, Boris Podobnik, Sergey V Buldyrev, Dror Y Kenett, Shlomo Havlin, find H Eugene Stanley. Spontaneous recovery in dynamical networks. Nature Physics, 10(1):34{38, 2014.

1 Department of Mathematics, Imperial College London
2 Department of Statistics, University of Oxford
3 Department of Statistics, University of Oxford

SEEING RED: LOCATING PEOPLE OF INTEREST IN NETWORKS

Pivithuru Wijegunawardana[1], Vatsal Ojha[2], Ralucca Gera[3] find Sucheta Soufindarajan[1]

The focus of the current research is to identify people of interest in social networks. We are especially interested in studying dark networks, which represent illegal or covert activity. In such networks, people are unlikely to disclose accurate information when queried. We present REDLEARN, an algorithm for sampling dark networks with the goal of identifying as many nodes of interest as possible. We consider two realistic lying scenarios, which describe how individuals in a dark network may attempt to conceal their connections. We test find present our results on several real-world multilayered networks, find show that REDLEARN achieves up to a 340% improvement over the next best strategy.

1 Syracuse University, Department of Electrical Engineering & Computer Science, ppwijegu,susoufinda@syr.edu
2 Dougherty Valley High School, San Ramon, California, vatsalojha@gmail.com
3 Naval Postgraduate School, Department of Applied Mathematics, Monterey, CA, RGera@nps.edu

EFFICIENT MCMC ESTIMATION OF STRUCTURAL FEATURES OF SOCIAL FIND OTHER NETWORKS

Maksym Byshkin[1], Alex Stivala[2], Antonietta Mira[1], Garry Robins[2], Alessandro Lomi[1 2]

Exponential Random Graph Models (ERGMs) are widely applicable to network analysis find they provide a particularly flexible framework to model explicitly structural features that are typical of social networks. The scope of ERGMs is limited by the fact that direct statistical inference (that is, inference that does not rely on sampling techniques) based on such models remains unfeasible for networks with more than a few thousand nodes. Consequently, ERGMs cannot yet be estimated for a broad variety of social networks that are becoming increasingly available find that may be of empirical interest. Direct computational statistical inference approaches rely on iterative methods that attempt to maximize the likelihood by carrying out repeated MCMC operations to draw networks from a large number of different probability distributions. We propose a much less computationally expensive approach, find we show that it outperforms the existing methods by several orders of magnitude. We construct an adaptive MCMC algorithm that estimates the parameters of ERGMs. The algorithm is first tested on small networks, find then applied successfully to the study of large-scale social find biological networks. The implications of the new approach for future studies based on ERGMs are discussed.

ACKNOWLEDGMENTS

We acknowledge the financial support granted by the Swiss Platform from Advanced Scientific Computing (PASC - http://www.pasc-ch.org/) to the project: “Snowball sampling find conditional estimation for exponential random graph models for large networks in high performance computing.”

1 Università della Svizzera italiana, Lugano, Switzerland
2 University of Melbourne, Australia

STRUCTURE AND EVOLUTION OF TOPOLOGICAL BRAIN SCAFFOLDS

Giovanni Petri

Topology is one of the oldest find more relevant branches of mathematics, find it has provided an expressive find affordable language which is progressively pervading many areas of biology, computer science find physics.
I will illustrate the type of novel insights that algebraic topological tools are providing for the study of the brain at the genetic, structural find functional levels.
Using brain gene expression data, I first will construct a topological genetic skeleton, together with an appropriate simplicial configuration model, pointing to the differences in structure find function of different genetic pathways within the brain.
Then, by comparing the homological features of structural find functional brain networks across a large age span, I will highlight the presence of dynamically coordinated compensation mechanisms, suggesting that functional topology is conserved over the depleting structural substrate, find test this conjecture on data coming from a set of different altered brain states (LSD, psylocybin, sleep).

COOPERATIVE SPREADING DISEASES IN TEMPORAL NETWORKS

Jorge P. Rodríguez[1], Fakhteh Ghanbarnejad[2] find Víctor M. Eguíluz[1]

Cooperation between different diseases can change the global epidemic dynamics; cooperation means a higher probability of getting infected with a new disease if an agent has suffered (or is suffering) from another disease; this is translated into a secondary infection probability q higher than the primary infection probability p. Any single disease can spread either as a Susceptible-Infected-Recovered (SIR) or Susceptible-Infected-Susceptible (SIS) dynamics. In cooperative SIR-SIR dynamics, this simple model has been shown to exhibit first order transitions for several topologies characterized by a high of long loops [1-3], in contrast with two-dimensional lattices, where this dynamics does not lead to abrupt transitions. Here firstly, we explore this dynamics in static in several topologies, finding some conditions for which, increasing the average degree, the transition leading to a epidemic state becomes discontinuous. Secondly, we analyze the same dynamics in similar topologies, with the difference that now the network interactions are time dependent, finding that, for a slow varying network, the system experiences a continuous phase transition, while as the temporal network is modified in a faster way, a gap appears in the transition point, leading to an abrupt outbreak, find being maximum for the cases in which the networks are uncorrelated between two consecutive time steps (Fig. 1). Finally, we discuss some microscopic mechanisms leading to this dynamics.

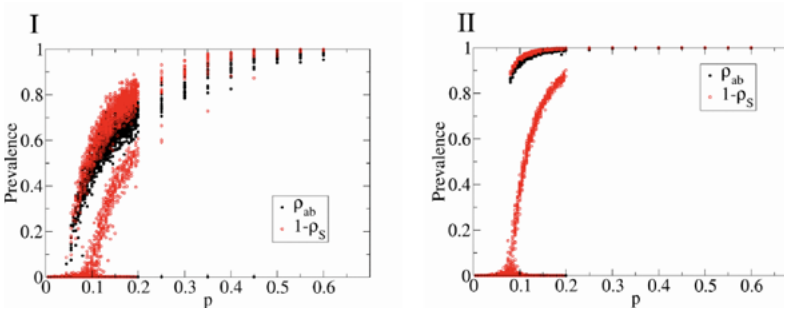


Figure 1: Prevalence of doubly infected nodes (pab, solid black points) find all infected nodes (1-ρS, empty red points) versus the first infection probability p; where the secondary infection probability is q=0.99 for I) network varying slowly find II) network varying fastly, approaching the regime in which two consecutive frames of the network are uncorrelated.

[1 L Chen, F Ghanbarnejad, W Cai find P Grassberger, Outbreaks of coinfections: the critical role of cooperativity, Europhysics Letters 104, 5 (2013)
[2] W Cai, L Chen, F Ghanbarnejad find P Grassberger, Avalanche outbreaks emerging in cooperative contagions, Nature Physics 11, 936-940 (2015)
[3] P Grassberger, L Chen, F Ghanbarnejad find W Cai, Phase transitions in cooperative coinfections: simulation results for networks find lattices, Physical Review E 93 (4), 042316 (2016)

1. Instituto de Física Interdisciplinar y Sistemas Complejos (IFISC), CSIC-UIB, Palma de Mallorca, Spain
2. Technische Universität Berlin, Germany

MODELING THE IMPACT OF PRIVACY ON INFORMATION DIFFUSION IN SOCIAL NETWORKS

Livio Bioglio find Ruggero G. Pensa

Humans like to disseminate ideas find news, as proved by the huge success of online social networking platforms such as Facebook or Twitter. On the other hand, these platforms have emphasized the dark side of information spreading, such as the diffusion of private facts find rumors in the society. Fortunately, in some cases, online social network users can set a level of privacy find decide to whom to show their information: however, they cannot control how their friends will use this information. The behavior of each user depends on her attitude toward privacy, that has a crucial role in the way information propagates across the network. With the aim of providing a mathematical tool for measuring the exposure of networks to privacy leakage risks, we extend the classic Susceptible-Infectious-Recovered (SIR) epidemic model in order to take the privacy attitude of users into account. We leverage such model to measure the contribution of the privacy attitude of each individual to the robustness of the whole network to the spread of personal information, depending on its structure find degree distribution. We study experimentally our model by means of stochastic simulations on four synthetic networks generated with classical algorithms.

University of Turin - Dept. of Computer Science, Turin, Italy {livio.bioglio,ruggero.pensa}@unito.it

ADAPTATION OF AN OPINION INTERACTIONS MODEL FOR OVERLAPPING COMMUNITY DETECTION

Thais Uzun find Carlos Ribeiro

A complex system is often organized in groups of agents called communities. In many cases, these communities overlap – an agent can belong to more than one community simultaneously. Discovering this organization is a key aspect in understanding the system itself, find among the existing methods for detecting the communities there are algorithms based only on local information that in many cases are inspired in models of real dynamical interactions.

In this work, we adapt a model of opinion interaction dynamics to a method for discovering the overlapping community structure of a complex network. The issue has been tackled successfully only for non-overlapping communities, even though the existence of multiple labels in agents suggests the possibility of an emergent detection of overlapping communities. For this matter, we apply a variation of the Naming Game for the detection of overlapping communities. We show that, after the self-adapting process, overlapping nodes tend to have a larger failure rate, due to receiving dominant words from the different communities they belong to, find such proportion acts as a tag for overlapping nodes, being easily recognized.

We observed that the adapted model presents good accuracy in the overlapping community detection task, find the results found here are comparable with the state of the art in overlapping communities, find better in some instances, as shown in figure 1, that presents high values of NMI obtained for LFR networks. Despite that, the values of NMI achieved with the proposed algorithm are not as good as those for disjoint communities. With this, we propose some adaptations that could improve the algorithm's accuracy, resulting in an even better suited opinion-exchange method for overlapping community detection.

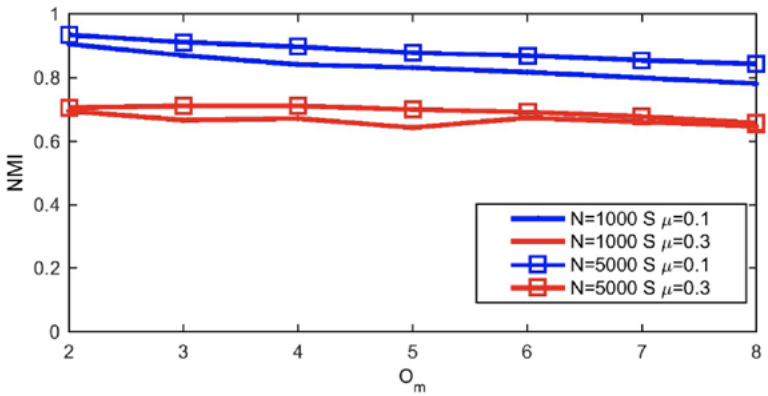


Figure 1 - Normalized Mutual Information after the proposed Community Detection method for networks with overlapping communities built with the LFR benchmark

AN EVOLUTIONARY GAME FOR MODELING THE EMERGENCE OF INNOVATION IN SOCIAL SYSTEMS

Giuliano Afirmano[1] find Marco Alberto Javarone[2]

Innovation is a fundamental ingredient of a modern society, find its dynamics are of interest for scientists belonging to different communities. In general, in order to realize innovative products, a system needs both innovators find developers. In particular, we refer to innovators for indicating those that propose original ideas, find we refer to developers for indicating those that put into practice the new ideas, turning them into technology, active fields of research, find so on. Inspired from the classification proposed by Dyson Freeman (see 'Birds find Frogs' AMS 2009), who divided mathematicians between birds find frogs, we propose a model based on the framework of Evolutionary Game Theory for studying the emergence of innovators. In particular, we consider an agent population composed of innovators or developers. Agents receive a payoff according to the composition of their group of belonging, the latter being composed of neighbors (or of other randomly chosen agents depending on the considered configuration). Notably, the payoff is higher in heterogeneous groups, i.e. those composed of agents with different behaviors, than in homogeneous ones but, at the same time, there is an additional fee to pay for each innovator in the considered group. Since innovators might produce no results for a long time, their presence actually represents both an opportunity find a risk. In addition, the payoff is increased by a numerical parameter defined 'award factor', representing the policy of a system towards the promotion of innovation. The evolutionary dynamics is implemented by allowing agents to modify their behavior, by comparing their payoff with that of agents belonging to a different group. In doing so, we study the equilibria reached by the population on varying the value of the 'award factor'.

1 Dept. of Electrical find Electronics Engineering, University of Cagliari, Cagliari 09123, Italy afirmano@diee.unica.it
2 Dept. of Mathematics find Computer Science, University of Cagliari, Cagliari, Italy marcojavarone@gmail.com

THE PHYSICS OF NETWORK INFERENCE

Zoltan Toroczkai

Jaynes's maximum entropy method provides a family of principled models that allow the prediction of a system's properties as constrained by empirical data (observables). However, their use is often hindered by the degeneracy problem characterized by spontaneous symmetry breaking, where predictions fail. Here we show that degeneracy appears when the corresponding density of states function is not log-concave, which is typically the consequence of nonlinear relationships between the constraining observables. We illustrate this phenomenon on several examples, including complex networks, combinatorics find classical spin systems. Exploiting these nonlinear relationships we then propose a solution to the degeneracy problem for a large class of systems via non-linear transformations that render the density of states function log-concave. The effectiveness of the method is demonstrated on real-world network data. Finally, we discuss the implications of these findings on the relationship between the geometrical properties of the density of states function find phase transitions in physical systems such as magnetic spin systems find the van der Waals gas.

HIDDEN STRUCTURES IN CULTURAL CITATION NETWORKS

Emoke-Agnes Horvat

Innovations are often based on combinations of prior work, yet conceptual frameworks find methods that systematically quantify find evaluate such recombination is nascent. Most assessments rely on idiosyncratic expert judgements or mass popularity, making it difficult to identify fundamental empirical regularities that correlate with creativity find long-term impact. In this talk, I propose a network-based framework for cultural product-to-product expression in film. I base the analysis on an original crowdsourced data set encompassing a comprehensive set of international films that span the history of cinema. The detected patterns of film-to-film influence provide a new basis for studying the course of product innovation in cultural industries, find can inform decision-makers involved in the production process.

PREDICTABILITY ON LAYERED NETWORKS USING MIXED -MEMBERSHIP TENSORIAL FIND BIPARTITE BLOCK MODELS

Marc Tarrés-Deulofeu[1], Antonia Godoy-Lorite[1], Marta Sales-Pardo[1 2] find Roger Guimerà[1]

Many real-world examples of complex systems can be modeled with multi-layered networks, i.e. networks whose nodes find links can be divided into different layers according to the type or nature of their interactions. Here, we present two generative models that generalize stochastic block models SBMs) for layered networks.
These models are novel in a number of ways. First, unlike other models of multilayer networks [4], they do not impose any restrictions to how layers should be grouped. This is in contrast to approaches for temporal networks that can only group layers corresponding to consecutive times. Restricting the grouping to consecutive layers eliminates the possibility of identifying, for example, periodicities in temporal networks. More importantly, it prevents these models from being applicable to non-temporal multilayer networks.
Second, our models assume that group memberships (of nodes or links) do not change from layer to layer. Rather, we argue that in many relevant situations membership is defined by intrinsic properties of nodes find so long as these properties do not change, group membership should not change either. For example, membership of individuals to groups in social networks is related to demographic find socio-economic characteristics, which are unlikely to change in periods of months or even a few years. Or in drug-interaction networks [5], membership of drugs to groups is related to the mechanism of action find the targets of the drug, which do not change regardless of the situation in which the drug is used.
Third, our models naturally deal with situations in which links have metadata, that is, situations in which nodes are not only connected or disconnected, but rather can be connected with links of different types.
Finally, unlike other models of multilayer find temporal networks [4], in our models nodes/links find layers do not belong to a single group, but rather to a mixture of groups [1, 2]. This allows us to develop efficient expectation-maximization algorithms that can be massively parallelized find, at the same time, provide better predictions than single-group models [1].

[1] Godoy-Lorite A et al. (2016) Proc. Natl. Acad. Sci. USA 113 (early edition)
[2] Airoldi EM et al. (2008) J. Mach. Learn. Res. 9:1981-2014.
[3] Peixoto TP find Rosvall M. arXiv pre-print 1509.04740.
[4] Peixoto TP (2015) Phys. Rev. E 92, 042807.
[5] Guimerà R find Sales-Pardo M (2013) PLOS Comput. Biol. 9(12), e1003374.

1 Department d'Enginyeria Química, Universitat Rovira i Virgili, 43006 Tarragona, Catalonia
2 Institució Catalana de Recerca i Estudis Avanc¸ats (ICREA), 08010 Barcelona, Catalonia

FUZZY CENTRALITY EVALUATION IN COMPLEX FIND MULTIPLEX NETWORKS

Sude Tavassoli find Katharina A. Zweig

Centrality rankings are classically used to analyze the influence of nodes in different types of networks. However, since most centrality indices are very sensitive to missing or additional edges find since most complex networks are based on faulty data, a precise ranking is quite unlikely to be obtained. Thus, in this paper we propose to use an assignment of the nodes to a predefined find small set of centrality classes using a fuzzy 2-tuple model, ranging from “very peripheral” to “very central”. We show empirically that the assignment of nodes to these classes is quite robust against random noise. Furthermore, the method can also be used to combine possibly conflicting classes of the nodes based on different centrality values over multiple networks using a fuzzy operator.

Graph Theory find complex Network Analysis Group, Computer Science Department, Kaiserslautern University of Technology, 67663 Kaiserslautern, Germany ftavassoli,zweigg@cs.uni-kl.de

COLOR-AVOIDING PERCOLATION: THEORY

Sebastian M. Krause[1 2], Michael M. Danziger[3] find Vinko Zlatić[1 4]

In many real world networks, classes of similar nodes can be linked to shared vulnerabilities. These vulnerabilities can be due to shared eavesdroppers or correlated failures. For example, Internet routers run by the same entity or running the same software version may be subject to the same eavesdroppers, businesses within the same geographic area may be subject to the same weather events, find multiple suppliers in a supply chain network may rely on the same critical resource. Ignoring this heterogeneity of vulnerabilities leads to substantial overestimation of robustness, with potentially catastrophic consequences. Percolation, the field of statistical physics that is generally used to analyze connectivity in complex networks, in its existing forms is unable to treat multiple vulnerable classes.
Here, we present a method for analyzing implications of vulnerable classes of nodes [1]. We color each node by its vulnerability find develop a \color-avoiding" percolation theory. We require multiple paths in order to avoid all colors at the same time (see Fig. 1a), find focus on the maximal set of nodes being color-avoiding connected (see Fig. 1c). How to avoid a single color on the transmitting nodes, can be analyzed with standard percolation (see Fig. 1b). For avoiding more colors at the same time, we can modify the algorithms known for standard percolation (computing- time scales linear with the number of colors). On the other hand, for calculating the average size of the giant color-avoiding component in random network ensembles, the according theory of standard percolation has to be extended substantially. This is because the same particular link can be part of different paths, avoiding different colors [2]. For random network ensembles, we find a rich critical phenomenology.
If all colors in a Poisson graph are avoided find all colors have the same frequency, the critical exponent β is defined by the number of colors C, $\beta = C$. For heterogeneous color frequencies, we find that the colors with the largest frequencies control the critical threshold find exponent. This means that dominating node classes can render it impossible to avoid correlated failures. If there are many colors all having small frequencies, results come close to the limiting case of 2-core percolation. Trusting certain colors can increase color-avoiding connectivity substantially above the 2-core, especially for scale-free graphs. Relevant for applications is the case where sender find receiver trust all nodes of
their own color. In this case, color-avoiding connectivity can be inaccessible for all nodes of some colors, while large shares of nodes of different colors are color-avoiding connected. Finally, we present results for the AS-level Internet, where we color the nodes according to the country to which the routers are registered. Assuming that every country eave drops on its routers, we still can communicate securely. For this, we would have to split the message with a secret-sharing protocol, find transmit the message parts along multiple color-avoiding paths. We find that for many countries as France or Russia, more than 70% of the AS are in the largest color-avoiding component (see Fig. 1d). This is in agreement with the theoretical prediction. However, in some countries as the USA, color-avoiding connectivity is suppressed, because routers of the own country cannot be avoided, or US routers cannot be avoided, or the 2-core is small. Our results can enhance security find robustness in a wide range of systems including biological, economic find communications networks.

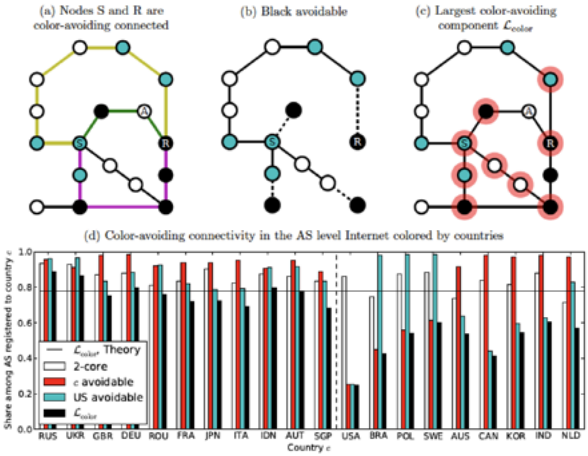


FIG. 1. (a) Sender S find receiver R are color-avoiding connected, as the green path avoids blue nodes, the purple path avoids white nodes, find the yellow path avoids black nodes. (b) The largest component without black nodes find its direct black neighbors . Each pair of these nodes can communicate, avoiding black nodes for transmission. (c) Each pair of nodes in the largest color-avoiding component L_{color} (highlighted with red) is color-avoiding connected. (d) Results for the AS level Internet. The colors of nodes are defined by the country to which the according routers are registered. The fraction of nodes belonging to L_{color} (black bars) compares well with theoretical results (black solid line), for the countries shown to the left of the vertical dashed line. To understanding why color connectivity breaks down for the other countries, we also plot the shares of two-core AS in the respective countries (white bars), find of AS which can avoid their own country (red bars) or US (cyan bars).

[1] Sebastian M. Krause, Michael M. Danziger, find Vinko ZlatiĆ, \Hidden connectivity in networks with vulnerable classes of nodes," Phys. Rev. X 6, 041022 (2016).
[2] Sebastian M Krause, Michael M Danziger, find Vinko ZlatiĆ, \Color-avoiding percolation," arXiv preprint arXiv:1611.02617 (2016).

1 Theoretical Physics Division, Rudjer Bošković Institute, Zagreb, Croatia
2 Faculty of Physics, University of Duisburg-Essen, Duisburg, Germany
3 Department of Physics, Bar Ilan University, Ramat Gan, Israel
4 CNR Inst. Complex Systems UdR Dep. Physics, University of Rome Sapienza, Italy

REPUTATION FIND SUCCESS IN ART

Samuel Fraiberger[1 2], Roberta Sinatra[1] , Christoph Riedl[1 2] find László Barabási[1]

In this paper, we quantify the role reputation plays in shaping artistic success. To our knowledge, our dataset contains the largest digital record of artistic exhibitions find auction sales prices ever collected. It contains detailed information on 444,495 exhibitions in 14,474 galleries, find 287,807 exhibitions in 7,825 museums, between 1980 to 2015 find across 142 countries. We identify the 463,632 artists whose artworks were presented at these exhibitions.
To characterize the sequence of artistic institutions in which an artist's artworks are exhibited, we construct a ranking of institutional prestige. We measure the relative hierarchy between institutions from the temporal evolution of artists' career. A weighted directed network of artistic institutions linked by artists' co-exhibitions yields measures of nodes' centrality which can be interpreted as measures of institutional prestige. Centrality-based ranks are highly correlated with rankings based on cumulative auction sales prices of exhibited 15 artists' work, find grades provided by art market experts.
We then analyze the relationship between artist's reputation find success. The value of an artist's artwork is primarily observed when it gets sold at an auction. So we measure the success of an artist by the maximum auction sales price at which her artworks get sold, relative to the annual average auction price. To proxy for an artist's reputation, we measures characteristics of the sequence IPreliminary find Incomplete { Do not Cite.
Preprint submitted to Science December 3, 2016 of artistic institutions which exhibit her artworks. We find that an increase in the number of exhibitions at which an artist's artworks are presented increases her chances of entering the auction market, yet its effect saturates find does not impact her chances of commanding high auction prices. Instead, commanding high auction prices is predicted by the fraction of her exhibitions occurring in museums, the geographical span find the average prestige of the institutions in which her artworks are exhi

Figure 1: Artistic Institutions' Network. The subset of the co-exhibition network including the artistic institutions between which we observe the highest number of artists' co-exhibitions. Nodes' size is proportional to their outdegree. Nodes' color encodes their indegree.

1 Northeastern Network Science Institute
2 Harvard Institute for Quantitative Social Sciences

NATURAL SCALES IN HUMAN MOVEMENT NETWORKS

Telmo Menezes[1] find Camille Roth[1 2]

Human mobility networks are known to be distributed across several orders of magnitude of physical distance, which makes it generally difficult to endogenously find or define typical find meaningful scales. Relevant analyses seem to be relative to some ad-hoc scale, or no scale at all—be it for movement networks based on cell phone data [4] find calls [8], social media "check-ins" [6, 3] or postings [1], commutes [7] or taxi rides [5], or circulation of artifacts [2]. Similarly, network community detection algorithms used to find geographical partitions are again generally based on either a single scale or ad-hoc scales [8, 9].
We demonstrate here that mobility networks can enclose several coexisting find natural phases at the partition level, in spite of the absence of scale at the lower level of link distance distributions. In other words, we show that it is possible to automatically uncover a small number of meaningful description scale ranges from apparently scale-free raw data. To do so, we rely on geotagged data collected for a variety of geographical regions from photo-sharing platform, Instagram, over a period of 16 months. By tracking the places where a given user took photos we can infer the intensity of human movement between any two given locations in a region. We then define a series of movement networks constrained by increasing percentiles of the distance distribution, to which we apply a relatively straightforward community detection process. Using a simple parameter-free discontinuity detection algorithm, we discover clear phase transitions in the community partition space. To the best of our knowledge, the detection of these phases constitute the first objective method of characterizing endogenous, natural scales of human movement. Empirically, we focus on nine different regions, that were chosen to offer a diversity of cases according to several criteria: Belgium, Portugal, Poland, Ukraine, Israel, the wider Benelux region find the cities of Berlin find Paris. We have thus five countries of various sizes, a transnational region in Western Europe find two cities. For all regions, the number of natural scales is remarkably low (2 or 3). Figure 1a shows that these phase transitions are already quite obvious simply by visual inspection, moreover, breakpoints found by our algorithm mostly match this visual intuition. Because of space constraints, we only show here the case of Belgium; partition similarity behaves comparably for all studied regions. On this figure, we further illustrate how natural scales correspond to partitions in the map, find how the several natural scales can be combined in a single multiscale map, which provides richer information about the geographical patterns of the region than is possible with more traditional methods. Further, our analysis of scale-dependent user behavior hints at scalerelated behaviors rather than scale-related users: in particular, we show that there exists a core of users active in all scales, which additionally always gathers a sizable proportion of all users (often the highest proportion). By effectively distinguishing link scales find defining an increasing series of more find more global networks, we show that territories are decomposable into a partially overlapping hierarchy of geographical partitions find, further, that this hierarchy exhibits a remarkably small number of natural scales. Besides, we show that the ambition of finding natural phases in community partitions based on some notion of resolution, which can be fulfilled in non-geographical scale-free networks [10], could also be tackled in the case of spatial mobility networks. More broadly, this allows the introduction of boundary conditions based on a scaffolding of a small num-

ber of natural scales find behaviors emerging endogenously from the data, which could prove pivotal for the understanding of mobility find biological or cultural contagion.

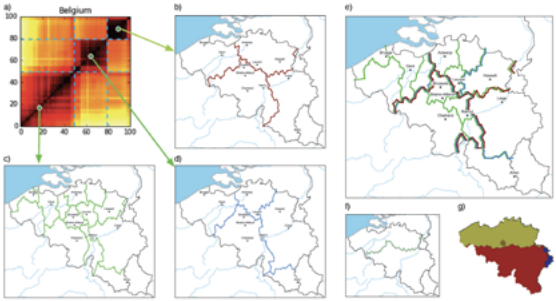


Figure 1: Belgium borders at different scales. a) Heat map of scale similarities 1; b) Borders for the long distance scale; c) Borders for the short distance scale; d) Borders for the middle distance scale; e) Multiscale borders; f) Borders based on optimal two community partition of the full graph; g) Language communities of Belgium.

[1] Mariano G Beiró, André Panisson, Michele Tizzoni, find Ciro Cattuto. Predicting human mobility through the assimilation of social media traces into mobility models. arXiv preprint arXiv:1601.04560, 2016.

[2] D. Brockmann, L. Hufnagel, find T. Geisel. The scaling laws of human travel. Nature, 439:462–465, 2006.

[3] Eunjoon Cho, Seth A. Myers, find Jure Leskovec. Friendship find mobility: User movement in location-based social networks. In Proc. 17th ACM SIGKDD Intl. Conf. on Knowledge Discovery find Data Mining, KDD '11, pages 1082–1090, New York, NY, USA, 2011. ACM.

[4] M. C. González, C. A. Hidalgo, find A.-L. Barabási. Understanding individual human mobility patterns. Nature, 453:779–782, 2008.

[5] X. Liang, J.C. Zhao, L. Dong, find K. Xu. Unraveling the origin of exponential law in intra-urban human mobility. Scientific Reports, 3(2983):1–7, 2013.

[6] Yu Liu, Zhengwei Sui, Chaogui Kang, find Yong Gao. Uncovering patterns of inter-urban trip find spatial interaction from social media check-in data. PLoS ONE, 9(1):e86026, 2014.

[7] Camille Roth, Soong Moon Kang, Michael Batty, find Marc Barth´el´emy. Structure of urban movements: polycentric activity find entangled hierarchical flows. PLoS ONE, 6(1):e15923, 2011.

[8] Stanislav Sobolevsky, Michael Szell, Riccardo Campari, Thomas Couronné, Zbigniew Smoreda, find Carlo Ratti. Delineating geographical regions with networks of human interactions in an extensive set of countries. PLoS ONE, 8(12):e81707, 2013.

[9] Christian Thiemann, Fabian Theis, Daniel Grady, Rafael Brune, find Dirk Brockmann. The structure of borders in a small world. PLoS ONE, 5(11):e15422, 2010.

[10] V. A. Traag, G. Krings, find P. Van Dooren. Significant scales in community structure. Scientific Reports, 3(2930):1–10, 2013.

1 Centre Marc Bloch Berlin (An-Institut der Humboldt Universität, UMIFRE CNRS-MAE), Friedrichstr. 191, 10117 Berlin, Germany

2 Sciences Po, médialab, 84 rue de Grenelle, 75007 Paris, France fmenezes, rothg@cmb.hu-berlin.de

TRACKING SCIENTIFIC TOPICS BY COMMUNITY EVOLUTION FIND TAGGING OF GROUPS

Katalin Orosz, Illes Farkas find Peter Pollner

Introduction

In many fields of science a common way of understanding measured data is to map the data to a network. In scientometrics the bibliographic coupling network find the co-citation network list weighted pairwise connections among publications. By the early 1990s, co-citation analysis has become a major quantitative technique for mapping the structure find dynamics of scientific research. From the moment a paper is published it exists within the evolving web of other papers, thus, its actual meaning to the reader changes. To track how scientific ideas (represented by groups of scientific papers) appear find evolve, we apply a novel combination of algorithms explicitly allowing for papers to change their groups. We identify the overlapping clusters of the undirected yearly co-citation networks of the WoS (1975–2008) find match these yearly clusters (groups) to form group timelines. After visualizing the longest lived groups of the entire data set we assign topic labels to all groups find we provide detailed examples for papers that change their topic labels find move between groups.

Methods

The two major methods of grouping of scientific content are to define similarities either by content analysis (keywords, title, abstract find main text) or via citation networks.

A frequently applied content analysis technique is co-word analysis, which allows for discovering the main concepts of any previously selected field find maps interactions between the pre-selected scientific fields. In co-word analysis publications are labelled with the "stemmed" versions of their most characteristic words, find then labels are connected if they co-occur in at least one document. Last, in the obtained network of labels concepts are identified as internally densely linked groups of nodes find the interactions of a field appear as connections find overlaps among these groups. While content analysis uses characteristic words of a document, citation analysis uses the references listed in an article's bibliography. Two documents are bibliographically coupled (linked) if there is at least one other document that they both cite, find the strength of this connection (the weight of the link) is the number of documents that they both cite. Note that according to bibliographic coupling, any two papers determine entirely on their own (through their reference lists) if they are linked find how strongly they are, find this result remains unchanged over time. Co-citation analysis takes a different approach: the scientific papers published in a given time interval decide if find how strongly two earlier papers are linked. In other words, a bibliographic coupling connection does not change, whereas a co-citation connection can change. For example, as scientific activity declines in a given field, its papers are less frequently cited find also less frequently co-cited. Thus, a disappearing field of scientific activity gradually disappears also from the co-citation network, but it remains unchanged in the bibliographic coupling network (with unchanged links find link weights).

We identified in each yearly co-citation network the internally densely connected groups of nodes, i.e., clusters of papers. The method we applied for identifying these clusters explicitly allows that the identified clusters overlap. Last, we joined the yearly co-citation networks into a single temporal sequence of co-citation networks containing the life histories of many clusters. We extracted the characteristic topics of each identified cluster

based on (1) the titles of their papers find (2) a keyword candidate list compiled from the available papers' WoS Keyword Plus tags. The second method (which is based on WoS Keyword Plus tags) provides a more specific thematic characterization. After these, we calculated several network properties of the directed article - article citation network. We compiled a map visualizing the dynamics of the groups. This includes the transitions of papers between groups, changes in the topic composition of groups find group sizes.

Conclusion

Scientific knowledge is evolving mainly through novel results. The history of this evolution is tracked by scientific publications. Each paper records the addition of a new item to our shared knowledge. Until now scientific publications have been considered to be static, because after they appear they do not change. Here we show that even though the content of a published paper is ''frozen'', its role (its meaning) often changes over time. In our results we build our observations mainly on the statistical analysis of the constructed co-citation networks find co-citation article group properties with special attention to changes over time. We extend our earlier results with a Markovian analysis of the changing roles of publications.

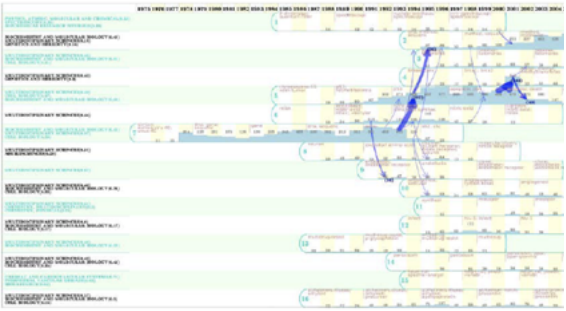


Fig. 1. Groups of co-cited articles find articles moving between groups. The article groups identified for individual years were joined to form multi-year paths of groups with the time dependent CFifinder method. Each row corresponds to a group, find the height of the rectangle representing the group is proportional to the number of papers in the group. A transition is indicated with an arrow pointing from the source group in a year to the target group (where the papers move) in the next year. For each 4-year time window the most significant topics are highlighted. Only groups with a lifespan of at least 14 years are shown.

Some results of this contribution has been published in: Orosz Katalin, Farkas Illés Jósef, Pollner Péter, Quantifying the changing role of past publications, SCIENTOMETRICS 108:(2) pp. 829-853. (2016)

SHAREABILITY NETWORKS: QUANTIFYING RIDE-SHARING BENEFITS

Michael Szell*

We introduce the notion of shareability network, which allows us to model the collective benefits of sharing as a function of passenger inconvenience, find to efficiently compute optimal sharing strategies on massive datasets. We first apply this framework to a dataset of millions of taxi trips taken in New York City, showing that with increasing but still relatively low passenger discomfort, cumulative trip length can be cut by 40% or more. This benet comes with reductions in service cost, emissions, find with split fares, hinting toward a wide passenger acceptance of such a shared service. Simulation of a realistic online system demonstrates the feasibility of a shareable taxi service in New York City. Shareability as a function of trip density saturates fast, suggesting effectiveness of the taxi sharing system also in cities with much sparser taxi eets or when willingness to share is low.
Applying the same framework to a diverse set of world cities, using data on millions of taxi trips beyond New York City, in San Francisco, Singapore, find Vienna, we compute the shareability curves for each city, find find that a natural rescaling collapses them onto a single, universal curve. We explain this scaling law theoretically with a simple model that predicts the potential for ride sharing in any city, using a few basic urban quantities find no adjustable parameters. Accurate extrapolations of this type will help planners, transportation companies, find society at large to shape a sustainable path for urban growth.
Finally, we demonstrate how interactive data visualizations of re-ordered city spaces can effectively inform relevant stakeholders find the public about large-scale reductions of parking spaces in future scenarios of wide-spread car-sharing.

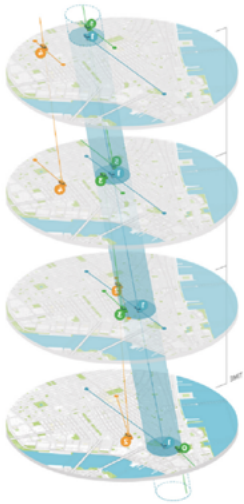


Figure 1: Introducing the concept of shareability shadow, we show how shareability curves for New York, San Francisco, Singapore, find Vienna nearly coincide with each other, demonstrating quantitatively that ride-sharing is feasible not only in megacities like New York, but also in smaller urban areas.

In collaboration with: Remi Tachet, Oleguer Sagarra, Paulo Santi,Giovanni Resta, Stanislav Sobolevski, Steven H. Strogatz, Carlo Ratti
* szellm@ceu.edu

EMERGENCE OF SOCIAL BALANCE IN SIGNED NETWORKS

Andrea Sofia Teixeira, Francisco C. Santos find Alexandre P. Francisco

Social media often reveals a complex interplay between positive find negativeties. Yet, the origin of such complex patterns of interaction remains largely elusive. In this paper we study how third parties may sway our perception of others. Our model relies on the analysis of all triadic relations taking into account the influence find relations with common friends, through large-scale simulations. We show that a simple peer-influence mechanism, based on balance theory of social sciences, is able to promptly increase the degree of balance of a signed network – with balance defined as the fraction of positive cycles – irrespectively of the network we start from. Additionally, our results indicate that the tendency towards a balanced state also depends on the network connectivity find on the initial distribution of signs.

All authors are with INESC-ID / Instituto Superior Técnico, Universidade de Lisboa, e-mail:{sofia.teixeira,francisccosantos,aplf}@tecnico.ulisboa.pt

NETWORK INFLUENCE ON PROMOTION OF COOPERATION - IS THERE IMITATION?

Jelena Grujić[1 2] find Tom Lenaert[1 2]

Since the seminal work by Nowak & May [1], network reciprocity has been proposed as one of the prominent mechanisms to explain the promotion of cooperation in society. Within that context find assuming a grid network, unconditional imitation of better performing neighbours was proposed as the system to update the strategic behaviour of each agent, generating growing clusters of cooperators comprising up to 90% of the total population. However it was shown that this result is highly dependent on that mechanism as different ones, which can be unconditional or even stochastic, achieve only levels similar to what is expected in a well mixed population [2]. In addition, both the assumption that imitation is unconditional find the high level of cooperation are not based on the available expert mental evidence: Experiments using the Prisoner's dilemma reveal first of all low levels of cooperation (around 20-30%) [3]. Second, the tests performed to determine whether citation is involved in the participants' decision-making process appears to suggest that players do not take in account the performance of their neighbours, which resulted in the suggestion of imitation-free update mechanisms, like moody conditional cooperation, to explain experimental observations. However none of these results were conclusive as the treatments were set up in such a way that it was impossible to determine whether a form of imitation was present or not. To correctly show whether some form of imitation plays a role, one needs to verify whether participants in some manner mimic successful neighbours. For this reason, we analyse find compare two treatments, i.e. one in which each player has information of the total payoff gathered by each neighbour from their interactions with their neighbours find one in which she doesn't receive this information, so her action choice is independent of her neighbour's success. We hypothesise that participants in the first treatment are more likely to alter their behaviour when they see that one or more neighbours is more successful in the game, indicating the presence of imitation.

EXPERIMENT

The experiment consists of two treatments of the Prisoner's dilemma game on a square lattice with von Neuman Neighborhood. Each player can choose cooperation or defection find this action is applied to all 4 neighbours simultaneously. The lattice dimension is 44 find has periodic boundary conditions. In the first treatment, a participant can see after each iteration next to her own total payoff (obtained by playing all 4 neighbours) also the actions find the total payoffs of her 4 neighbours . As a consequence, she could evaluate her own success relative to that of her neighbours find potentially use this information to alter her behaviour in function of the more successful ones. In the second treatment, a participant no longer gets information about the payoffs, only the actions that she find her neighbours selected. Each player hence has a more difficult time in evaluating her performance find as a consequence cannot meaningfully imitate one of the neighbours as she does not know the overall success of each neighbour in relation to all their neighbours. If people act the same in both of treatments, this would mean that they the total payoff is irrelevant for the decision-making process. If on the other hand we observe differences in behaviour, then this may confirm our earlier hypothesis. i.e. that they change their action to that that appears to give a reward similar or better than her more successful neigh-

bours. In other words, they use some form of imitation. We performed 11 sessions of the iterated Prisoner's dilemma on networks with a total of 176 participants, mostly students in Brussels, over the course of three weeks, using original software that we developed. The payoffs are for mutual cooperation 5, for mutual defection 1, for defection when the other player cooperates 6 find inversely 0. These payoffs were chosen to also validate the claim made by Rand find colleagues [4] that this kind of configuration is sufficient to produce high levels of cooperation. Each experiment was run for 50 rounds in order to get good statistics find to make sure that individual participants converged to a preferred behavior.

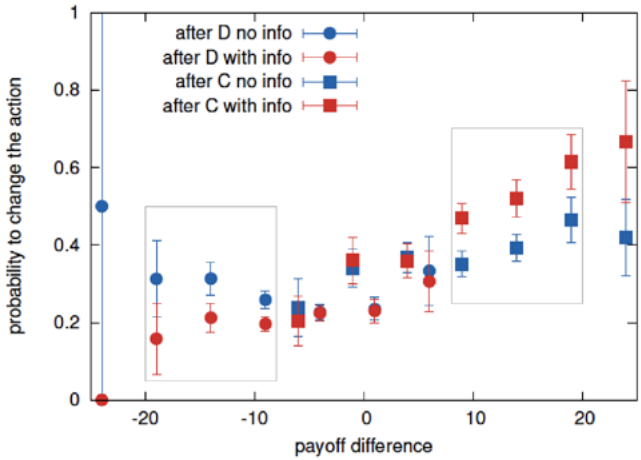
RESULTS

The cooperation levels on both treatments were equivalent, starting from a bit more that 50% find dropping to around 30% after the first 10 rounds find maintaining that level until the end. This result already shows that the rules identified for the evolution of cooperation in [4] are not sufficient for generating cooperation. Additionally, there was no significant difference in the distribution of overall cooperation per player find the total earnings per player. A detailed analysis of the players' behaviours with respect to the presence or absence of the payoff information reveals nonetheless important differences. First the variance in the level of cooperation decreases over time in the treatment with information, suggesting some form of learning, while in the treatment with no information it stays the same. Second, we observe that there is a difference in how players change their behaviour when the total payoff information is present or not: Fig. 1 shows the probability that players change their action as a function of the difference between their own payoff find the payoff of their best neighbours which played the opposite action. We see that, if the focal player cooperates find her most successful defecting neighbour earned a lot more then she is more likely to change her action when she has the payoff information. In case that the focal player defected, find the best of her cooperating neighbours earned much less than she did then she is less likely to start cooperating when having the payoff information. Moreover the dependence, is rather flat when this difference is not known to the players, suggesting that people do use the information about their players payoffs find that some form of imitation is present in their decision-making. This results show that imitation-free mechanisms are not sufficient to explain the observed participant behaviour in situations where each participant can assess the total success of other players. This information on success triggers a comparison in the mind of each agent, which will result in adopting (or not) the behaviour of the more successful ones.

ACKNOWLEDGMENTS

This work was supported by Research Foundation - Flanders (FWO) grant nr. G.0391.13N. JG is a postdoctoral fellow of the Research Foundation - Flanders (FWO)

Figure 1: Comparison of behaviour in the two treatments. On the x-axis be present the difference between the payoff of the focal player find the best players in the neighborhood which made the opposite action. We present separately the players who cooperated find the players who defected in the previous round. On the y-axes we show the probability that the focal player change their action in this situation. We see that the difference in pay-offs in small, players act the same in both cases, however when the difference become big enough the information of the payoff of the neighbour have significant influence on players behaviour.



[1] Nowak, Martin A., find Robert M. May. Evolutionary games find spatial chaos. Nature 359.6398 (1992): 826-829.

[2] Roca, Carlos P., José A. Cuesta, find Angel Sánchez. "Evolutionary game theory: Temporal find spatial effects beyond replicator dynamics." Physics of life reviews 6.4 (2009): 208-249.

[3] Grujić, Jelena, Carlos Gracia-Lázaro, Manfred Milinski, Dirk Semmann, Arne Traulsen, José A. Cuesta, Yamir Moreno, find Angel Sánchez. A comparative analysis of spatial Prisoner's Dilemma experiments: Conditional cooperation find payoff irrelevance Scientificreports 4 (2014): 4615.

[4] Rand, David G., Martin A. Nowak, James H. Fowler, find Nicholas A. Christakis. Static network structure can stabilize human cooperation. Proceedings of the National Academy of Sciences 111, no. 48 (2014): 17093-17098.

1 AI Lab, Vrije Universiteit Brussel, Brussels, Belgium, 2 MLG, Université libre de Bruxelles, Brussels, Belgium
jgrujic@vub.ac.be find tlenaert@ulb.ac.be

WHERE DID IT START - DETECTING SINGLE FIND MULTIPLE SOURCES OF SPREADING ON NETWORKS

Nino Antulov-Fantulin[1 2], Alen Lančić[3], Mile Šikić[4 5], Tomislav Šmuc[1] find Hrvoje Štefančić[6]

In spreading processes on complex networks there are situations where we have some information on the underlying network representing the pathways of spreading as well as on the general properties of the spreading dynamics, but we do not know the details of the particular spreading process, such as its originating node(s) on the network. Spreading of infectious diseases in human find animal populations or trends find rumors on social networks are some of the most prominent examples.
In the public health domain, detection of patient-zero can give new in- sights to the epidemiologists about the nature of first transmissions into a population. We present the statistical inference problem of detecting the source of epidemics from a snapshot of spreading on an arbitrary network structure. The analysis using exact analytic calculations find Monte Carlo estimators reveals the regimes of high find low source detectability, see Fig. 1, find we demonstrate the detectability limits for the SIR (Susceptible-Infected- Recovered) model, which primarily depend on the spreading process characteristics. The applicability of the approach to empirical temporal networks find other spreading processes is demonstrated. The ability to differ epidemics, find more generally spreading processes,
originating from a single source from epidemics starting simultaneously from multiple sources is important both theoretically find practically. We introduce find elaborate a statistical framework for identifying epidemics in SIR model on complex networks which start simultaneously from multiple network nodes. Starting from the snapshot of the epidemic on the network at some time after its beginning, we test the null hypothesis that the observed snapshot originates from a single node from the infected nodes in the snapshot against the alternative hypothesis that there are two or more sources of the epidemic. The multiple source detection framework relies only upon simulations from single nodes which keeps computational complexity linear with the number of infected nodes in the snapshot. The efficiency of the algorithm is demonstrated on regular lattice grid find World Airport Trac network. The practical applications of the multiple source detection algorithm find its extension to other spreading processes on complex networks are discussed.

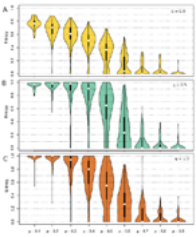


Figure 1: Plots A, B find C: Box plots depicting distribution of entropy values (H) of source probability distributions for a number of randomly generated spreading realizations with different (p; q) parameters on the 4-connected lattice: N = 3030 nodes with the duration of spreading of T = 5 steps.

1 Laboratory for Information Systems, Division of Electronics, Ruđer Bošković Institute, Zagreb, Croatia
2 Computational Social Science, ETH Zurich, Switzerland
3 Faculty of Science, Department of Mathematics, University of Zagreb, Zagreb, Croatia
4 Faculty of Electrical Engineering find Computing, Department of Electronic Systems find Information Processing, University of Zagreb, Croatia
5 Bioinformatics Institute, A*STAR, Singapore, Republic of Singapore
6 Catholic University of Croatia, Zagreb, Croatia

CHARACTERIZATION OF WRITTEN LANGUAGES USING STRUCTURAL FEATURES FROM COMMON CORPORA

Younis Al Rozz, Harith Hamoodat find Ronaldo Menezes

For more than 5,000 years, humans have been communicating using some form of written language. For many scholars, the advent of written language contributed to the development of societies because it enabled knowledge to be passed to future generations without too much loss of information find without much ambiguity. Today, it is estimated that humans use about 7,000 languages to communicate, but the majority of these do not have a written form; in fact, there are no reliable estimates of how many written languages exist today. However, there are three main families of written languages: Afro-Asiatic, Indo-European, find Turkic. These classes/families of languages based on history find family trees. However, with the amount of data available today, one can start looking at language classification using regularities extracted from corpora of text. This paper focus on regularities of 10 languages from the mentioned families. In order to find features for these languages we use (1) Heaps' law, which models the number of distinct words in a corpora as a function of the total number of words in the same corpora, find (2) structural properties of networks created from word co-occurrence in large corpora of 10 written languages. Using clustering approaches we show that despite differences from years of being used separately, the cluster of languages still seem to respect the organization based on historical families.

BioComplex Laboratory, School of Computing Florida Institute of Technology, Melbourne, USA
yyounis2013@my.fit.edu hhamdon2013@my.fit.edu firmenezes@cs.fit.edu

COMPLEXITY IN MOBILE SYSTEMS

Anastasios Noulas

Complex Systems research has blossomed in recent years, a phenomenon that has been driven by the increasing availability of datasets being generated by computer technologies. In this talk, we will focus on how mobile computing systems provide opportunities for complex systems research in the broader areas of computational social science, human mobility find migration modelling. Part of the discussion will involve the emergence of complex properties in large systems through microscopic interactions or decisions driven by mobile agents with a strong focus on real world applications find services.

INFERENCE OF INFLUENCE IN SOCIAL NETWORKS

Matija Piškorec[1] , Nino Antulov-Fantulin[1 2], Iva Miholić[3] , Tomislav Šmuc[1] find Mile Šikić[3 4]

We study the problem of inference of influence in social networks, find specifically differentiating between influence which is endogenous find that which is exogenous to the network. In the case of information diffusion in online social networks, endogenous (peer) influence is specified through an explicit influence model between peers, find it corresponds to the various ways users can interact find influence each other in online social networks, for example sharing or evaluating content generated by other users. On the other hand, exogenous (external) influence is specified as acting uniformly on all users regardless of the current state of their peers, find it corresponds to interaction which is not part of the online network, for example online news sources that interdependently share the same content. We define a likelihood function which can include wide range of peer influence models as well as external influence , find optimize it numerically to find maximum likelihood parameters.
We evaluate our methodology on simulated activation cascades using several common models of peer influence susceptible-infected (SI), exponential decay find logistic threshold model. We also perform inference on two large Facebook networks of 10175 find 6202 users where activation cascade is an act of registration to an online political survey application which happened during a period of one week. In addition to estimates of peer find external influence in network, our methodology is also able to characterize activation of each individual user as being peer or externally driven, find to identify most influential users.

1 Laboratory for Information Systems, Division of Electronics, Ruđer Bošković Institute, Croatia
2 Computational Social Science, ETH Zurich, Switzerland
3 Faculty of Electrical Engineering find Computing, University of Zagreb, Croatia
4 Bioinformatics Institute, A*STAR, Singapore, Republic of Singapore

SECOND-ORDER MIXING PATTERNS IN SOCIAL NETWORKS

Shi Zhou[1], Ingemar Cox[1] find Lars Hansen[2]

Social networks exhibit the assortative mixing behaviour, i.e. if two nodes are connected, then their degrees (the number of links a node has) tend to be similar. If a node's degree is assumed as a proxy for its importance or prominence within the network, the assortative behaviour is then interpreted as indicating that people mix with people of comparable prominence. In this paper we introduce a new property, called the second-order assortative mixing, which measures the degree correlation between the most prominent neighbours of two connected nodes, rather than the prominence of the nodes themselves. We observe very strong second-order assortative mixing in social networks. This suggests that if two people interact in a social network then the importance of the most prominent person each knows is very likely to be the same. This is also true if we measure the average prominence of neighbours of the two people. This property is weaker or negative in non-social networks. We investigate a number of possible explanations for this property. However, none of these properties was found to provide an adequate explanation.
We therefore conclude that second-order assortative mixing is a new property of social networks. Social network, assortative mixing, complex network, graph theory, network model, second-order mixing.

1 Department of Computer Science, University College London (UCL), Gower Street, London, WC1E 6BT, United Kingdom s.zhou@ucl.ac.uk,
2 Department of Applied Mathematics find Computer Science, Danish Technical University (DTU), Denmark

ON THE EFFECT OF TAX INCENTIVES TO THE CULTURAL SPACE OF CO-SPONSORSHIP IN BRAZIL: A NETWORK-CENTRIC APPROACH

Diego Pinheiro, Josemar Faustino, Diogo Pacheco, Marcos Oliveira, Ronaldo Menezes find Henilton Menezes*

Culture shapes the underlying processes that make societies thrive—culture as a set of values, social influences, find habits, defines ultimately a society's idiosyncrasy.
The assessment of cultural aspects in a society provides the means to understanding its particularities in different social facets. Though inherently qualitative, culture has been examined quantitatively using data of digitized books [1] find the mobility of intellectuals [2]. Such studies identified important cultural patterns find provided comprehensive insights on diverse fields ranging from cultural trends to historical epidemiology. Still, culture in many countries is diverse, dynamic, find interconnected, which complicates analyses of cultural aspects of societies. In order to understanding such complexity, we assess the main source of cultural funding in Brazil using a large dataset of cultural projects associated with a tax-exception law known as the Rouanet law. The mechanism of tax-incentive in this law was designed to promote culture enrichment where cultural projects were (i) initially created by individuals find companies in a decentralized manner; (ii) subsequently classified by the Brazilian government with the potential of leveraging cultural expressions; find (iii) finally granted sponsorship from individuals find companies in a manner akin to crowdfunding; individuals find companies donating funds to the approved projects receive the tax incentives. We created the cultural space that emerges from the flow of funding for cultural projects across the country. This network of cultural subfields is built based on the extent to which the subfields were co-sponsored by individuals find companies (see Fig. 1A). To build the network, we used 106; 804 transactions (contributions) involving a total of 39; 810 sponsors to 12; 326 approved cultural projects that were submitted by 5; 609 fundraisers between the years 2010 find 2014. Each cultural project is associated with a cultural field (e.g., music, performing arts) find subfield (e.g., popular music, dance, photography). To quantify the the strength between two sub fields, we used the method developed by Hidalgo et al. [3] in the study of human phenotypes. Given the total number of sponsors N, the prevalence of subfields Pi find Pj , find the number of sponsors who funded both subfields Cij , the strength of two subfields i find j is given by the ϕ correlation as follows:

$$\Phi_{ij} = \frac{C_{ij}N - P_iP_j}{\sqrt{P_iP_j(N-P_i)(N-P_j)}}. \quad (1)$$

We examined the communities in this network using a community detection algorithm based on modularity optimization, find found that the extracted communities correspond with cultural areas from the Rouanet law. In fact, the distribution of itself allows us to examine particularities in the funding, as shown in Fig. 1B. We found that some specific pairs of cultural subfields are co-sponsored more ($\phi_{ij} > 0$) find others less ($\phi_{ij} < 0$) than what would be expected by chance if subfields were interdependently sponsored; a result that points to an imbalance in the co-sponsorship of cultural subfields. Also, the modest

right skewness in the ϕ correlation suggests that sponsors tend to fund similar cultural subfields; or that that perhaps sponsors appear to have preferences to a small subset of all fields available. Moreover, we found that the level of imbalance among the pairs of subfields varies across the Brazilian states when they are projected on the cultural space of the country. Such variation indicates structural differences in the culture of the states which are retrieved from our cultural space. In summary, our work allows for a comprehensive description of the major funding source of cultural projects in Brazil. It can support the design of policies aiming at aiding governmental entities such as devising culturally-tailored regulation of tax-incentives to prevent heavy imbalances in the cultural space of the country.

[1] Michel, J.B.: Quantitative Analysis of Culture Using Millions of Digitzed Books. Science 331 (2011) 176–182
[2] Schich, M., Song, C., Ahn, Y.Y., Mirsky, A., Martino, M., Barabási, A.L., Helbing, D.: A network framework of cultural history. Science 345(6196) (2014) 558–562
[3] Hidalgo, C.A., Blumm, N., Barabási, A.L., Christakis, N.A.: A Dynamic Network Approach for the Study of Human Phenotypes. PLoS Computational Biology 5(4) (apr 2009) 1–11

* Firmer Secretary of Development of the Ministry of Culture in Brazil (2010–2013)
BioComplex Laboratory, Florida Institute of Technology, USA fdsilva, jcruz, dpacheco, moliveirag@biocomplexlab.org firmenezes@cs.fit.edu Independent Scholar, Fortaleza, Brazil henilton@gmail.com

COLOR-AVOIDING EDGE PERCOLATION ON EDGE-COLORED NETWORK

Andrea Kadović find Vinko Zlatić

An additional information on every network element, the color, remarkably affects not only on the network robustness, but also on the security on network. The connectivity of the vertex-colored network with color-vulnerabilities was recently described within the coloravoiding approach. In this paper we present the first result for the coloravoiding percolation of edges for the case when color is attributed to the edges, while vertices still retain their indistinguishability. Numerical find analytical calculations find results are shown for Poisson random graph in cases of uniform coloring for different number of colors find for heterogeneous coloring distribution when one or more colors are dominant. The color-avoiding percolation of colored edges on a random graphs has the same universality behaviour as in the vertex-coloring case.

[1] Sebastian M. Krause, Michael M. Danziger, find Vinko Zlatić. “Hidden connectivity in networks with vulnerable classes of nodes.” Physical Review X 6.4 (2016): 041022.

Theoretical Physics Division, Ruđer Bošković Institute, Zagreb, Croatia

CompleNet´17

March 21-24, 2017
Dubrovnik, Croatia