A Review on Network Robustness from an Information Theory Perspective

Tiago Schieber[1,2](✉), Martín Ravetti[1], and Panos M. Pardalos[3,4]

[1] Departamento de Engenharia de Produção,
Universidade Federal de Minas Gerais, Belo Horizonte, MG, Brazil
tischieber@gmail.com

[2] Departamento de Engenharia de Produção,
Pontifícia Universidade Católica de Minas Gerais, Belo Horizonte, MG, Brazil

[3] Center for Applied Optimization, Industrial and Systems Engineering,
University of Florida, Gainesville, Florida, USA

[4] Laboratory of Algorithms and Technologies for Network Analysis,
National Research University Higher School of Economics, Nizhny Novgorod, Russia

Abstract. The understanding of how a networked system behaves and keeps its topological features when facing element failures is essential in several applications ranging from biological to social networks. In this context, one of the most discussed and important topics is the ability to distinguish similarities between networks. A probabilistic approach already showed useful in graph comparisons when representing the network structure as a set of probability distributions, and, together with the Jensen-Shannon divergence, allows to quantify dissimilarities between graphs. The goal of this article is to compare these methodologies for the analysis of network comparisons and robustness.

1 Introduction

Quantification of dissimilarities between graphs has been a central subject in graph theory for many decades. With the complex networks field, we witness a burst of applications on real systems where the measure of graph or subgraph similarities have played a major role. Several methods for this quantification have become increasingly addressed, where most approaches are based on invariant measurements under graph isomorphism [1–6]. Although there exists in the literature a quasi-polynomial time algorithm to solve graph isomorphism [7], still, an efficient way to decide if two structures are isomorphic continues an open problem, as the search for efficient pseudo-distances between networks.

Representing a network as a set of stochastic measures (probability distributions associated with a given set of measurements) showed useful to characterize network evolution, robustness and efficiently treat the graph isomorphism problem [6,8–10].

These characteristics are useful to define a pseudo-metric between networks via the Jensen-Shannon divergence, an Information Theory quantifier that

© Springer International Publishing Switzerland 2016

Y. Kochetov et al. (Eds.): DOOR 2016, LNCS 9869, pp. 50–60, 2016.
DOI: 10.1007/978-3-319-44914-2_5

already showed very effective in measuring small network topology changes [6,9–11]. When comparing n probability distributions, it is given by the Shannon entropy of the average minus the average of the Shannon entropies and, it was proven to be a bounded *square of a metric* between probability distributions [28], here defined for the discrete case:

$$JS(\mathbf{P}_1, \mathbf{P}_2, \ldots, \mathbf{P}_n) = H\left(\frac{\sum_{i=1}^{n} \mathbf{P}_i}{n}\right) - \frac{\sum_{i=1}^{n} H(\mathbf{P}_i)}{n} \tag{1}$$

being $H(\mathbf{P}) = -\sum_i p_i \log p_i$ the Shannon entropy of $\mathbf{P}$.

The JS divergence (Eq. 1) possesses a lower bound equals zero and an upper bound equals $\log n$. The zero value means that all probabilities are equal to the same distribution $\mathbf{P}_1 = \mathbf{P}_2 = \cdots = \mathbf{P}_n = \mathbf{P}$. A $\log n$ value gives the biggest uncertainty when comparing $\mathbf{P}_1,\ \mathbf{P}_2, \ldots, \mathbf{P}_n$ since $\log n$ is the biggest entropy value achieved only by the uniform distribution.

The metric property of the square root of the JS divergence, together with stochastic measures on networks, allows to define two pseudo-metrics between networks: one given only by global properties (D^g) representing the network as a single probability distribution and, the other, more precise but more computationally expensive (D), considering local network characteristics by representing the network as a set of probability distributions.

The analysis of properties of complex networks, therefore, relies on using stochastic measurements capable of expressing the most relevant topological features. Depending on the network and application, a specific set of stochastic measures could be chosen. This article presents a survey of such measurements. It includes classical complex network measurements, applications on network evolution, comparisons and robustness.

2 Methodology

A network G is a pair $(V, \mathbb{E})$, where V is a set of nodes (or vertices), and $\mathbb{E}$ is a set of ordered pairs of distinct nodes, which we call edges. A *weighted network* associates a weight (ω_e) to every edge $e \in \mathbb{E}$, characterizing not only the connections among vertices but also the strength of these connections.

Exists, in the literature, several measurements representing network connectivity. In particular, most real networks present small average distance between elements and high-density communities.

The in-degree (out-degree) of a node, k^{in} (k^{out}), is the number of incoming (outgoing) edges. The in-weight (out-weight) of a node, ω^{in} (ω^{out}), is the sum of all incoming (outgoing) edge weights. Following [29] it is possible to define a degree centrality measure considering both degree and weight by relating them to a tuning parameter $\alpha \in [0, 1]$ as:

$$\kappa_\alpha^{in}(v) = (k_v^{in})^{1-\alpha}(\omega_v^{in})^\alpha \qquad \text{and} \qquad \kappa_\alpha^{out}(v) = (k_v^{out})^{1-\alpha}(\omega_v^{out})^\alpha. \tag{2}$$

If $\alpha = 0$, the weights are forgotten to obtain the node degree. As α increases the number of connections loses in importance and, when α reaches 1, the centrality is given by the total vertex weight.

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, if there is no path between them, $d(i,j) = \infty$. In a weighted network, there are several distances measures in literature because the strength of these connections sometimes implies in small distances between the nodes. In an e-mail network, a bigger edge weight value may represent a frequent communication and, therefore, a small distance between them. Here, we consider the same approach used in [12] transforming weights into costs by inverting them and computing shortest paths between pairs of nodes. Readers should refer to [13] for a deeper discussion on the topic.

The network diameter (average path length) is the maximum (average) distance between all pairs of connected nodes.

The clustering coefficient (C), also known as transitivity, characterizes triangles in the network. It is the fraction of the number of triangles and the number of connected triples. Thus, a complete graph possesses $C = 1$ and, a tree graph, $C = 0$. Analogously, the vertex clustering coefficient, C_v, is given by:

$$C_v = \frac{3n_{\Delta}(v)}{n_e(v)},$$

being, $n_{\Delta}(v)$ the number of triangles involving node i and $n_3(v)$ the number of connected triples having v as a central vertex. A node clustering coefficient value equals 1 means that there is a connection between all pairs of its first neighbors, and a zero value represents the lack edges between them.

The closeness centrality measure of a node is the sum of the inverse of all pairs of distances from it:

$$c_v = \sum_{j,\ j \neq v} \frac{1}{d(v,j)}.$$

A high closeness centrality value means that the node possesses a lower total distance from all other nodes.

Betweenness centrality quantifies node importance in terms of interactions via the shortest paths among all other nodes:

$$B_v = \sum_{i \neq j \in V(G)} \frac{n(i,j,v)}{2n(i,j)},$$

being, $n(i,j)$ the number of shortest paths connecting i and j and $n(i,j,x)$ the number of shortest paths connecting i and j passing through x.

See Table 1 for space and time computational complexity of the above mentioned measures.

Given two networks G_1 and G_2 and two stochastic measurements $\mathbf{P}_{G_1}$ and $\mathbf{P}_{G_2}$, the global pseudo-metric

$$D^g_{\mathbf{P}}(G_1, G_2) = \sqrt{\frac{JS(\mathbf{P}_{G_1}, \mathbf{P}_{G_2})}{\log 2}}. \tag{3}$$

measures how far away two networks are via probability distributions.

Table 1. Space/time computational complexity in a network with N nodes and E edges.

	Space	Time
Degree	$O(N)$	$O(N^2)$
All pairs of distances (unweighted)	$O(N^2)$	$O(N^2 + NE)$
Local clustering coefficient	$O(N)$	$O(N^3)$
Closeness	$O(N)$	$O(NE)$
Betweenness	$O(N)$	$O(NE)$

The degree distribution $\mathbf{P}_{deg}(k)$ is the fraction of nodes with degree k. The network distance distribution, $\mathbf{P}_{\delta}(d)$, gives the fraction of pairs of nodes at distance d. Analogously, $\mathbf{P}_{B_v}$, $\mathbf{P}_c$ and $\mathbf{P}_C$ are given, respectively, by distributions of the betweennesss, closeness and local clustering coefficient.

Here, we consider five variations of the D^g function (Eq. (3)) associated with the stochastic measures given by the degree ($D^g_{\mathbf{P}_{deg}}$), distance ($D^g_{\mathbf{P}_{\delta}}$), closeness ($D^g_{\mathbf{P}_c}$), betweenness ($D^g_{\mathbf{P}_{bet}}$) and clustering coefficient ($D^g_{\mathbf{P}_C}$) distributions.

We can also obtain local information from the stochastic measure. We focus our attention on the node distance distribution ($\mathbf{P}_{\delta,v}(d)$) given by fraction of nodes at distance d from each node v. The network node dispersion (NND), a network quantifier related to the heterogeneity of nodes, introduced in [10] to a network G of size n:

$$NND(G) = JS(\mathbf{P}_{\delta,1}, \mathbf{P}_{\delta,2}, \ldots, \mathbf{P}_{\delta,n})$$

allows, together with the global pseudo-metric associated with the distance distribution ($D^g_{\mathbf{P}_{\delta}}$), to have an efficient size independent pseudo-metric between networks:

$$D(G_1, G_2) = \frac{1}{2} D^g_{\mathbf{P}_{\delta}}(G_1, G_2) + \frac{1}{2} \left| \sqrt{\frac{NND(G_1)}{\log n}} - \sqrt{\frac{NND(G_2)}{\log m}} \right|, \qquad (4)$$

being, n and m, the sizes of networks G_1 and G_2, respectively.

Each global dissimilarity measure captures different characteristics. Most real networks present a degree distribution following a power-law $\mathbf{P}_{deg}(k) \sim k^{-\gamma}$ [16] but, there exist several networks with different topologies sharing the same degree distribution. The clustering based dissimilarity measures how far away two networks are comparing connected communities densities but, it fails to characterize properly tree-like structures. Distance based measures capture important features on networks: from the distance distribution, it is possible to obtain the network diameter, average path length, and average degree. From the node distance distribution perspective, as more information are available, we also get the node degree, closeness centrality, among others.

3 Applications

3.1 Distance Between Null Models

Here we compare how well-known networks null models are away from each other using the D^g and D functions. We consider four of the most commonly used models: K-regular [14], Erdös-Renyi (ER) [15], Barabási-Albert (BA) [16], Exponential (EXP) [17] and Watts-Strogatz rewiring model (WS) [18].

The K-regular consists in generating random networks with a constant degree K. ER is the random graph generation given by a connection probability $p \in [0,1]$. Both BA and EXP are models of evolving networks: at each time step a new node is added and connected to m other existing nodes but, in the Exponential model, the new node is connected at random and the BA uses a preferential attachment mechanism[1]. WS model generates random networks by rewiring, with a given probability, links from a regular lattice.

The experiment consists in generate 10000 independent samples of each model with a fixed size $N = 1000$ computing averaged stochastic measures for each null model and then get comparisons via D^g and D. We set the parameters aiming to preserve the average degree of all generated networks: 10-Regular, BA and EXP with parameter $m = 5$, ER with $p = 10/999$ and WS with $k = 5$ and different rewiring probabilities $p = 0.2,\ 0.4,\ 0.6,\ 0.8$. Figure 1 shows the multidimensional scaling map [19] performed over the outcomes.

All of the analyzed measures were able to capture the scale-free behavior of the BA model ($P(k) \sim k^{-3}$) identifying significant structural differences even when compared with a similar growing model like the EXP, highlighting how different is the preferential attachment procedure in growing networks. It is also possible to see that bigger rewiring probability values imply higher proximity between WS and ER models [11]. As p increases, the randomness of the WS networks also increases. Figures 1B and C show the dissimilarity function importance: B shows that the average of the distance distributions of the ER network approaches the distance distribution of the regular graph meaning that, on average, a random graph behaves like a regular one but, the NND value is zero in most regular networks (Fig. 1C).

3.2 Critical Element Detection Problem and Network Robustness

The knowledge about how the network behaves after failures is of paramount importance and, therefore, the detection critical elements are important to plan efficient strategies to protect or even to destroy networks.

Given a network and an integer k, the critical element detection problem is to find a set of at most k elements (nodes or edges), whose deletion generates the biggest topological difference when comparing the residual and the original networks [20–22].

Here, we consider finding the critical 3 nodes in the Infectious Sociopatterns network whose deletion generates the biggest $D^g_{\mathbf{P}_{deg}}$, $D^g_{\mathbf{P}_{\delta}}$, $D^g_{\mathbf{P}_{bet}}$, $D^g_{\mathbf{P}_{c}}$, $D^g_{\mathbf{P}_{C}}$ and

[1] Higher degree nodes have a bigger probability of getting new connections.

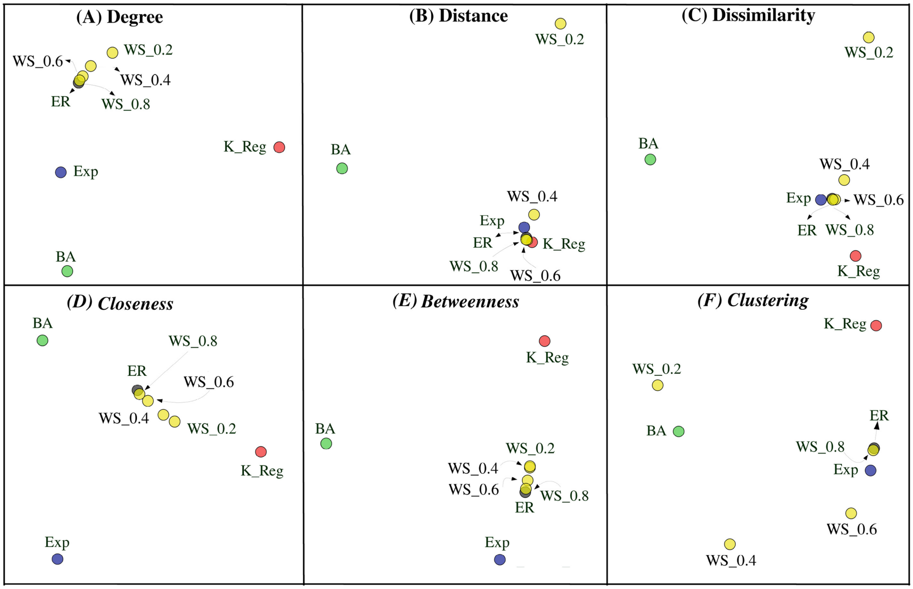

Fig. 1. Multidimensional scaling map performed over the outcomes of (A) $D^g_{\mathbf{P}_{deg}}$, (B) $D^g_{\mathbf{P}_\delta}$, (C) D, (D) $D^g_{\mathbf{P}_c}$, (E) $D^g_{\mathbf{P}_{bet}}$ and (F) $D^g_{\mathbf{P}_C}$ between all pairs of network null models: BA, EXP, K-regular and WS for different rewiring probability values ($WS\_0.2$, $WS\_0.4$, $WS\_0.6$ and $WS\_0.8$ consider the rewiring probability given by 0.2, 0.4, 0.6 and 0.8, respectively).

D values. The Infectious Sociopatterns network consists the face-to-face behavior of people during the exhibition INFECTIOUS: STAY AWAY in 2009 at the Science Gallery in Dublin. Nodes represent exhibition visitors; edges represent face-to-face contacts that were active for at least 20 seconds. The network has the data from the day with the highest number of interactions and is consider undirected and unweighted [23,24]. Figure 2A shows the outcomes. It is interesting to see that the betweenness and distance distributions share the same 3 critical elements. The dissimilarity function, on the other way, shares only two elements with the betweenness distribution sharing the third element with the clustering coefficient distribution. Figure 2B shows the degraded network after the removal of the critical elements found in A. When comparing the original and the degraded network, the last possesses a larger diameter (11), average path length (4.213771) and a small global clustering coefficient (0.436811).

The critical element detection problem is proven to be NP-hard in the general case for nodes and/or edges and, thus, the real case problems usually need heuristic approaches. The most common in the literature [25] is the strategy given by attacking the most central nodes (targeted attack[2]). Table 2 compares the values obtained by using 4 strategies of targeted attacks: higher degree,

[2] The nodes fail in decreasing order of centrality.

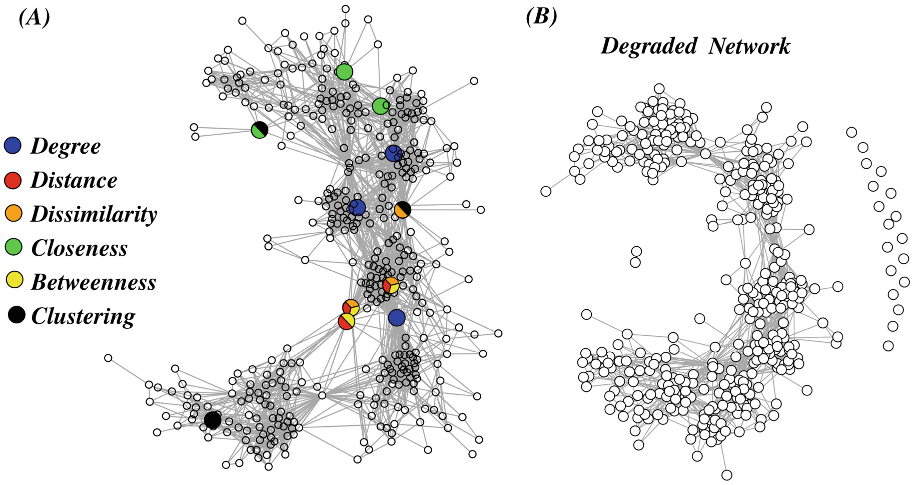

Fig. 2. (A) Critical 3 nodes in the Infectious Sociopatterns network for the degree ($D^g_{\mathbf{P}_{deg}}$), distance ($D^g_{\mathbf{P}_\delta}$), betweenness ($D^g_{\mathbf{P}_{bet}}$), closeness ($D^g_{\mathbf{P}_c}$), clustering ($D^g_{\mathbf{P}_C}$) and dissimilarity (D). (B) the degraded network obtained by the disconnecting the critical nodes.

closeness, betweenness, and clustering coefficient and the strategy of selecting the best combination of elements, we call it *Best* and it is computed by a brute force algorithm.

None of the above-mentioned targeted attack strategies achieved the network degradation given by the *Best* strategy, indicating that only one centrality measure is not enough as strategy to efficiently destroy the network.

Table 2. Comparing D^g and D values between targeted attacks (degree, closeness, clustering and betweeness) and the best strategy given by the critical node detection problem solution.

	$D^g_{\mathbf{P}_{deg}}$	$D^g_{\mathbf{P}_\delta}$	D	$D^g_{\mathbf{P}_c}$	$D^g_{\mathbf{P}_{bet}}$	$D^g_{\mathbf{P}_C}$
Degree	0.1468	0.1290	0.0745	0.2413	0.0860	0.0293
Closeness	0.1228	0.1790	0.0968	0.2471	0.0952	0.0333
Betweenness	0.1204	0.1666	0.0932	0.2462	0.1040	0.0333
Clustering	0.0638	0.0858	0.0509	0.2295	0.0115	0.0285
Best	0.1867	0.2288	0.12563	0.3811	0.1291	0.0456

Network failures may not occur all at once, but, at different time instances. Two sequences of failures may result in the same degraded network, even though, one may have caused a bigger topological destruction at the beginning of the attack. Therefore, the critical element detection problem fails in capturing this time-dependence of the failures.

In order to capture this time dependence of the failure process, following [9], a sequence of failures is defined as a sequence of time-indexed networks (G_t) where $G_0 = 0$ and G'_t is a subgraph of G_t for all $t' > t$ (as time increases, the network became more degraded).

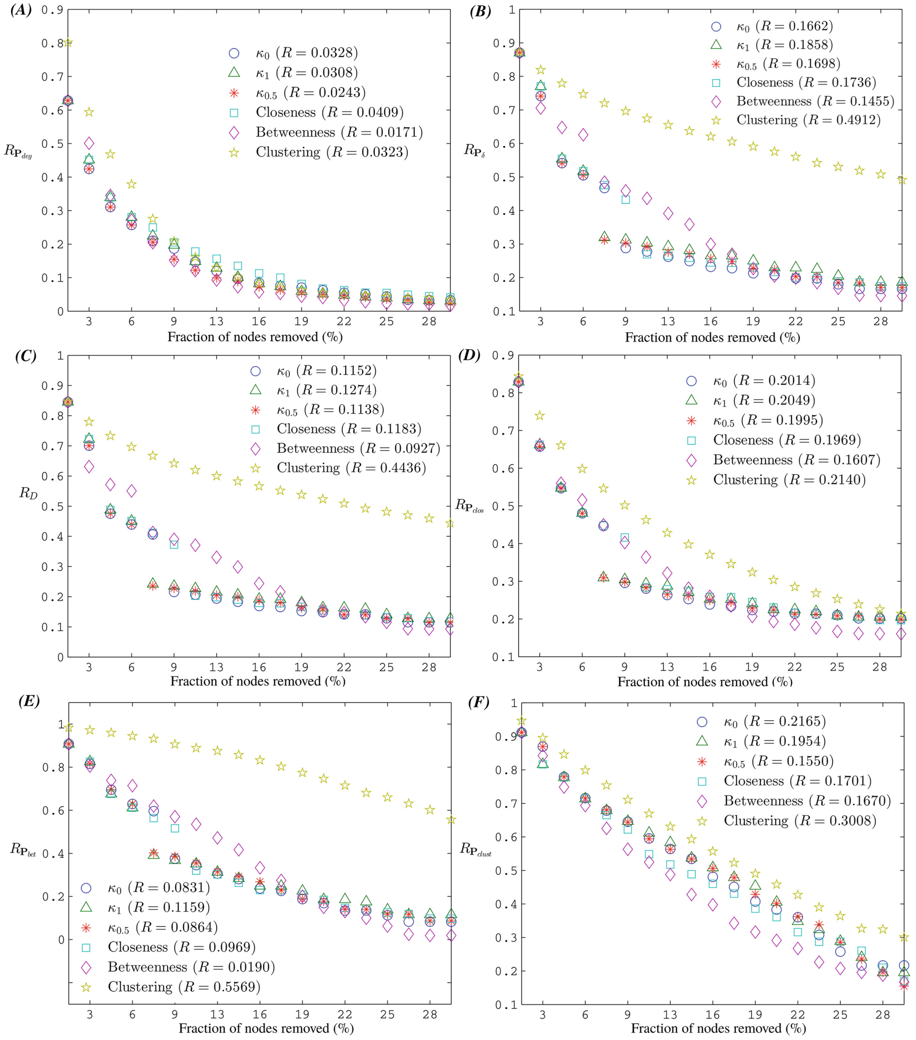

Fig. 3. Targeted attacks on the Train Bombing network. (A) $R_{\mathbf{P}_{deg}}$, (B) $R_{\mathbf{P}_\delta}$, (C) $R_{\mathbf{P}_D}$, (D) $R_{\mathbf{P}_c}$, (E) $R_{\mathbf{P}_{bet}}$ and (F) $R_{\mathbf{P}_C}$.

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

$$R_P(G|(G_t)_{t\in\{1,\,2,\ldots,n\}}) = \prod_{t=1}^{n} R(G_{t-1}|G_t), \tag{5}$$

being $R(G_t|(G_{t-1})_{t\in\{1,\,2,\ldots,n\}}) = 1 - D(G_t, G_{t-1})$.

This 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 cumulatively during a sequence of failures.

Here, we analyze the robustness of the Train bombing network under targeted attacks. This undirected and weighted network contains contacts between suspected terrorists involved in the train bombing in Madrid on March 11, 2004, as reconstructed from newspapers. A node represents a terrorist and an edge between two terrorists shows that there was a contact between the two terrorists. The edge's weight denotes how "strong" a connection was. This includes friendship and co-participation in training camps or previous attacks [23,26,27].

The experiment consists in attacking, at each time step, one node of the Train bombing network by a decreasing centrality value until the disconnection of 30 % of the nodes. Figure 3 shows the outcomes considering the robustness measure computed using $D^g_{\mathbf{P}_{deg}}$, $D^g_{\mathbf{P}_\delta}$, $D^g_{\mathbf{P}_{bet}}$, $D^g_{\mathbf{P}_c}$, $D^g_{\mathbf{P}_C}$ and D values. The targeted attacks are performed in decreasing order of: degree (κ^0), weight (κ^1), degree and weight with importance ($\kappa^{0.5}$), closeness, betweenness and clustering coefficient. In most cases, targeting the nodes with the highest betweenness centrality value generates the highest degradation in most of the analyzed measures. The only exemption is for $R_{\mathbf{P}_C}$, where the best strategy is given by attacking nodes considering $\kappa^{0.5}$ values.

Table 3 also shows that the best strategy after the degradation of 30 % of the network is not necessary the best when considering 20 % or 10 %. For example, in the case of $\mathbf{P}_{bet}$, the best strategy is considering the nodes' weight when 10 % of the nodes are removed, the degree attack for 20 % and, the betweenness centrality strategy for 30 %.

Table 3. Best targeted attack strategy for the Train Bombing network.

	Fraction of nodes removed		
	10 %	20 %	30 %
$R_{\mathbf{P}_{deg}}$	Betweenness	Betweenness	Betweenness
$R_{\mathbf{P}_\delta}$	Degree	Degree	Betweenness
R_D	Degree	Degree	Betweenness
$R_{\mathbf{P}_c}$	Degree	Betweenness	Betweenness
$R_{\mathbf{P}_{bet}}$	Weight	Degree	Betweenness
$R_{\mathbf{P}_C}$	Betweenness	Betweenness	$\kappa^{0.5}$
	Best attack strategy		

4 Concluding Remarks

In this work, we review a methodology to quantify graph dissimilarities based on Information Theory quantifiers that possess important properties. One of them is the flexibility of choosing the network measurement depending on the purpose of the analysis or application.

Acknowledgments. Research is partially by supported by the Laboratory of Algorithms and Technologies for Network Analysis, National Research University Higher School of Economics, CNPq and FAPEMIG, Brazil.

References

1. Bunke, H.: Recent developments in graph matching. In: Proceedings of the 15th International Conference on Pattern Recognition, vol. 2 (2000). http://dx.doi.org/10.1109/ICPR.2000.906030
2. Dehmer, M., Emmert-Streib, F., Kilian, J.: A similarity measure for graphs with low computational complexity. Appl. Math. Comput. **182**(1), 447–459 (2006)
3. Rodrigues, L., Travieso, G., Boas, P.R.V.: Characterization of complex networks: a survey of measurements. Adv. Phys. **56**(1), 167–242 (2006)
4. Schaeffer, S.E.: Survey: graph clustering. Comput. Sci. Rev. **1**(1), 27–64 (2007)
5. Bai, L., Hancock, E.R.: Graph kernels from the Jensen-Shannon divergence. J. Math. Imaging Vis. **47**(1–2), 60–69 (2013)
6. Schieber, T.A., Ravetti, M.G.: Simulating the dynamics of scale-free networks via optimization. PLoS ONE **8**(12), e80783 (2013)
7. Babai, L.: Graph isomorphism in quasipolynomial time. Arxiv, January 2016. http://arxiv.org/abs/1512.03547
8. Carpi, L.C., Saco, P.M., Rosso, O.A., Ravetti, M.G.: Structural evolution of the tropical pacific climate network. Eur. Phys. J. B **85**(11), 1–7 (2012). http://dx.doi.org/10.1140/epjb/e2012-30413-7
9. Schieber, T.A., Carpi, L., Frery, A.C., Rosso, O.A., Pardalos, P.M., Ravetti, M.: Information theory perspective on network robustness. Phys. Lett. A **380**(3), 359–364 (2016)
10. Schieber, T.A., Carpi, L., Ravetti, M., Pardalos, P.M., Massoler, C., Diaz Guilera, A.: A size independent network difference measure based on information theory quantifiers (2016, Unpublished)
11. Carpi, L.C., Rosso, O.A., Saco, P.M., Ravetti, M.: Analyzing complex networks evolution through information theory quantifiers. Phys. Lett. A **375**(4), 801–804 (2011). http://www.sciencedirect.com/science/article/pii/S037596011001577X
12. Newman, M.E.J., Strogatz, S.H., Watts, D.J.: Random graphs with arbitrary degree distributions and their applications. Phys. Rev. E **64**, 026118 (2001)
13. Deza, M.M., Deza, E.: Encyclopedia of Distances, p. 590. Springer, Heidelberg (2009)
14. Lewis, T.G.: Network Science: Theory and Applications. Wiley Publishing, Hoboken (2009)
15. Erdös, P., Rényi, A.: On random graphs. Publ. Math. **6**(290), 290–297 (1959)
16. Albert, R., Barabási, A.: Statistical mechanics of complex networks. Rev. Mod. Phys. **74**, 47–97 (2002). http://arxiv.org/abs/cond-mat/0106096

17. Frank, O., Strauss, D.: Markov graphs. J. Am. Stat. Assoc. **81**(395), 832–842 (1986)
18. Watts, D.J., Strogatz, S.H.: Collective dynamics of small-world networks. Nature **393**(1), 440–442 (1998)
19. Cox, T.F., Cox, T.F.: Multidimensional Scaling, 2nd edn. Chapman and Hall/CRC, Boca Raton (2000). http://www.amazon.com/Multidimensional-Scaling-Second-Trevor-Cox/dp/1584880945
20. Arulselvan, A., Commander, C.W., Elefteriadou, L., Pardalos, P.M.: Detecting critical nodes in sparse graphs. Comput. Oper. Res. **36**(7), 2193–2200 (2009). http://dx.doi.org/10.1016/j.cor.2008.08.016
21. Dinh, T.N., Xuan, Y., Thai, M.T., Pardalos, P.M., Znati, T.: On new approaches of assessing network vulnerability: hardness and approximation. IEEE/ACM Trans. Netw. **20**(2), 609–619 (2012)
22. Walteros, J.L., Pardalos, P.M.: A decomposition approach for solving critical clique detection problems. In: Klasing, R. (ed.) SEA 2012. LNCS, vol. 7276, pp. 393–404. Springer, Heidelberg (2012)
23. Kunegis, J.: KONECT - the Koblenz network collection. In: Proceedings of International Web Observatory Workshop (2013)
24. Isella, L., Stehlé, J., Barrat, A., Cattuto, C., Pinton, J.F., den Broeck, W.V.: What's in a crowd? analysis of face-to-face behavioral networks. J. Theor. Biol. **271**(1), 166–180 (2011)
25. Iyer, S., Killingback, T., Sundaram, B., Wang, Z.: Attack robustness and centrality of complex networks. PLoS ONE **8**(4), e59613 (2013)
26. Train bombing network dataset - KONECT, January 2016
27. Hayes, B.: Connecting the dots. Can the tools of graph theory and social-network studies unravel the next big plot? Am. Sci. **94**(5), 400–404 (2006)
28. Lin, J.: Divergence measures based on the Shannon entropy. IEEE Trans. Inf. Theory **37**(1), 145–151 (1991)
29. Opsahl, T., Agneessens, F., Skvoretz, J.: Node centrality in weighted networks: generalizing degree and shortest paths. Soc. Netw. **32**(3), 245–251 (2010)