Physics Letters A 380 (2016) 359–364

Contents lists available at ScienceDirect

Physics Letters A

www.elsevier.com/locate/pla

Information theory perspective on network robustness

Tiago A. Schieber [a,b], Laura Carpi [c], Alejandro C. Frery [d], Osvaldo A. Rosso [e,f], Panos M. Pardalos [b], Martín G. Ravetti [a,g,*]

[a] Departmento de Engenharia de Produção, Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
[b] Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA
[c] Departament de Física i Enginyeria Nuclear, Universitat Politècnica de Catalunya, Colom 11, Terrassa, 08222, Barcelona, Spain
[d] Laboratório de Computação Científica e Análise Numérica (LaCCAN), Universidade Federal de Alagoas, Maceió, Alagoas, Brazil
[e] Instituto de Física, Universidade Federal de Alagoas, Maceió, Alagoas, Brazil
[f] Instituto Tecnológico de Buenos Aires (ITBA), Ciudad Autónoma de Buenos Aires, Argentina
[g] Departament de Física Fonamental, Universitat de Barcelona, Barcelona, Spain

ARTICLE INFO

Article history:
Received 26 June 2015
Received in revised form 20 October 2015
Accepted 22 October 2015
Available online 3 November 2015
Communicated by C.R. Doering

Keywords:
Network robustness
Complex networks
Information theory

ABSTRACT

A crucial challenge in network theory is the study of the robustness of a network when facing a sequence of failures. In this work, we propose a dynamical definition of network robustness based on Information Theory, that considers measurements of the structural changes caused by failures of the network's components. Failures are defined here as a temporal process defined in a sequence. Robustness is then evaluated by measuring dissimilarities between topologies after each time step of the sequence, providing a dynamical information about the topological damage. We thoroughly analyze the efficiency of the method in capturing small perturbations by considering different probability distributions on networks. In particular, we find that distributions based on distances are more consistent in capturing network structural deviations, as better reflect the consequences of the failures. Theoretical examples and real networks are used to study the performance of this methodology.

1. Introduction

There are several works dealing with the concept of robustness, however, there is still no consensus on a definitive definition. Robustness is usually described as the ability of the network to continue performing [1], or, as the capacity in maintaining its functionality after failures or attacks [13,14,16]. These general definitions are perhaps the most used in the literature, however, they cannot exactly grasp the complexity of the concept that network's robustness could have. Some other works describe robustness as the capacity of the network in maintaining its efficiency in the presence of failures [11,12]. In some sense, this definition provides more information about the network's topological structure, as its efficiency depends on the network's shortest path lengths [17].

The study of how robust a network is when facing random failures or targeted attacks is a major challenge in network theory. Several methodologies have been proposed to measure network robustness. Approaches based on information routing [4,24,25], structural controllability [19,23] or in the proposal of a more destructive attack strategy in networks [3,22] can be found in the literature; being the most popular those based on percolation theory [1,7,9], and on the size of the biggest connected component (BC) [2,15,16,26]. Although these measures showed to be useful in many cases, they are not as sensitive as they should, to the detection of failures that do not disconnect the network or that do not modify its diameter. Depending on the network structure, it is possible to attack great part of it, keeping these measures blind to the changes.

In this work, we propose a measure for network robustness based on the Jensen–Shannon divergence, an Information Theory quantifier that already showed to be very effective in measuring small topological changes in a network [8,27,28]. This method considers failures occurring in a temporal sequence capturing, in some sense, the dynamics of the role of the remaining links after each single failure. The Jensen–Shannon divergence quantifies the topological damage of each time step due to failures, and the robustness measure provides the cumulative information of these sequential topological damages. It is worth noticing that this approach does not consider the consequences of the dynamical process operating through the network.

[*] Corresponding author.
E-mail address: martin.ravetti@dep.ufmg.br (M.G. Ravetti).

http://dx.doi.org/10.1016/j.physleta.2015.10.055
0375-9601/© 2015 Elsevier B.V. All rights reserved.

2. Methodology

Quantification of network robustness could be thought as the distance that a given topology is apart from itself after a failure. We assume that the robustness value ranges from 0, the greatest variation, to 1, unchanged characteristics. In other words, a higher robustness value implies in smaller structural changes. In this work we consider a link failure, its removal, and a node failure in the removal of all it incident links.

Let G be a network defined by a set $V(G)$ of N nodes, a set $\mathcal{E}(G)$ of M links and a set $W(E(G))$ containing the edges strengths. A network failure event f is defined as the removal of a subset of edges $f \subset \mathcal{E}(G)$. A time-ordered sequence of failures $\mathcal{F} = \{f_{t_1}, f_{t_2}, \ldots, f_{t_n}\}$ in G can be interpreted as a sequence of the resulting networks after each event $(G_{t_i})_{i\in\{0,\,1,\ldots,n\}}$ such that $G_{t_0} = G$ and G_{t_i} is the network obtained after the failure f_{t_i} in $G_{t_{i-1}}$. For simplicity, here we consider only discrete time intervals given by $t_i = i$.

Considering the set $\mathcal{N}_{\mathcal{F}}$ of all possible sequences of failures in a network G, a *robustness function with respect to G is a function* $R\colon \mathcal{N}_{\mathcal{F}} \to [0, 1]$. The distance between two networks is computed as the distance between probability distributions used to characterize them. Without loss of generality, discrete distributions will be considered henceforth.

The *Jensen–Shannon divergence* between two probability distributions P and Q is defined as the Shannon entropy of the average minus the average of the entropies. This measure was proven to be the *square of a metric* between probability distributions [18], bounded by 1, and defined as:

$$\mathcal{J}^H(P, Q) = H\left(\frac{P+Q}{2}\right) - \frac{H(P)+H(Q)}{2},$$

being $H(P) = -\sum_i p_i \log_2 p_i$, the entropy that measures the *amount of uncertainty* in a probability distribution. Readers are referred to the Supplementary Information material (SI) for a discussion on the continuous case.

It is possible then, to define the *robustness* of G, for any given sequence of n failures $(G_t)_{t\in\{1,\,2,\ldots,n\}}$ and probability distribution P as:

$$R_P(G|(G_t)_{t\in\{1,\,2,\ldots,n\}}) = \prod_{t=1}^{n}\left[1 - \mathcal{J}^H(P(G_t), P(G_{t-1}))\right], \qquad (1)$$

being $G_0 = G$.

A more suitable form of equation (1) can be obtained via recurrence relation:

$$R_P(G|(G_t)_{t\in\{1,\,2,\ldots,n\}}) = \prod_{t=1}^{n} R_P(G_{t-1}|G_t), \qquad (2)$$

in which, for each time step, $R_P(G_{t-1}|G_t)$ indicates how affected the topology of the network G_{t-1} is after a single failure resulting in G_t. The robustness function depends on the network's topology, and also on the sequence of failures. The same link possesses different importance (effect) in the topology, depending on its position in the failure sequence. The use of the product of the temporal fluctuations ($R_P(G_{t-1}|G_t)$) allows us to have a perception of the temporal damage due the sequence of failures.

It is important to notice that the computation of the network robustness, when defined as in equation (1) could consider any probability distribution able to represent features of the network. This work specifically considers the degree distribution, commonly used to characterize network's structures, and the distance distribution that contains rich information about the graph structure. The degree and distance distributions are here defined for unweighted and undirected networks. In Section I of the SI readers can find a discussion other network robustness measurements and in section VI the analysis of a directed and weighted network is performed.

Given a node i, its *degree*, represented by k_i, is the number of edges incident to it. Then, the *degree distribution* $P_{deg}(k)$ is the fraction of nodes with degree k. The *distance* from the node i to node j, $d_{i,j}$, is the length of the shortest path from i to j. If there is no such path from i to j, $d_{i,j}$ equals ∞. Then, the *distance distribution* $P_\delta(d)$ is the fraction of pairs of nodes at distance d. Both degree and distance distributions are discrete and defined on the sets $\{0,\ 1, \ldots,\ N-1\}$ and $\{1,\ 2, \ldots,\ N-1,\ \infty\}$, respectively.

3. Discussions and computational experiments

With the aim of comparing the performance of the methodology based on $R_{P_{deg}}$ and R_{P_δ} with commonly used methods based on the biggest connected component (R_{bc}) and percolation (R_{π_d}), we consider the deletion of a single link on a complete graph with N nodes, the most robust unweighted and undirected graph.

R_{bc} is obtained by computing the fraction of nodes belonging to the biggest connected component. In this case, R_{bc} does not notice the removal of any link, as no disconnection is achieved. It is possible to strategically remove $N^2/2 - 2N + 1$ links, leaving just the minimum spanning tree, where robustness measures based on the biggest connected component remain blind to these attacks.

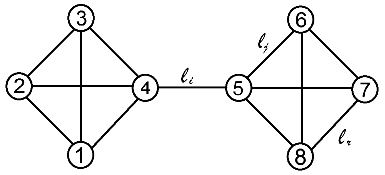
Percolation based measures correlate the robustness value of the network with the critical percolation threshold and can be computed in several ways. One of the most common methods depends on the number of links removed until increasing the diameter of the network. R_{π_d} indicates the variation of the original diameter d_0 with respect to diameter d after a sequence of failures, computed by $R_{\pi_d} = d_0/d$.

In the case of the complete graph, the deletion of a single link increases the network diameter in one unit, however, after the first attack R_{π_d} may become unable to detect subsequent events. In order to increase the network's diameter in one more unit, the removal of $N-2$ specific links are needed.

The proposed robustness measure is able to detect the removal of any single link of the network, independently of which probability distribution (P_{deg}, P_δ) is evaluated. Values for $R_{P_{deg}}$ and R_{P_δ} can be then, easily computed as functions of N. The removal of a single link in a complete graph with 10^6 nodes, implies in changes of the order of 10^{-6} for $R_{P_{deg}}$ and 10^{-13} for R_{P_δ}. For a complete graph of $N = 10^7$, changes are of the order of 10^{-7} for $R_{P_{deg}}$ and 10^{-15} for R_{P_δ}.

Among the measures here considered, only $R_{P_{deg}}$ and R_{P_δ} showed to be capable of capturing the removal of any single link, showing a gradual decrease in the robustness values, as more links are removed from the network. This could be of relevance in situations in which it is necessary to plan the inclusion of new links to improve the robustness of the network. Methodologies based on the size of the biggest connected component, or on the percolation threshold, are not able to properly guide in this purpose. It is important to point out, that the robustness measure here proposed, depends not only on the network topology but on the sequence of failures over time, aiming to quantify the vulnerability of a given structure under a series of deterministic or stochastic failures. The process of fixing failures cannot be measured in the same way, but the degree and distance probability distributions seem to be adequate to this purpose.

There are interesting differences between P_{deg} and P_δ. The computational complexity to obtain the degree distribution is linear, plus a constant cost to update it, after any link removal. The best known algorithm for obtaining P_δ requires $\mathcal{O}(N^{2.376})$ in time complexity [29], and the computational cost of the PDF update depends on the link removed. However new algorithms as the



Edge removed	R_{P_δ}	$R_{P_{deg}}$	R_{bc}	R_{π_d}
ℓ_i	0.447	0.862	0.500	0.000
ℓ_j	0.943	0.922	1.000	0.750
ℓ_r	0.998	0.857	1.000	1.000

Fig. 1. Computation of the structural robustness for three different single edge removal: ℓ_i, ℓ_j and ℓ_r, respectively.

ANF or HyperANF (algorithms based on HyperLogLog counters) offer an extremely fast and precise approach [5,6,10,21], obtaining very good approximations of the distance probability distribution for graphs with millions of nodes in a few seconds. In the SI readers can find a table with the computational complexity of the most common methods.

Another important comparison is the information that can be assessed from both distributions, and their correlation with topological structures. The network's average degree, mean degree and the minimum and maximum degree are immediately obtained from the degree distribution. The network's efficiency, diameter, average path length, fraction of disconnected pairs of nodes and other distance features are easily obtained from the distance distribution.

Fig. 1 shows a simple network structure to analyze the correlation of the robustness values with different topological changes. Individual removal of links ℓ_i, ℓ_j and ℓ_r are performed. R_{bc} only detects the disconnection of the biggest connected component, being not sensitive to the removal of ℓ_j and ℓ_r. R_{π_d} detects the removal of ℓ_j and ℓ_i, but fails in capturing the removal of ℓ_r, as there is no modification in the diameter. The $R_{P_{deg}}$ detects every single failure, however, its value does not properly reflect the network's disconnection (ℓ_i). The measure based on the distance distribution (R_{P_δ}) captures, in a more appropriate way, each of the above-mentioned network failures, especially those aspects concerning disconnections on a connected network. This example captures important advantages and disadvantages of each robustness measure. In the SI, other quantifiers to measure robustness are also evaluated. However, the use of the distance distribution shows to be the most adequate for this analysis.

Let us now analyze two sequences of failures considering ℓ_i and ℓ_j. If link ℓ_i fails at instant $t = 1$ and link ℓ_j fails at instant $t = 2$, $R_{P_\delta} = 0.4377$. Now, if the sequence is inverted considering link ℓ_j failing at instant $t = 1$ and link ℓ_i failing at instant $t = 2$, $R_{P_\delta} = 0.4564$. This example depicts how the roles and topological importance of the remaining links after a failure are reflected by R_{P_δ}.

We test the proposed methodology on several real networks, nevertheless, only the results for two of them are depicted in the main text, the Dolphin Social Network [20] and the Western States Power Grid of the United States network [30]. Readers are referred to the SI section V for applications on other networks.

The Dolphin network is an undirected social network of bottlenose dolphins (*genus Tursiops*). The nodes are the bottlenose dolphins of a community from New Zealand, where an edge indicates a frequent association between dolphin pairs occurring more often than expected by chance [20]. The dolphins were observed between 1994 and 2001. It presents $N = 62$, $M = 159$, an average degree of 5.13, an average path length of 3.357, and a clustering coefficient of 0.258.

The Power Grid Network is the undirected and unweighted representation of the topology of the Western States Power Grid of the United States, compiled by Duncan Watts and Steven Strogatz [30]. It presents, $N = 4941$, $M = 6594$, an average degree of 2.67, an average path length of 18.99, and a clustering coefficient of 0.103.

In both cases, at each time step a single link is randomly removed until the global disconnection of approximately 10% of their links. Thirty independent experiments were performed and, at each time step, the robustness measure for each experiment is computed. Fig. 2 depicts the composition of violin plots of the robustness value, where R_P^m indicates the minimum robustness value found at each time step.

It is possible to see from Figs. 2(a) and 2(c) that the robustness measure computed from the degree distribution shows a smoother behavior, as it is unable to detect cluster disconnections. This is not the case for the distance distribution, in which the fraction of disconnected pairs of nodes is detected (see Figs. 2(b) and 2(d)). The large decrease in the R_{P_δ} values usually represents cluster disconnections from the network. As we are analyzing average values, the disconnection may occur in a fraction of the thirty independent experiments only.

The large variability of single robustness values for the Dolphin network reflects the extent of the damage that certain failures can cause, showing the Dolphin network more susceptible to random failures than the US Power Grid. The robustness measures, in particular those computed through the distance distributions, Figs. 2(b) and 2(d) also show big leaps when the link removal is around 6% and 9% for Dolphin and 3% and 6% for the Power Grid, indicating network's disconnections.

Fig. 3 compares the R_P^m values with two sequences of failures for each experiment; the sequences presenting the lowest robustness value at the end of the attack ($R_{P_\delta}(16)$ for Dolphin and $R_{P_\delta}(660)$ for Power Grid) and the sequences with the lowest robustness value at the first time step ($R_{P_\delta}(1)$). Note that the sequences of failures resulting in lower robustness values are not the most efficient in destroying the network at the beginning of the process. This behavior occurs because the robustness measure provides cumulative information about the evolution of the state of the network (see Equation (2)). A small $R_P(G_t|G_{t-1})$ value indicates that, at time t, the failure of certain links is critical, generating bigger changes in the topology.

This methodology could also be applied to detect critical elements, such as the nodes and links in the US power grid network that, when individually removed, cause a major disturbance in the network's structure. Fig. 4 shows the 10 most critical links and nodes that produce the largest robustness values variation (see table in Fig. 4). Critical elements for US power grid network identified by $R_{P_{deg}}$, as well as results for Dolphin network, can be found in SI; *cf.* Figs. S3–S5.

The knowledge of critical elements is of great importance to plan strategies either to protect or to efficiently attack networks. In both scenarios, the knowledge about how the network continues to perform after failures is of paramount importance. It is interesting noticing that the problem of finding the best sequence of links to destroy the network can be solved through combinatorial optimization approaches. Readers are referred to SI, section VI for a computational experiment considering targeted attacks in two directed and weighted real networks.

4. Final remarks

We propose a novel methodology to measure the robustness of a network to component failures or targeted attacks. This mathematical formulation is based on the consideration that the network robustness is a measure related to the distance that a given topology is apart from itself after a sequence of failures, rather that a

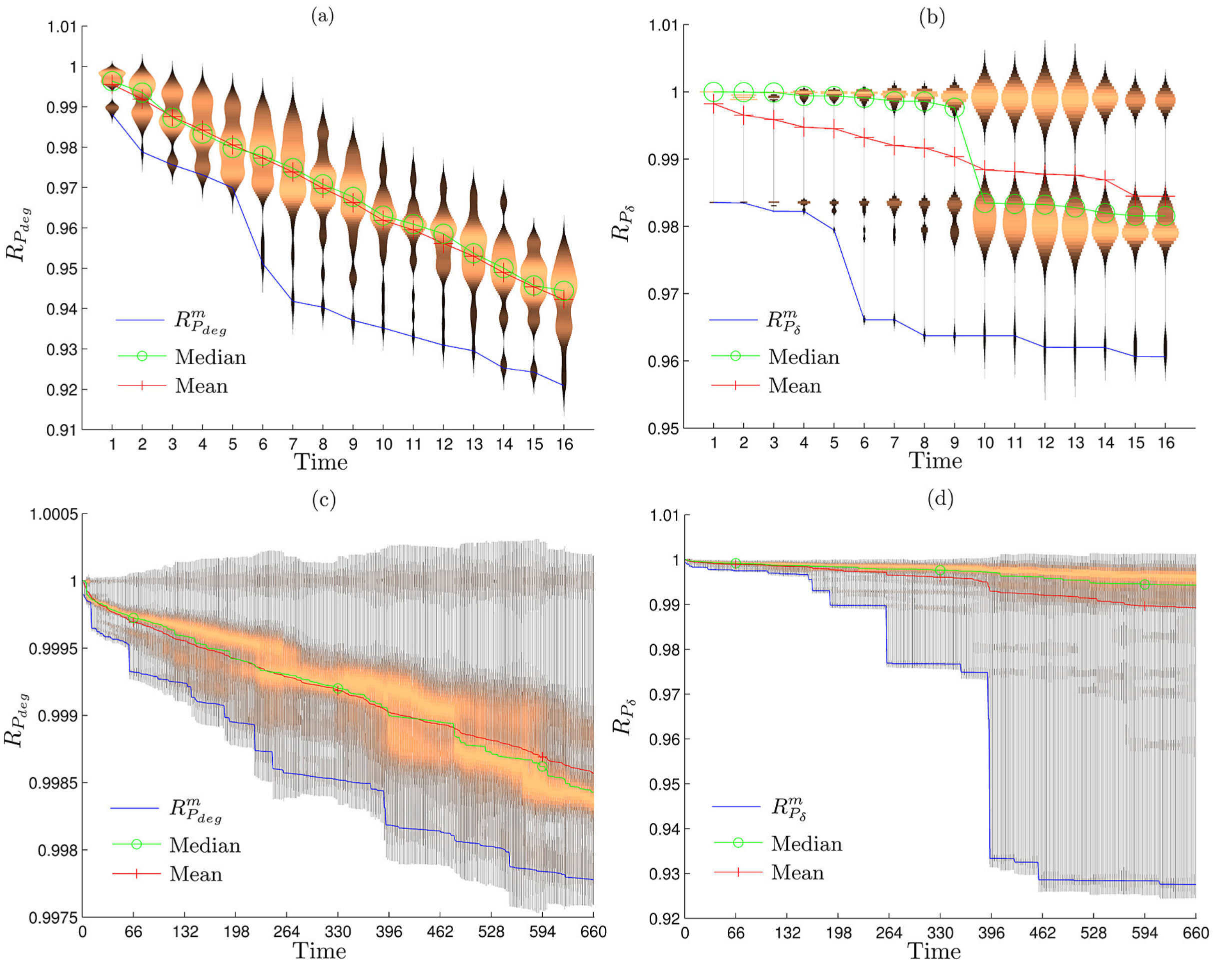


Fig. 2. Robustness measures under random failures for Dolphin and Power Grid networks. At each time step, a random edge is disconnected from the network and $R_{P_{deg}}$, R_{P_δ} functions are computed. The experiment is independently executed 30 times. Results for the Dolphin networks are depicted in (a) R_{Pdeg} and (b) R_{P_δ}. Results for the Power Grid network are presented in (c) $R_{P_{deg}}$ and (d) R_{P_δ}. In all cases, at each time step, the minimum robustness values are also indicated by R^m_P.

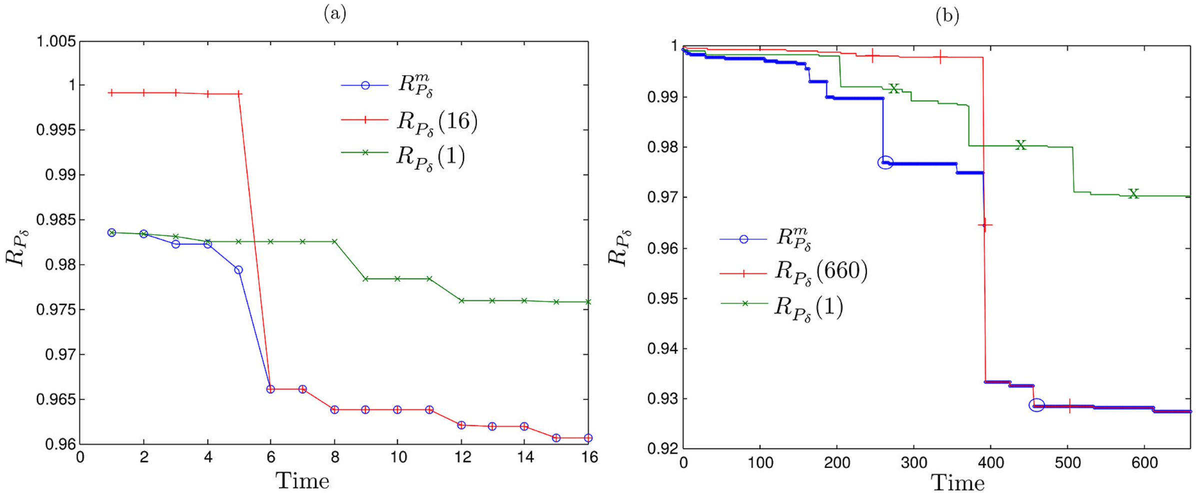
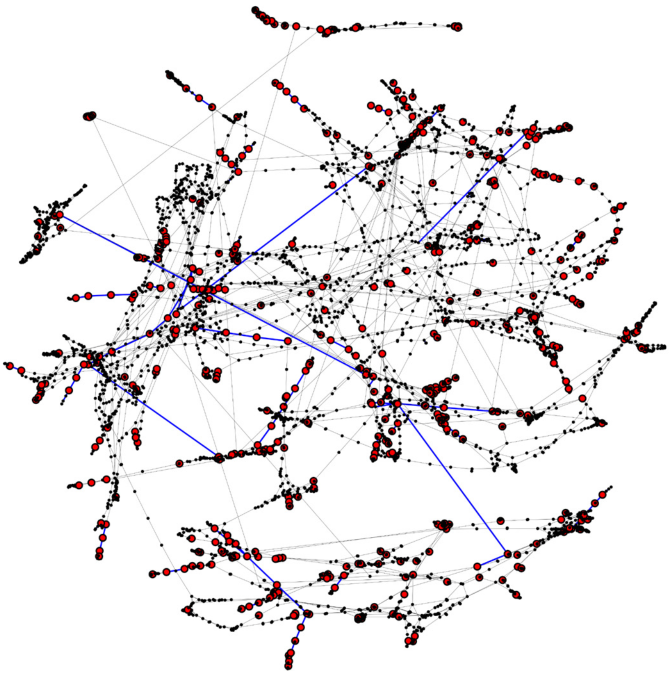


Fig. 3. Evolution of the R_{P_δ} of two different sequences of failures: the sequence that ends with the lowest robustness value, $R_{P_\delta}(16)$ and $R_{P_\delta}(660)$, and the sequence in which the first removal is most effective, $R_{P_\delta}(1)$. In both cases, at each time step, the minimum robustness values are also indicated by $R^m_{P_\delta}$. (a) Dolphin network and (b) US Power Grid network.



Edge ID	R_δ
4220 − 2544	0.9695
4220 − 4165	0.9960
3046 − 2523	0.9962
3046 − 3045	0.9964
3048 − 2523	0.9964
3048 − 3047	0.9968
363 − 270	0.9974
3074 − 3047	0.9974
347 − 270	0.9978
347 − 342	0.9980
Node ID	R_δ
4220	0.9675
2544	0.9677
727	0.9786
693	0.9887
2529	0.9914
2523	0.9919
2605	0.9941
4165	0.9946
363	0.9949
427	0.9955

Fig. 4. Detection of the ten percent of the most critical links and nodes, considering R_{P_δ} over the Western States Power Grid of the United States network. Wider nodes (red in the web version) represent the fraction of 10 of network's vertices such that its single disconnection causes a big reduction on the R_{P_δ} value. Wider links (blue in the web version) represent the fraction of 10 of network's edges such that its single disconnection causes a big reduction on the R_{P_δ} value. The table shows the robustness values of the top 10 critical network elements.

single characteristic of the topology. This sequence is defined as a time dependent process in which, a subset of links is disconnected at each time step. The method provides a dynamic robustness profile that shows the response of the network's topology to each event, quantifying the vulnerability of these intermediate topologies.

Although the methodology is comprehensive enough to be used with different probability distributions, the use of distances shows to be more consistent in capturing network structural deviations, in the sense that their values are correlated with the consequences of the failures in the network topology. Different from the methods found in the literature, the method can efficiently work with disconnections, as the distance PDF is able to acknowledge the fraction of disconnected pairs of nodes. Furthermore, it is able to detect all changes, including those perceived by R_{bc} and R_{π_d}, resulting in a more general approach.

Acknowledgements

Research partially supported by FAPEMIG, CNPq and RSF grant 14-41-00039. OAR acknowledges support from CONICET.

Appendix A. Supplementary material

Supplementary material related to this article can be found online at http://dx.doi.org/10.1016/j.physleta.2015.10.055.

References

[1] R. Albert, H. Jeong, A.-L. Barabási, Error and attack tolerance of complex networks, Nature 406 (6794) (Jul. 2000) 378–382.

[2] S. Allesina, M. Pascual, Googling food webs: can an eigenvector measure species' importance for coextinctions?, PLoS Comput. Biol. 5 (9) (Sep. 2009) e1000494.

[3] A. Arulselvan, C.W. Commander, L. Elefteriadou, P.M. Pardalos, Detecting critical nodes in sparse graphs, Comput. Oper. Res. 36 (7) (Jul. 2009) 2193–2200.

[4] V. Boginski, C.W. Commander, T. Turko, Polynomial-time identification of robust network flows under uncertain arc failures, Optim. Lett. 3 (3) (2009) 461–473.

[5] P. Boldi, M. Rosa, S. Vigna, Robustness of social networks: comparative results based on distance distributions, in: Proceedings of the Third International Conference on Social Informatics, SocInfo'11, Springer-Verlag, Berlin, Heidelberg, 2011, pp. 8–21.

[6] P. Boldi, S. Vigna, In-core computation of geometric centralities with hyperball: a hundred billion nodes and beyond, CoRR, arXiv:1308.2144, 2013.

[7] D.S. Callaway, M.E.J. Newman, S.H. Strogatz, D.J. Watts, Network robustness and fragility: percolation on random graphs, Phys. Rev. Lett. 85 (Dec. 2000) 5468–5471.

[8] L.C. Carpi, O.A. Rosso, P.M. Saco, M. Ravetti, Analyzing complex networks evolution through information theory quantifiers, Phys. Lett. A 375 (4) (Jan. 2011) 801–804.

[9] R. Cohen, K. Erez, D. ben Avraham, S. Havlin, Resilience of the internet to random breakdowns, Phys. Rev. Lett. 85 (Nov. 2000) 4626–4628.

[10] P. Crescenzi, R. Grossi, L. Lanzi, A. Marino, A comparison of three algorithms for approximating the distance distribution in real-world graphs, in: A. Marchetti-Spaccamela, M. Segal (Eds.), Lecture Notes in Computer Science, vol. 6595, Springer, Berlin, Heidelberg, 2011, pp. 92–103.

[11] P. Crucitti, V. Latora, M. Marchiori, A. Rapisarda, Efficiency of scale-free networks: error and attack tolerance, Physica A: Stat. Mech. Appl. 320 (Mar. 2003) 622–642.

[12] P. Crucitti, V. Latora, M. Marchiori, A. Rapisarda, Error and attack tolerance of complex networks, Physica A: Stat. Mech. Appl. 340 (1–3) (2004) 388–394. News and expectations in thermostatistics.

[13] A.H. Dekker, B.D. Colbert, Network robustness and graph topology, in: Proceedings of the 27th Australasian Conference on Computer Science, ACSC '04, vol. 26, Australian Computer Society, Inc., Darlinghurst, Australia, 2004, pp. 359–368, URL http://dl.acm.org/citation.cfm?id=979922.979965.

[14] M. Fiedler, Algebraic connectivity of graphs, Czechoslov. Math. J. 23 (98) (1973) 298–305.

[15] P. Holme, B.J. Kim, C.N. Yoon, S.K. Han, Attack vulnerability of complex networks, Phys. Rev. E 65 (May 2002) 056109.

[16] S. Iyer, T. Killingback, B. Sundaram, Z. Wang, Attack robustness and centrality of complex networks, PLoS ONE 8 (4) (Apr. 2013) e59613.

[17] V. Latora, M. Marchiori, A measure of centrality based on network efficiency, New J. Phys. 9 (188) (2007).

[18] J. Lin, Divergence measures based on the Shannon entropy, IEEE Trans. Inf. Theory 37 (1) (Jan. 1991) 145–151.

[19] Y.-Y. Liu, J.-J. Slotine, A.-L. Barabasi, Controllability of complex networks, Nature 473 (7346) (May 2011) 167–173, URL http://dx.doi.org/10.1038/nature10011.

[20] D. Lusseau, K. Schneider, O. Boisseau, P. Haase, E. Slooten, S. Dawson, The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations, Behav. Ecol. Sociobiol. 54 (4) (2003) 396–405.

[21] C.R. Palmer, P.B. Gibbons, C. Faloutsos, Anf: a fast and scalable tool for data mining in massive graphs, in: Proceedings of the Eighth ACM SIGKDD Inter-

national Conference on Knowledge Discovery and Data Mining, KDD '02, ACM, New York, NY, USA, 2002, pp. 81–90.

[22] C.-L. Pu, S. Lia, A. Michaelsonb, J. Yanga, Iterative path attacks on networks, Phys. Lett. A 379 (August 2015) 1633–1638.

[23] C.-L. Pu, W.-J. Pei, A. Michaelson, Robustness analysis of network controllability, Physica A 391 (2012) 4420–4425.

[24] C.-L. Pu, Zhou Si-Yuan, K. Wang, Y.-F. Zhang, W.-J. Pei, Efficient and robust routing on scale-free networks, Physica A 391 (2012) 866–871.

[25] C.-L. Pu, J. Yang, W.-J. Pei, Y.-T. Tao, S.-H. Lan, Robustness analysis of static routing on networks, Physica A: Stat. Mech. Appl. 392 (15) (August 2013) 3293–3300.

[26] M. Salathé, M. Kazandjieva, J.W. Lee, P. Levis, M.W. Feldman, J.H. Jones, A high-resolution human contact network for infectious disease transmission, Proc. Natl. Acad. Sci. 107 (51) (12 2010) 22020–22025.

[27] T. Schieber, L. Carpi, M. Ravetti, Evaluation of the copycat model for predicting complex network growth, in: C. Vogiatzis, J.L. Walteros, P.M. Pardalos (Eds.), Dynamics of Information Systems, in: Springer Proceedings in Mathematics Statistics, vol. 105, Springer International Publishing, 2014, pp. 91–108.

[28] T.A. Schieber, M.G. Ravetti, Simulating the dynamics of scale-free networks via optimization, PLoS ONE 8 (12) (Dec. 2013) e80783.

[29] J. Tang, T. Wang, J. Wang, D. Wei, Efficient social network approximate analysis on blogosphere based on network structure characteristics, in: Proceedings of the 3rd Workshop on Social Network Mining and Analysis, SNA-KDD '09, ACM, New York, NY, USA, 2009, pp. 7:1–7:8.

[30] D.J. Watts, S.H. Strogatz, Collective dynamics of 'small-world' networks, Nature 393 (6684) (Jun. 1998) 440–442.

Supplementary Information: Information Theory Perspective on Network's Robustness

Tiago A. Schieber,[1, 2] Laura Carpi,[3] Alejandro C. Frery,[4] Osvaldo A Rosso,[5, 6] Panos M. Pardalos,[2] and Martín G. Ravetti[7, 1, *]

[1]*Departmento de Engenharia de Produção,*
Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
[2]*Industrial and Systems Engineering, University of Florida, Gainesville, FL, USA*
[3] *Departament de Física i Enginyeria Nuclear,*
Universitat Politècnica de Catalunya. Colom 11, Terrassa 08222, Barcelona, Spain
[4] *Laboratório de Computação Científica e Análise Numérica (LaCCAN),*
Universidade Federal de Alagoas, Maceió, Alagoas, Brazil
[5] *Instituto de Física, Universidade Federal de Alagoas, Maceió, Alagoas, Brazil*
[6]*Instituto Tecnológico de Buenos Aires (ITBA), Ciudad Autónoma de Buenos Aires, Argentina*
[7]*Departament de Física Fonamental, Universitat de Barcelona, Barcelona, Spain*

(Dated: October 20, 2015)

A crucial challenge in network theory is the study of the robustness of a network when facing a sequence of failures. In this work, we propose a dynamical definition of network robustness based on Information Theory, that considers measurements of the structural changes caused by failures of the network's components. Failures are defined here, as a temporal process defined in a sequence. Robustness is then evaluated by measuring dissimilarities between topologies after each time step of the sequence, providing a dynamical information about the topological damage. We thoroughly analyze the efficiency of the method in capturing small perturbations by considering different probability distributions on networks. In particular, we find that distributions based on distances are more consistent in capturing network structural deviations, as better reflects the consequences of the failures. Theoretical examples and real networks are used to study the performance of this methodology.

PACS numbers: 89.75.-k,89.75.Fb, 89.70.Cf

I. MEASUREMENTS ON COMPLEX NETWORKS

A. Classical complex networks measurements

There are several measures to characterize networks [1, 4] and, many of them, are important to deal with the network robustness problem [3]. Here, we give a brief definition of measures of connectivity, distances, betweenness, clustering and spectral analysis.

Given a network $G = (V(G), E(G))$, being $V(G)$ and $E(G)$ sets of vertices and edges, respectively, we denote by $BC(G)$ the size of the biggest connected component of G. For any two vertices $i,\ j \in V(G)$, the distance $d(i, j)$ is the length of the shortest path between i and j and, if there is no path between i and j, $d(i, j) = \infty$. The network diameter, d, is the maximum over the set $\{d(i, j) \mid i, j \in V(G)\}$ and the average path length (apl):

$$apl = \frac{2}{|V(G)|(|V(G)| - 1)} \sum_{x \in D^{fin}} x.$$

being $D^{\text{fin}} = \{d(i, j) \mid i, j \in V(G) \text{ and } d(i, j) < \infty\}$.

The closeness centrality measure of a vertex $i \in V(G)$ is the sum of the inverse of all pairs of distances from i:

$$Cl(i) = \sum_{j,\, j \neq i} \frac{1}{d(i, j)}$$

[*] Electronic address: martin.ravetti@dep.ufmg.br

The efficiency of a network ($\mathcal{E}$) is the average of the inverse of all pairs of distances:

$$\mathcal{E} = \frac{\sum_i Cl(i)}{|V(G)|(|V(G)| - 1)}.$$

The clustering coefficient (C) characterizes the presence of triangles in the network and is given by the fraction between the number of triangles and the number of connected triples present in the network. Thus, a complete graph possess $C = 1$ and, a tree graph, $C = 0$. The vertex clustering coefficient, $C(i)$, is given by:

$$C(i) = \frac{3\#_{Delta}(i)}{\#_e(i)},$$

where, $\#_{Delta}(i)$ is the number of triangles involving vertex i and $\#_3(i)$ is the number of connected triples having i as a central vertex.

The betweenness centrality quantifies the importance of an element (node or edge) in terms of interactions via the shortest paths among them. For a vertex $v \in V(G)$ or an edge $e \in E(G)$ the betweenness centrality is defined, respectively, by:

$$B_v = \sum_{i \neq j \in V(G)} \frac{n(i,j,x)}{2n(i,j)} \qquad B_e = \sum_{i \neq j \in E(G)} \frac{n(i,j,e)}{2n(i,j)},$$

where, $n(i,j)$ is the number of shortest paths connecting i and j and $n(i,j,x)$ is the number of shortest paths connecting i and j passing through x. The same holds for e. Here, we denote by $B_v^{\max} = \max\{B_v \ | v \in V(G)\}$, $B_e^{\max} = \max\{B_e \ | e \in E(G)\}$ and B_v^{av} and B_e^{av} the average of the sets $\{B_v \ | v \in V(G)\}$ and $\{B_e \ | v \in E(G)\}$, respectively.

The eigenvector centrality of a vertex i is the i-th position of the leading eigenvector of the graph adjacency matrix.

The second smallest eigenvalue of the Laplacian matrix (see [2] for a deeper discussion on the topic) is called algebraic connectivity, λ, and is related to network connectivity.

From the computational complexity point of view, Table I depicts the algorithm space/time complexity of some of the above-mentioned measures. For a deeper discussion on the topic readers are referred to [5]. Table II shows how these measures capture the individual removal of links l_i , l_j and l_r in Figure 2 in the main body of the paper.

TABLE I: Algorithm Space/Time Complexity

Algorithm	**Space**	**Time**
Shortest path Dijkstra	$O(\|V\|^2)$	$O(\|V\|\|E\|log\|E\| + \|V\|)$
Shortest Path (unweighted network)	$O(\|V\|^2)$	$O(\|V\|^2 + \|E\|\|V\|)$
Average Path Length	-	$O(\|V\|\|E\|)$
Diameter	-	$O(\|V\|\|E\|)$
Closeness Centrality	$O(\|V\|)$	$O(\|V\|\|E\|)$
Betweenness Centrality	$O(\|V\|)$	$O(\|V\|\|E\|)$
Edge Betweenness	$O(\|V\|)$	$O(\|V\|\|E\|)$
Eigenvector Centrality	$O(\|V\|)$	$O(\|V\|)$
Clustering coefficient	$O(\|V\|)$	$O(\|V\| \times$ av. degree$)$

TABLE II: Individual removal of links l_i, l_j and l_r in Figure 2 in the main body of the paper.

	BC	d	apl	C	$B_v^{\max}$	B_v^{av}	$B_e^{\max}$	B_e^{av}	λ	$\mathcal{E}$
Original	8	3	1.86	0.8	12	3	16	4	0.35	1.36
l_i	4	∞	1	1	0	0	1	1	0	0.86
l_j	8	4	2.04	0.72	12	3.63	16	4.75	0.32	1.29
l_r	8	3	1.89	0.69	12.5	3.13	16	4.42	0.35	1.32

B. Stochastic Measures

In terms of Information Theory quantifiers, given a network G and a set of measurements $\mathcal{M}$ on G, a stochastic measure associated to G via $\mathcal{M}$ is a probability distribution (histogram) associated with $\mathcal{M}$. Readers should refer to [6] for a deeper discussion on the topic.

In this manuscript, we consider that a stochastic measure is obtained by classical positive measurements computed over each network vertex or edge. Thus, for example, if B_v is the betweenness centrality of the vertex $v \in V(G) = \{1,\ 2, \ldots, N\}$, we can associate a probability distribution defined on $\{1,\ 2, \ldots, N+1\}$ given by

$$P_v(G) = \begin{cases} \frac{B_v}{\sum_{v \in V(G)} B_v}, & \text{if } \sum_{v \in V(G)} B_v > 0 \\ 0, & \text{otherwise} \end{cases} \qquad \text{for all } v \neq N+1$$

$$P_{N+1}(G) = 1 - \textstyle\sum_{v \in V(G)} P_v(G)$$

Given a stochastic measure, we use the Robustness measure, Equation (1) in the main text, to quantify how different their topologies are from each other. Table III shows how these measures capture the individual removal of links l_i, l_j and l_r from the network depicted in Figure 2 in the main body, considering four stochastic measures related to, clustering coefficient, betweenness, closeness and eigenvector centrality.

TABLE III: Network Robustness of Figure 2 in the main body of the paper considering the individual removal of links l_i, l_j and l_r for stochastic measures related to vertex betweenness, closeness, eigenvector and clustering coefficient.

	Betweenness	closeness	Eigenvector	Clustering
l_i	0	0.9958704	0.6870344	0.8774438
l_j	0.9079178	0.9987776	0.9640551	0.9394804
l_r	0.9898528	0.9998521	0.9763837	0.9636882

II. RANDOM FAILURE EXPERIMENT - CLASSICAL ROBUSTNESS MEASURES

In both cases (US Power Grid and Dolphin), 10% of their links were randomly removed and the fractional size of the giant component, the diameter of the giant component, network efficiency and clustering coefficient were computed. For each case, thirty independent experiments were performed, and Figures 1 and 2 depict the outcomes for Dolphin and US Power Grid networks, respectively.

III. CRITICAL NODES OF DOLPHIN AND POWER GRID NETWORKS

US Power Grid - Single Attack for $R_{P_{deg}}$: Figure 3 shows the detection of the ten most critical links and nodes, considering $R_{P_{deg}}$ for the western States Power Grid of the United States network.
Dolphin - Single Attack: Figures 4 and 5 show the detection of the ten percent of the most critical links and nodes, considering $R_{P_{deg}}$ and R_{P_δ}, respectively, in the Dolphin network.

IV. CONTINUOUS DISTRIBUTIONS

Let $\mathbf{P}$, $\mathbf{Q}$ and $\boldsymbol{\mu}$ probability distributions on a probability space $(\mathbf{X}, \mathcal{A})$ such that $\mathbf{P}$ and $\mathbf{Q}$ are absolutely continuous with respect to $\boldsymbol{\mu}$. The Jensen-Shannon divergence between $\mathbf{P}$ and $\mathbf{Q}$ is given by

$$\mathcal{J}^H(P,Q) = H\left(\frac{P+Q}{2}\right) - \frac{H(P)+H(Q)}{2},$$

being the Shannon entropy

$$H(P) = -\int_{\mathbf{X}} p(x) log_2(p(x)) d\boldsymbol{\mu}(x)$$

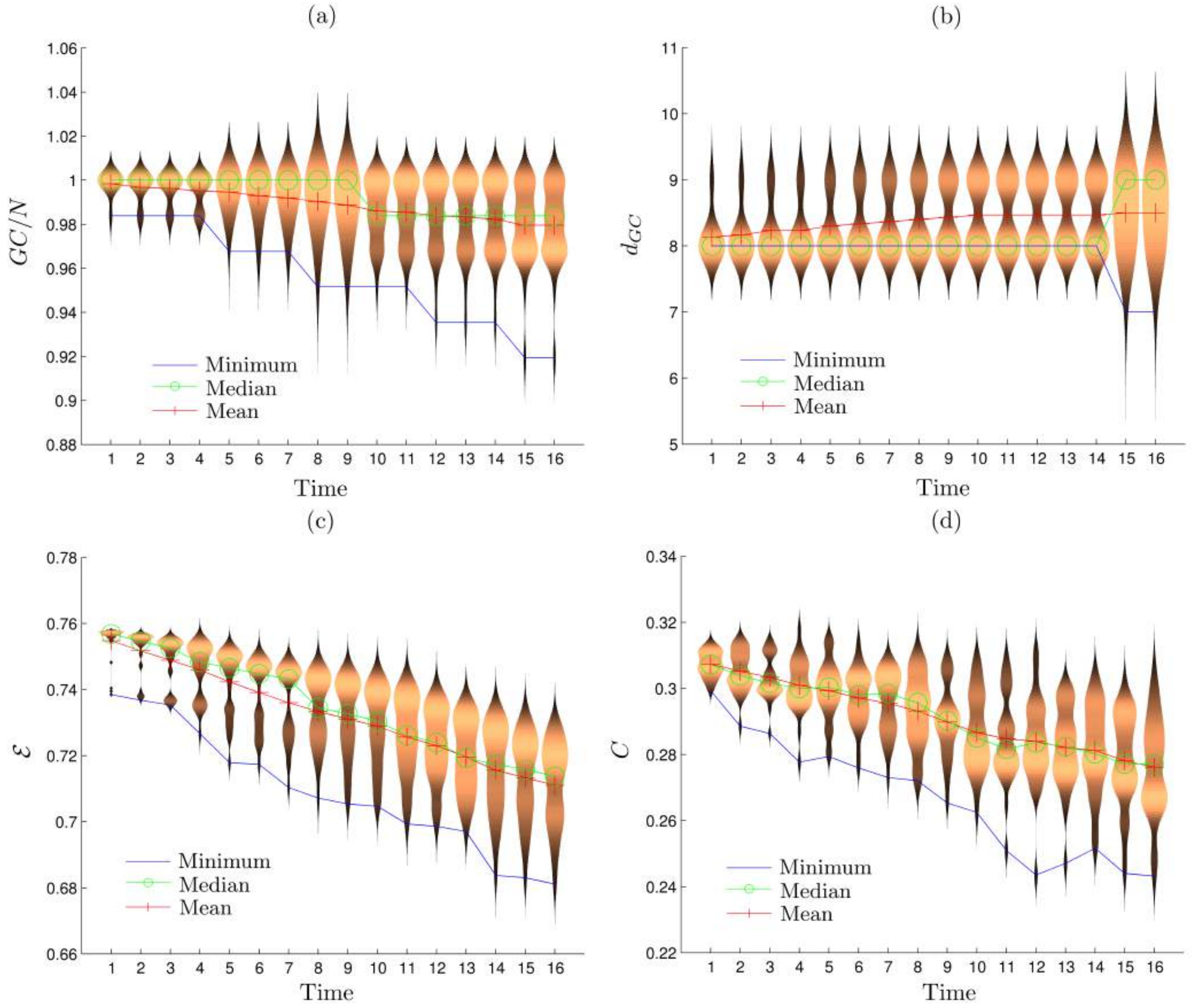


FIG. 1: Random failure experiment in the Dolphin network. Composition of violin plots for (a) fractional size of the giant component, (b) the diameter of the giant component, (c) network efficiency and (d) clustering coefficient with the minimum obtained value of each measure among 30 independent experiments.

Here, we repeat the computation of the random attack experiment considering continuous the distribution of the node betweenness centrality, P_b. Figure 6 depicts the outcomes.

V. ROBUSTNESS OF REAL NETWORKS - A RANDOM FAILURE EXPERIMENT

Here we present a random failure experiment performed on real networks. At each time step, 1% of the original links are randomly removed until the global disconnection of approximately 10% of their links. Thirty independent experiments were performed and the robustness measure for each experiment was computed considering the distance, degree and stochastic measures given by the clustering coefficient, betweenness, closeness and eigenvector centrality. All networks here presented are freely available at *The Koblenz Network Collection* [7]. Descriptions can also be found at the Koblenz website (http://konect.uni-koblenz.de/). Figures 7-11 depict the outcomes.

CAIDA: This is the undirected network of autonomous systems of the Internet connected with each other from the CAIDA project, collected in 2007. Nodes are autonomous systems (AS), and edges denote communication [7, 8]. It contains $26,475$ nodes, $53,381$ edges, clustering coefficient 0.007318732, average path length 3.875647 and diameter

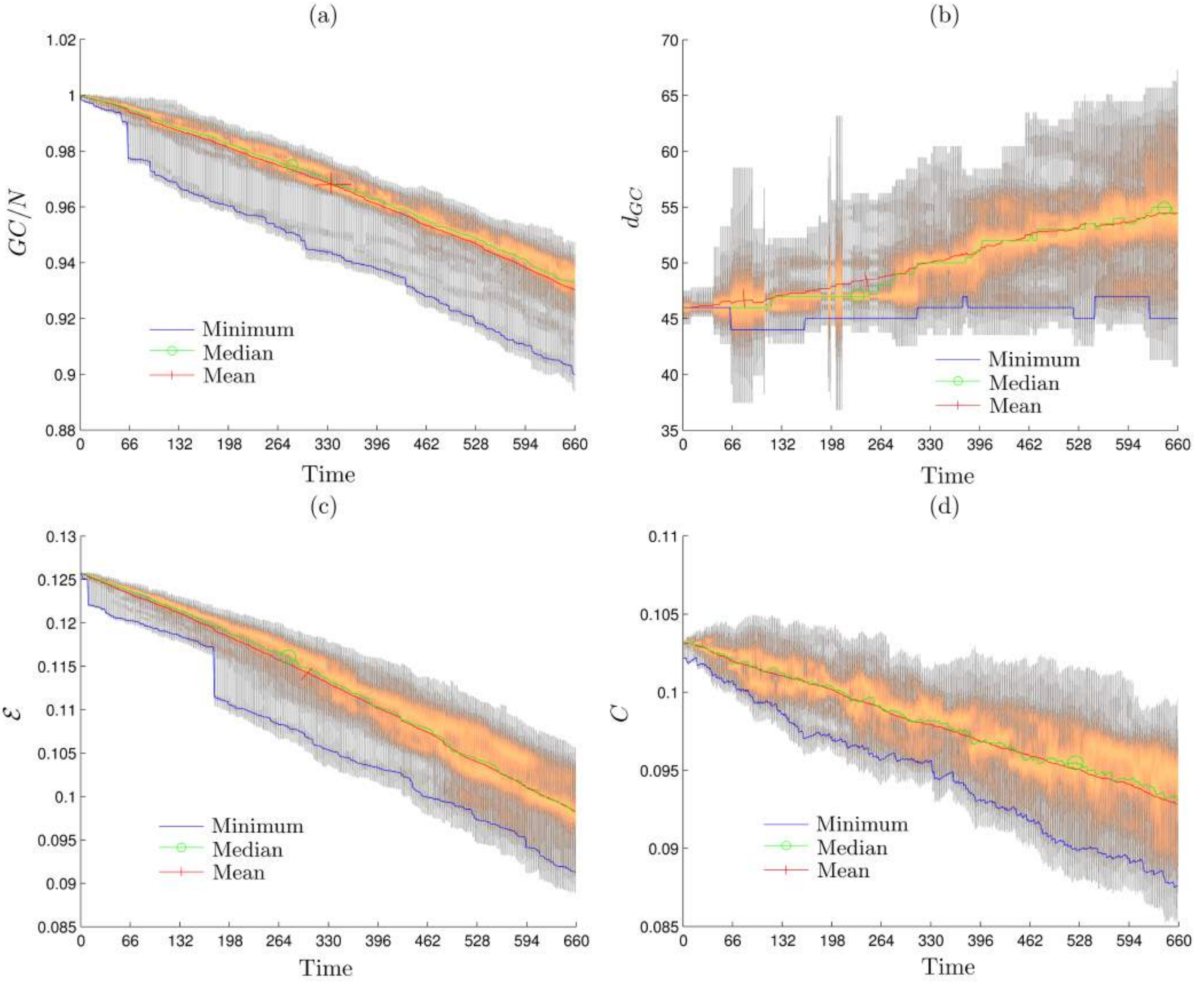


FIG. 2: Random failure experiment in the US Power Grid network. Composition of violin plots for (a) fractional size of the giant component, (b) the diameter of the giant component, (c) network efficiency and (d) clustering coefficient with the minimum obtained value of each measure among 30 independent experiments.

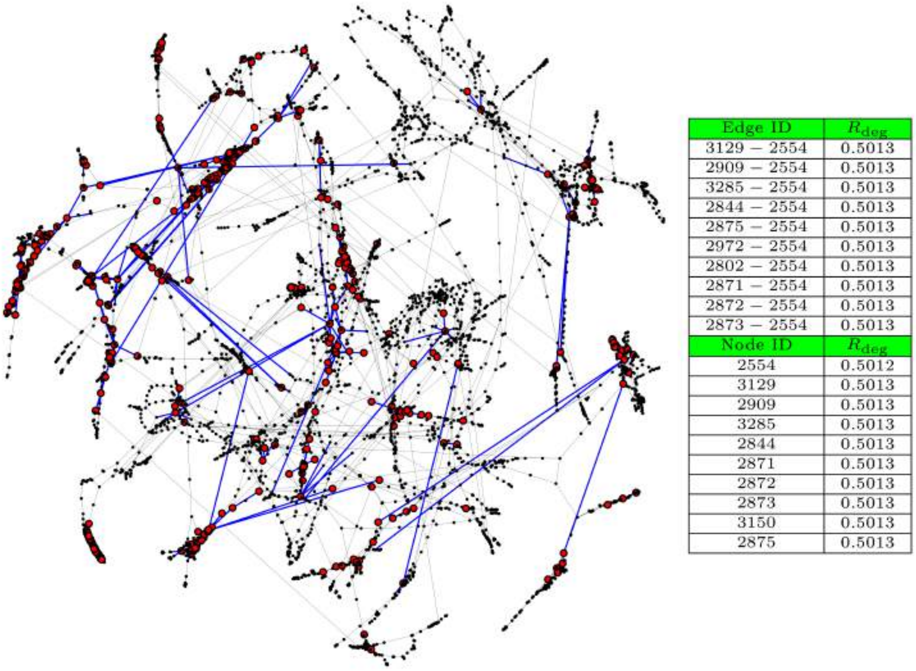
17. See Figure 7.

PGP: This is the interaction network of users of the Pretty Good Privacy (PGP) algorithm. The network contains only the giant connected component of the network [7, 9]. It contains 10, 680 nodes, 24, 316 edges, clustering coefficient 0.378024687, average path length 7.48554 and diameter 24. See Figure 8.

HOMO SAPIENS: This is a network of protein-protein interactions in the species *Homo sapiens*, i.e., in Humans. The data is curated by the Reactome project, an open online database of biological pathways [7, 20]. It contains 6.229 nodes, 146160 edges, clustering coefficient 0.6055493, average path length 4.213577 and diameter 24. See Figure 9.

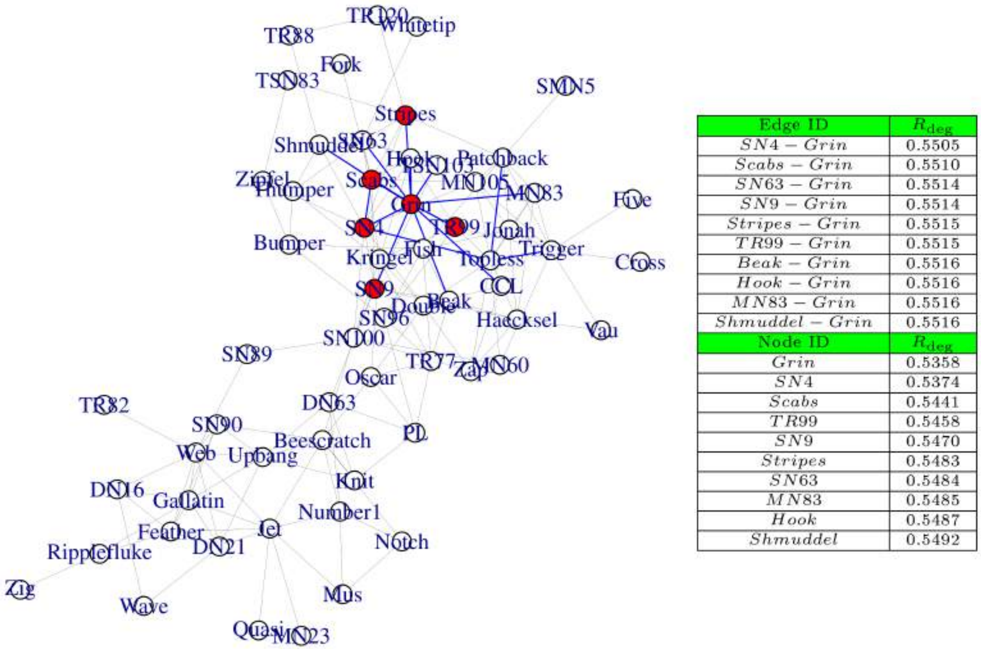
EGO-FACEBOOK: This network contains Facebook user-user friendships. A node represents a user. An edge indicates that two users are friends [7, 21]. It contains 2, 888 nodes, 2, 981 edges, clustering coefficient 0.0003593803, average path length 3.867421 and diameter 9. See Figure 10.

ASTROPH: This is the collaboration graph of authors of scientific papers from the arXiv's Astrophysics (astro-ph) section. An edge between two authors represents a common publication [7, 8]. It contains 18771 nodes, 198050 edges, clustering coefficient 0.3180016, average path length 4.193988 and diameter 14. See Figure 11.



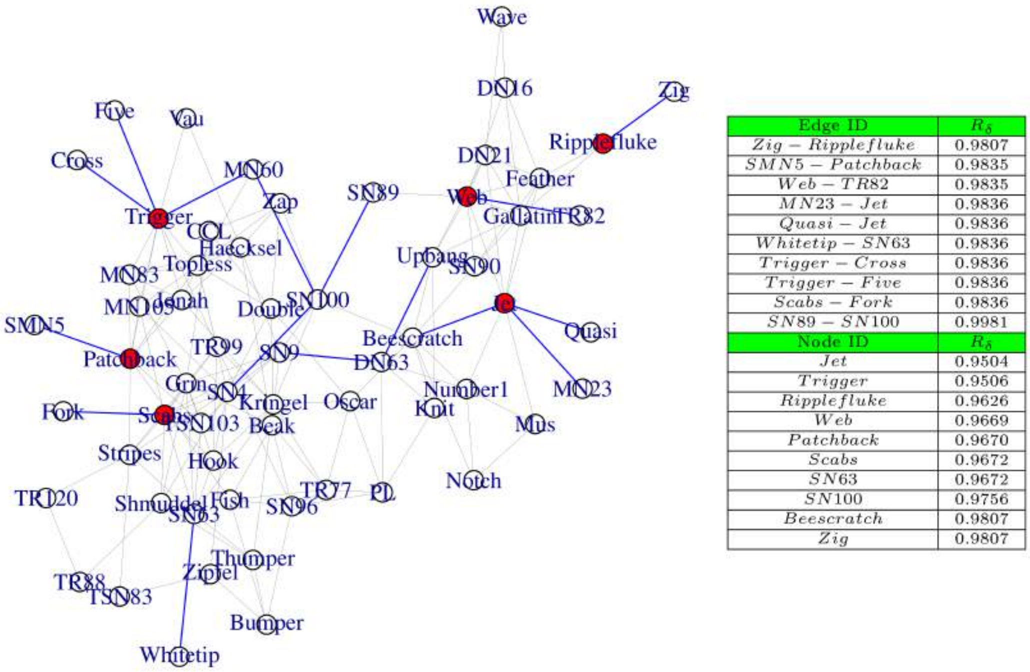
Edge ID	$R_{\deg}$
3129 − 2554	0.5013
2909 − 2554	0.5013
3285 − 2554	0.5013
2844 − 2554	0.5013
2875 − 2554	0.5013
2972 − 2554	0.5013
2802 − 2554	0.5013
2871 − 2554	0.5013
2872 − 2554	0.5013
2873 − 2554	0.5013
Node ID	**$R_{\deg}$**
2554	0.5012
3129	0.5013
2909	0.5013
3285	0.5013
2844	0.5013
2871	0.5013
2872	0.5013
2873	0.5013
3150	0.5013
2875	0.5013

FIG. 3: Detection of ten percent of the most critical links and nodes, considering $R_{P_{deg}}$ in the western States Power Grid of the United States network. Red nodes represent the fraction of 10% of network's vertices such that, its single disconnection causes a big reduction of the $R_{P_{deg}}$ value. Blue links represent the fraction of 10% of network's edges such that, its single disconnection causes a big reduction of the $R_{P_{deg}}$ value. The table shows the robustness values of the top 10 critical network elements.



Edge ID	$R_{\deg}$
$SN4 - Grin$	0.5505
$Scabs - Grin$	0.5510
$SN63 - Grin$	0.5514
$SN9 - Grin$	0.5514
$Stripes - Grin$	0.5515
$TR99 - Grin$	0.5515
$Beak - Grin$	0.5516
$Hook - Grin$	0.5516
$MN83 - Grin$	0.5516
$Shmuddel - Grin$	0.5516
Node ID	**$R_{\deg}$**
$Grin$	0.5358
$SN4$	0.5374
$Scabs$	0.5441
$TR99$	0.5458
$SN9$	0.5470
$Stripes$	0.5483
$SN63$	0.5484
$MN83$	0.5485
$Hook$	0.5487
$Shmuddel$	0.5492

FIG. 4: Detection of the ten percent of the most critical links and nodes, considering $R_{P_{deg}}$ in the Dolphin network. Red nodes represent the fraction of 10% of network's vertices such that, its single disconnection causes a big reduction of the $R_{P_{deg}}$ value. Blue links represent the fraction of 10% of network's edges such that, its single disconnection causes a big reduction of the $R_{P_{deg}}$ value. The table shows the robustness values of the top 10 critical network elements.



Edge ID	R_δ
$Zig - Ripplefluke$	0.9807
$SMN5 - Patchback$	0.9835
$Web - TR82$	0.9835
$MN23 - Jet$	0.9836
$Quasi - Jet$	0.9836
$Whitetip - SN63$	0.9836
$Trigger - Cross$	0.9836
$Trigger - Five$	0.9836
$Scabs - Fork$	0.9836
$SN89 - SN100$	0.9981
Node ID	R_δ
Jet	0.9504
$Trigger$	0.9506
$Ripplefluke$	0.9626
Web	0.9669
$Patchback$	0.9670
$Scabs$	0.9672
$SN63$	0.9672
$SN100$	0.9756
$Beescratch$	0.9807
Zig	0.9807

FIG. 5: Detection of the ten percent of the most critical links and nodes, considering R_{P_δ} in the Dolphin network. Red nodes represent the fraction of 10% of network's vertices such that, its single disconnection causes a big reduction of the R_{P_δ} value. Blue links represent the fraction of 10% of network's edges such that, its single disconnection causes a big reduction of the R_{P_δ} value. The table shows the robustness values of the top 10 critical network elements.

VI. WEIGHTED AND DIRECTED NETWORKS

Weighted networks has been the subject of interest of the scientific community in recent years because most real networks are represented not only by the connection among vertices but also by the strength of these connections. Several measures such as degree, distances between vertices, betweenness, closeness among others has been generalized to such cases [24]. In particular, Newman [23] transformed the weight, in a collaboration network, into costs by inverting them and computing shortest paths between pairs of vertices but, there exists several others distance measurements in networks (see [22] for a deeper discussion on the topic).

Regarding the degree of a vertex in a weighted network, following [24] it is possible to define a degree centrality measure considering both degree (k_v) and weight (s_v) by relating them with a tuning parameter α as

$$C_\alpha(v) = k_v^{1-\alpha} s_v^\alpha$$

and, thus if $\alpha = 0$ the centrality is given only by the degree centrality (the weights are forgotten). By setting $\alpha = 1$ the centrality is given by the total vertex weight (the connections are forgotten).

This methodology can be extended to directed networks as follows:

$$C_\alpha^{in}(v) = (k_v^{in})^{1-\alpha}(s_v^{in})^\alpha \qquad \text{and} \qquad C_\alpha^{out}(v) = (k_v^{out})^{1-\alpha}(s_v^{out})^\alpha \tag{1}$$

being $C_\alpha^{in}(v)$ and $C_\alpha^{out}(v)$ the centralities related to the in and out degrees, respectively.

In order to exemplify our methodology, two real networks were analyzed: Florida ecosystem wet and Florida ecosystem dry. Both networks contains the carbon exchanges in the cypress wetlands of South Florida during the wet and dry seasons, respectively. Nodes represent taxa and an edge denotes that a taxon uses another taxon as food with a given trophic factor (feeding level) [7, 25]. The networks possess 128 vertices. The experiment consists in the attack of the most central nodes of the network given by Equation (1) for a given α. In the beginning of the process the centralities are computed and the sequence of the attack is determined. At each time step, the most central vertex is disconnected from the network until its the complete disconnection. The robustness is computed via stochastic measure related with the betweenness centrality distribution, where distances were computed via costs given by the inverse of the weights. Figure 12 depicts the outcomes.

From Figure 12(C) it is possible to see that when 10% of the nodes are disconnected, for $\alpha = 0$ (without considering edges weights) the most efficient strategy to attack the network is given by the sequence of the most central in-degree

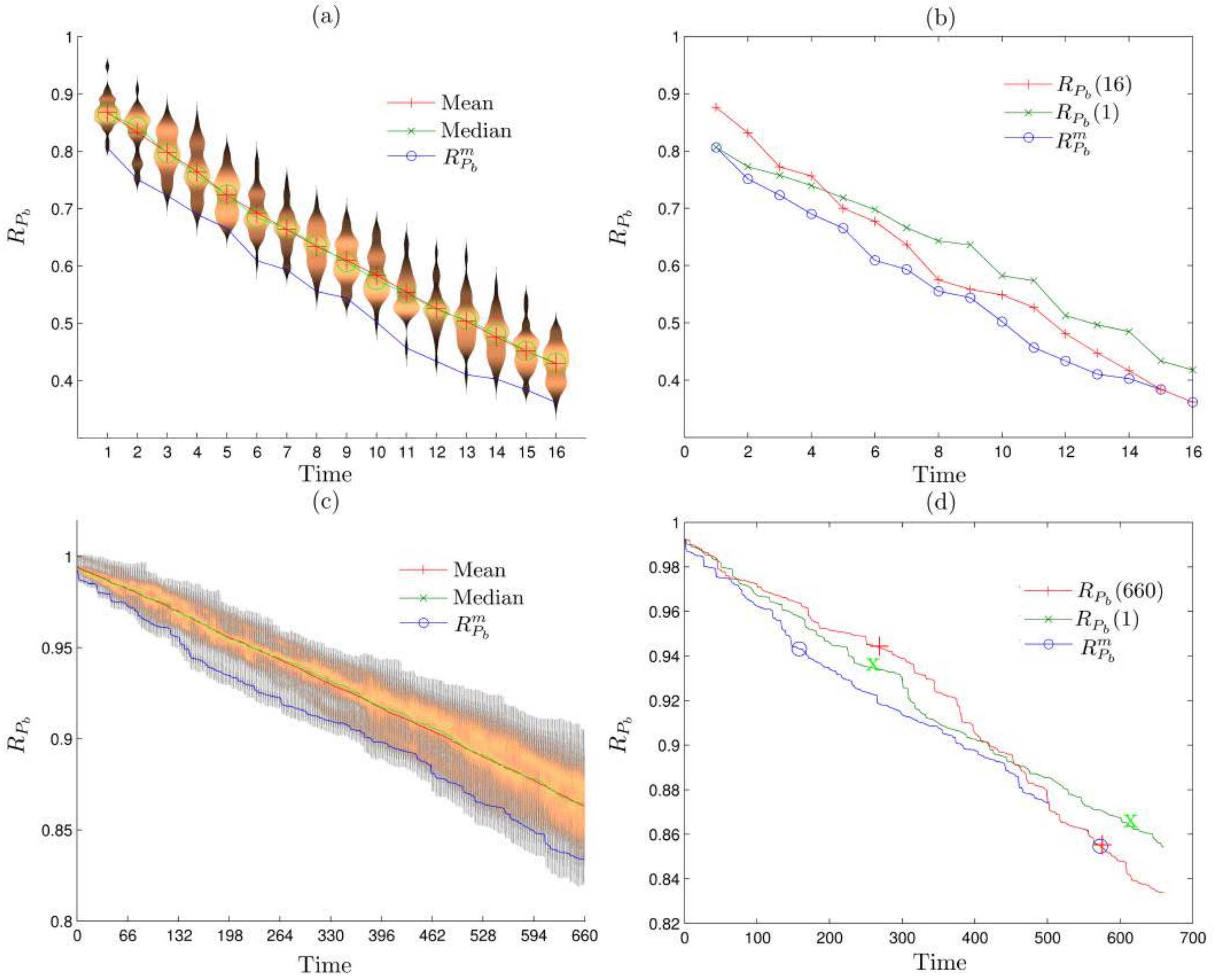


FIG. 6: Robustness measures considering the node betweenness centrality distribution under random failures in the Dolphin and Power Grid networks. At each time step, a random edge is disconnected from the network and R_{P_b} function is computed. The experiment is independently executed 30 times and the violin plots are presented for (a) for Dolphin and (c) for US Power Grid. (b) and (d) also show a comparison between two different attacks: one most effective to destroy the network at the beggining of the process, $R_{P_b}(1)$, and the other at the end ($R_{P_b}(16)$ for Dolphin and $R_{P_b}(660)$ for US Power Grid, respectively). In all cases, at each time step, the minimum robustness values are also indicated by R^m_P.

nodes in both dry and wet networks but the opposite occurs for $\alpha = 0.4$, for example. The minimum robustness measure for both networks is achieved for $\alpha = 0.2$ and in-attack strategy. From Figure 12(D) it is possible to see that when 50% of the nodes are disconnected, for $\alpha = 0$ (without considering edges weights) the most efficient strategy to attack the network is given by the sequence of the most central out-degree nodes in both dry and wet networks but the opposite occurs for $\alpha = 0.3$, for example. The minimum robustness measure for both networks is achieved for $\alpha = 0.1$ and out-attack strategy differently when only 10% of the nodes are disconnected.

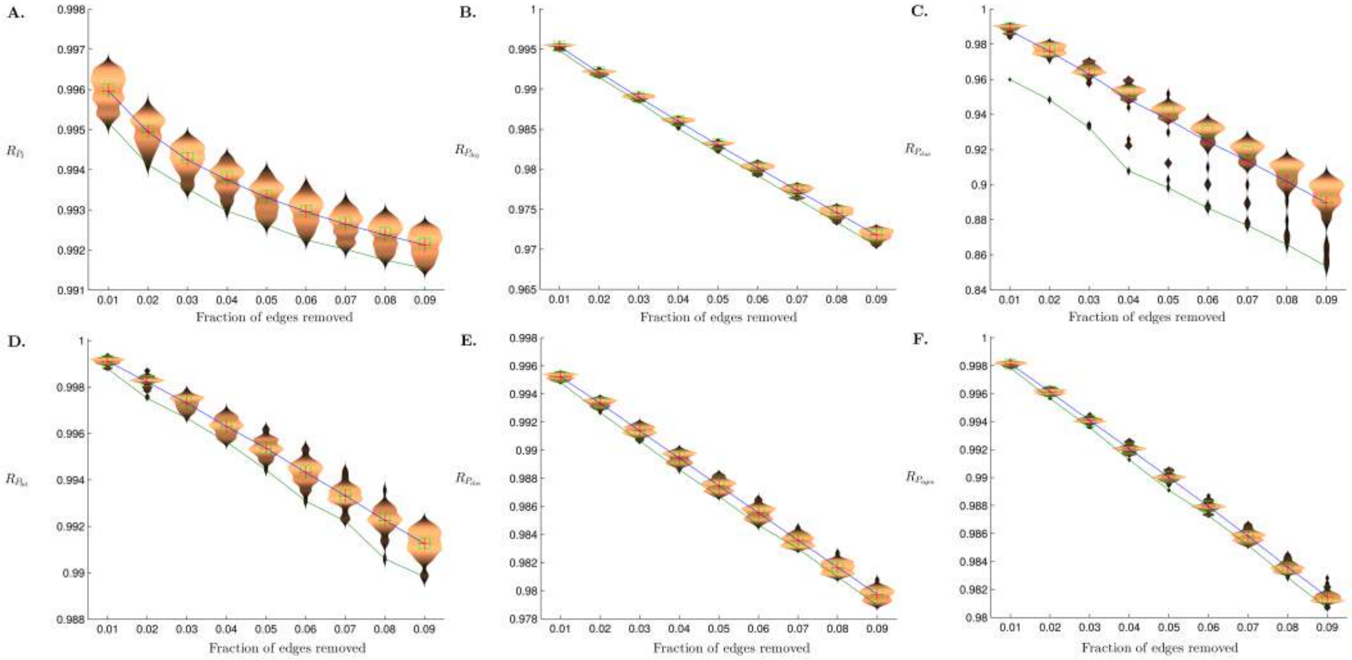


FIG. 7: Violin plots for robustness measures under random failures for CAIDA network considering: the distance distribution (A.), the degree distribution (B.), and stochastic measures related to clustering coefficient (C.), betweenness (D.), closeness (E.) and eigenvector centrality (F.). At each time step, 1% of the original links are randomly removed until the global disconnection of approximately 10% of their links.

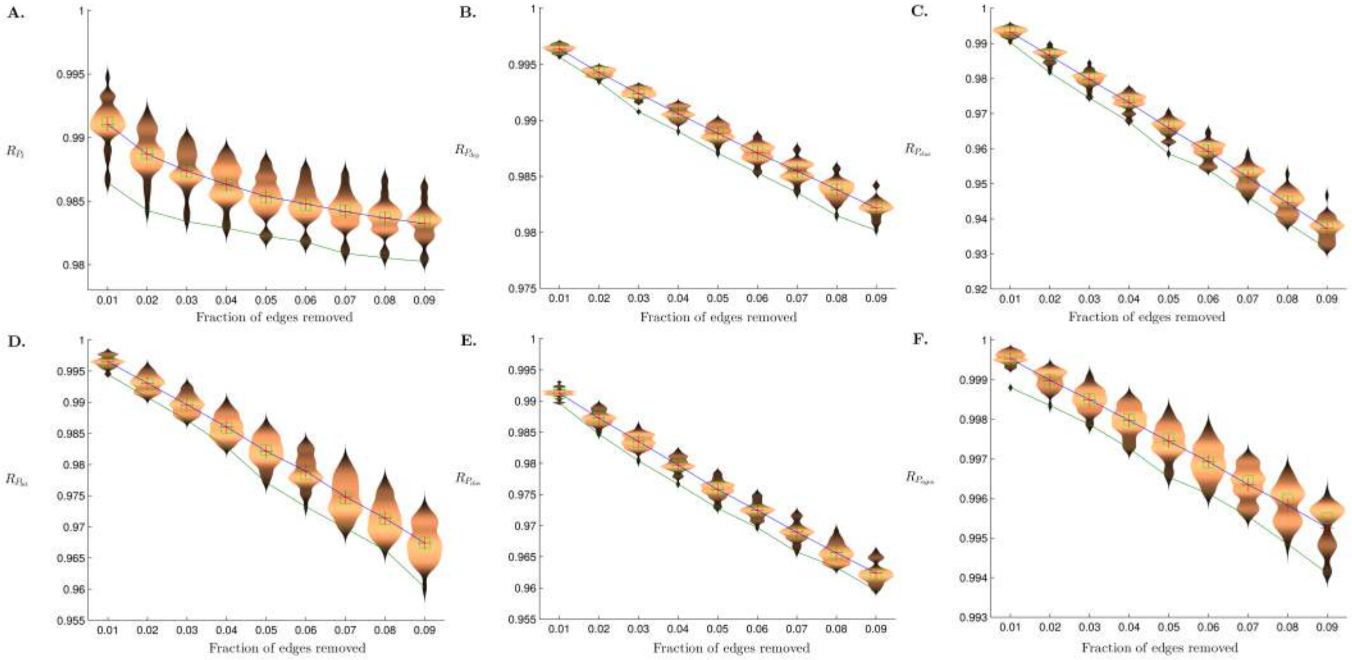


FIG. 8: Violin plots for robustness measures under random failures for PGP network considering: the distance distribution (A.), the degree distribution (B.), and stochastic measures related to clustering coefficient (C.), betweenness (D.), closeness (E.) and eigenvector centrality (F.). At each time step, 1% of the original links are randomly removed until the global disconnection of approximately 10% of their links.

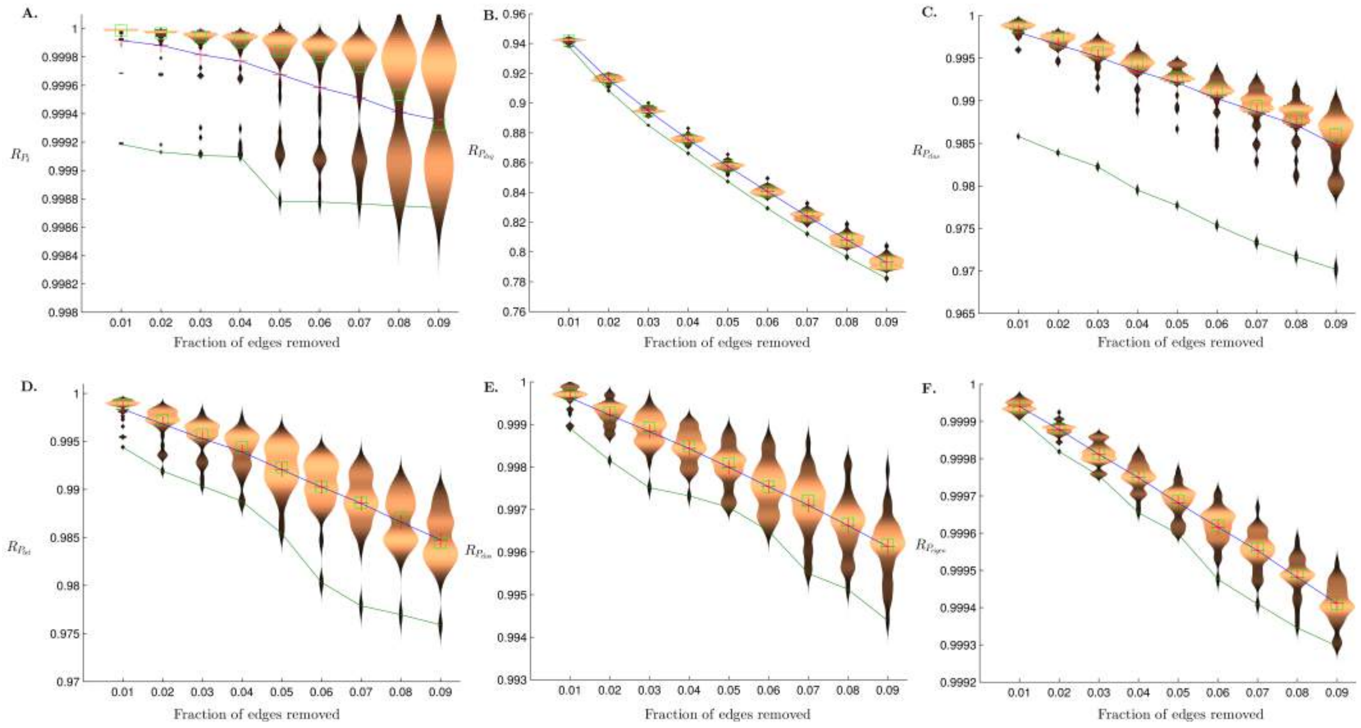


FIG. 9: Violin plots for robustness measures under random failures for HOMO SAPIENS network considering: the distance distribution (A.), the degree distribution (B.), and stochastic measures related to clustering coefficient (C.), betweenness (D.), closeness (E.) and eigenvector centrality (F.). At each time step, 1% of the original links are randomly removed until the global disconnection of approximately 10% of their links.

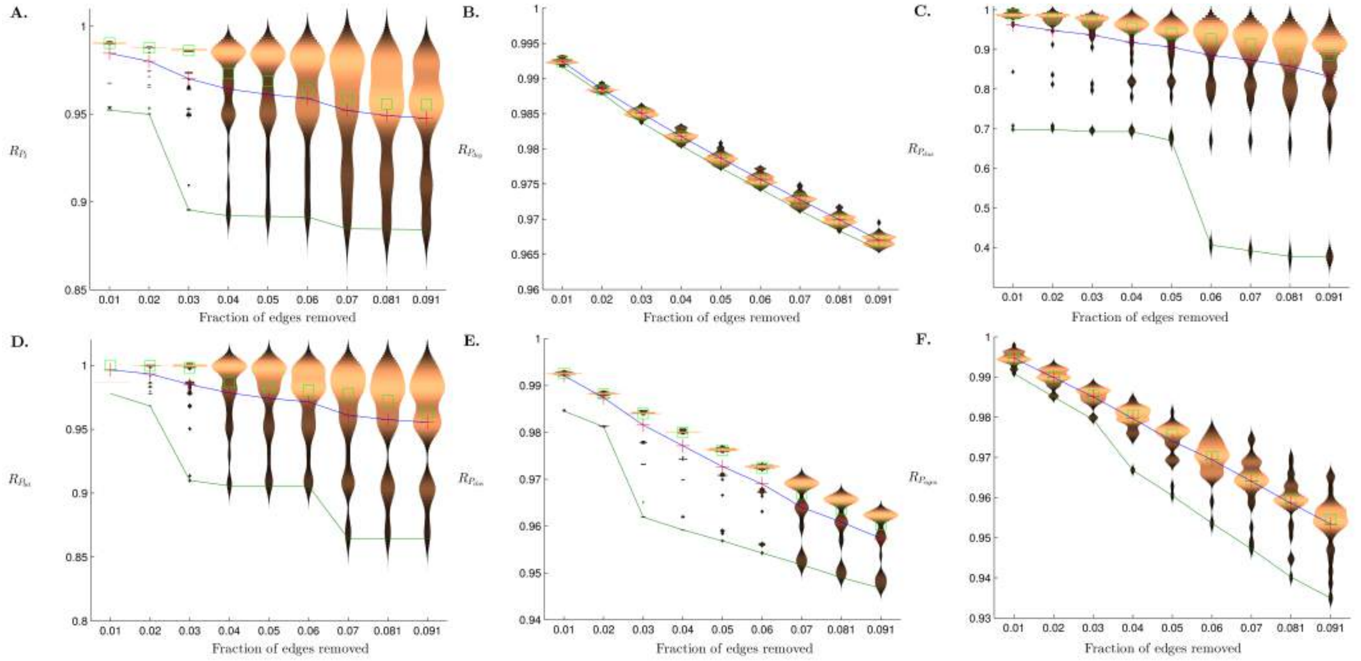


FIG. 10: Violin plots for robustness measures under random failures for EGO-FACEBOOK network considering: the distance distribution (A.), the degree distribution (B.), and stochastic measures related to clustering coefficient (C.), betweenness (D.), closeness (E.) and eigenvector centrality (F.). At each time step, 1% of the original links are randomly removed until the global disconnection of approximately 10% of their links.

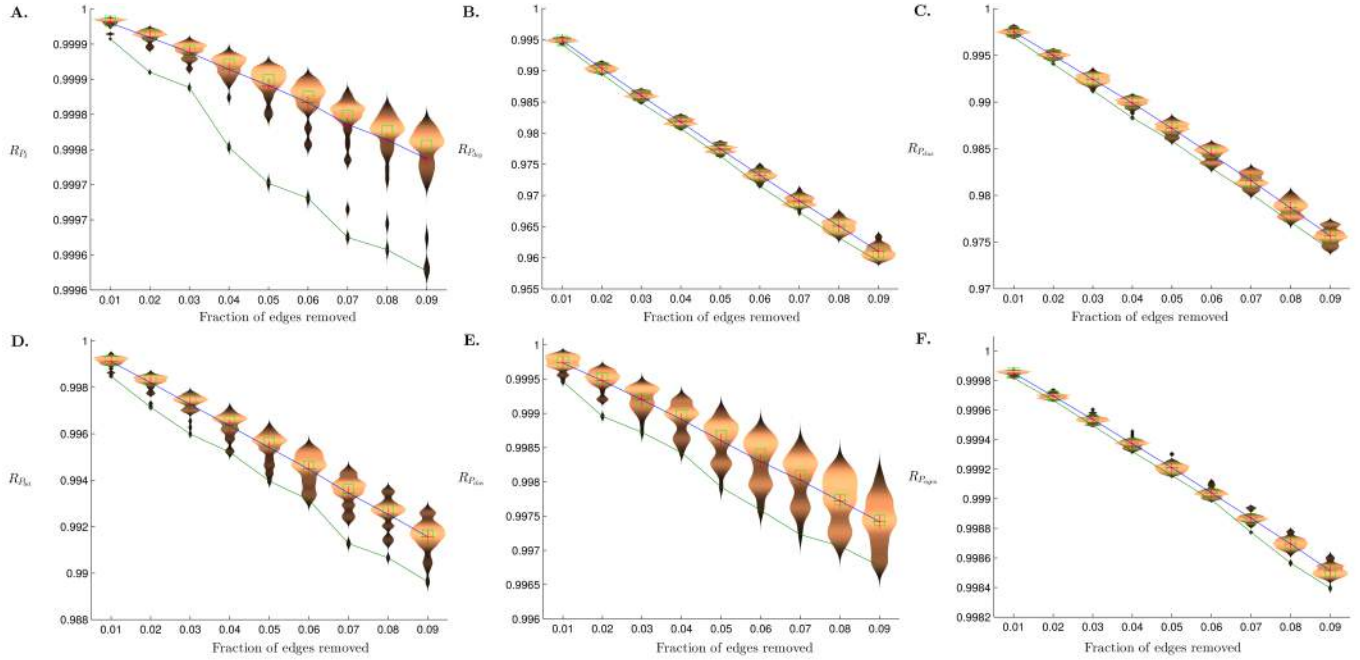


FIG. 11: Violin plots for robustness measures under random failures for ASTROPH network considering: the distance distribution (A.), the degree distribution (B.), and stochastic measures related to clustering coefficient (C.), betweenness (D.), closeness (E.) and eigenvector centrality (F.). At each time step, 1% of the original links are randomly removed until the global disconnection of approximately 10% of their links.

VII. ACKNOWLEDGMENTS

This work was partially supported by FAPEMIG, CNPq and grant RSF 14-41-00039

[1] R. Albert,H. Jeong and A.-L. Barabási, *Error and attack tolerance of complex networks*, Nature, 406 (2000), pp. 378–382.

[2] M. Fiedler, *Algebraic connectivity of graphs*, Czechoslovak Mathematical Journal, 23 (1973), pp. 298–305.

[3] A. H. Dekker and B. D. Colbert, *Network Robustness and Graph Topology* in Proceedings of the 27th Australasian Conference on Computer Science, Australian Computer Society, Inc., 26 (2004), pp. 359–368.

[4] L. Rodrigues, G. Travieso and P. R. Villas Boas, *Characterization of complex networks: A survey of measurements*, Advances in Physics, 56 (2006), pp. 167–242.

[5] Gabor Csardi and Tamas Nepusz, *The igraph software package for complex network research*, InterJournal, Complex Systems, (2006), pp. 1695.

[6] R.S. CABRAL and A.C FRERY. and J. A. RAMREZ. *Variability Analysis of Complex Networks Measures based on Stochastic Distances.* Physica A (Print), (2014) v. 415, pp. 73-86.

[7] Jérôme Kunegis. *KONECT - The Koblenz Network Collection.* Proc. Int. Web Observatory Workshop, (2013).

[8] J. Leskovec and J. Kleinberg and C. Faloutsos. *Graph Evolution: Densification and Shrinking Diameters.* ACM Trans. Knowledge Discovery from Data, (2007). v. 1, pp.1-40,

[9] M. Boguñá and R. Pastor-Satorras and A. Díaz-Guilera and A. Arenas. *Models of Social Networks based on Social Distance Attachment.* Phys. Rev. E, (2004), v. 70, pp.056122.

[10] R. Guimerà and L. Danon and A. DÍaz-Guilera and F. Giralt, and A. Arenas., *Self-similar community structure in a network of human interactions.* Phys. Rev. E., (2003), v.68, pp. 065103.

[11] T. Opsahl and P. Panzarasa. *Clustering in weighted networks.*, Social Networks, (2009), n. 2, pp. 155-163, v. 31.

[12] P. M. Gleiser and L. Danon. *Community structure in jazz.*, Advances in Complex Systems, (2003), n.4, pp. 565-573, v.6.

[13] W. Zachary., *An information flow model for conflict and fission in small groups.*, J. of Anthropological Research, (1977), pp. 452-473, v. 33.

[14] Lorenzo Isella and Juliette Stehlé and Alain Barrat and Ciro Cattuto and Jean-François Pinton and Wouter Van den Broeck.,*What's in a crowd? analysis of face-to-face behavioral networks.*, J. of Theoretical Biology, (2011), n.1, pp. 166-180, v. 271.

[15] Donald E. Knuth., *The Art of Computer Programming*, Addison-Wesley, (2008), v. 4.

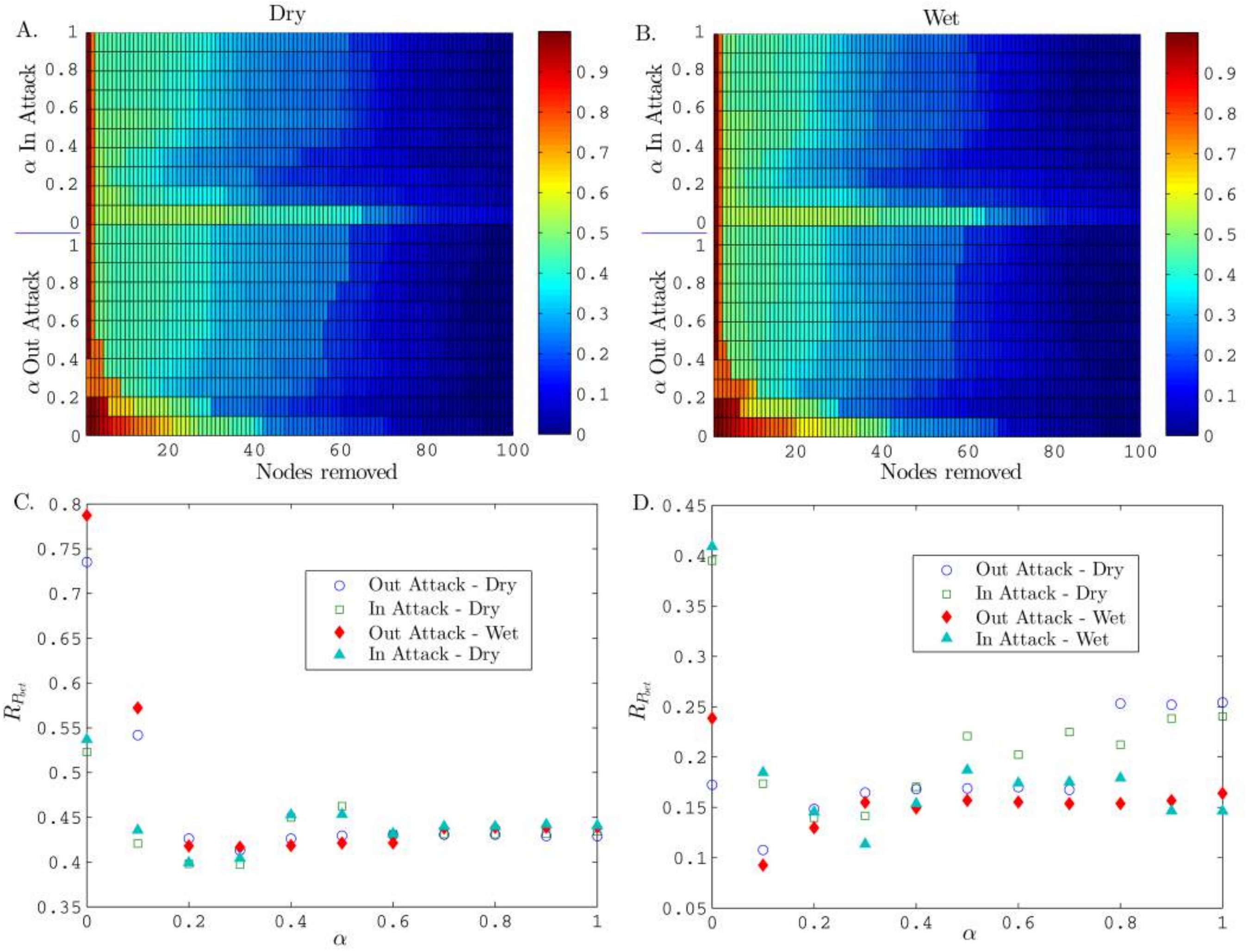


FIG. 12: Intentional attack of Florida ecosystem wet and dry networks. (A.) shows the $R_{P_{bet}}$ values for the Dry network for different values of α and attack strategies: bigger C_{α}^{in} (α in attack) and bigger C_{α}^{out} (α out attack). (B.) shows the $R_{P_{bet}}$ values for the Wet network for different values of α and attack strategies: bigger C_{α}^{in} (α in attack) and bigger C_{α}^{out} (α out attack). (C.) shows how the robustness values changes when approximately 10% of the nodes are disconnected for different values of α. (D.) shows how the robustness values changes when approximately 50% of the nodes are disconnected for different values of α.

[16] Lovro Subelj and Marko Bajec, *Robust network community detection using balanced propagation.*, Eur. Phys. J. B, (2011), n. 3, pp. 353-362, v. 81.

[17] Duncan J. Watts and Steven H. Strogatz., *Collective dynamics of small-world networks*, Nature, (1998), n. 1, pp. 440-442, v. 393.

[18] Mark E. J. Newman, *Finding community structure in networks using the eigenvectors of matrices*, Physical Review E, (2006), n. 3, v. 74.

[19] Han, Jing-Dong J and Dupuy, Denis and Bertin, Nicolas and Cusick, Michael E and Vidal, Marc, *Effect of Sampling on Topology Predictions of Protein-protein Interaction Networks*, Nature Biotechnology, (2005), n. 7, pp. 839–844, v. 23,

[20] Joshi-Tope, G. and Gillespie, Marc and Vastrik, Imre and D'Eustachio, Peter and Schmidt, Esther and de Bono, Bernard and Jassal, Bijay and Gopinath, GR and Wu, GR and Matthews, Lisa and others, *Reactome: A Knowledgebase of Biological Pathways*, Nucleic Acids Research, (2005), n. suppl 1, pp. D428–D432, v. 33,

[21] Julian McAuley and Jure Leskovec, *Learning to Discover Social Circles in Ego Networks*, Advances in Neural Information Processing Systems, (2012),

[22] M. M. Deza and E. Deza, *Distances in Networks*, Encyclopedia of Distances, Springer Berlin Heidelberg, (2014), pp. 413-428,

[23] M. E. J. Newman, *Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality.*, (2001), Physical Review E 64, 016132.

[24] T. Opsahl and F. Agneessens and J. Skvoretz, *Node centrality in weighted networks: Generalizing degree and shortest*

paths., (2010), Social Networks, n. 3, pp. 245-251, V. 32.

[25] R. E. Ulanowicz and J. J. Heymans and M. S. Egnotovich. *Network analysis of trophic dynamics in South Florida ecosystems*, FY 99: The graminoid ecosystem. Annual Report to the United States Geological Service Biological Resources Division Ref. No.[UMCES] CBL 00-0176, Chesapeake Biological Laboratory, University of Maryland, 2000.