Detecting Global Efficiency in Complex Networks Using Local Measurements and Neighborhood Information

Cinara Guellner Ghedini[1], **Carlos H. C. Ribeiro**[1]

[1]Technological Institute of Aeronautics
Computer Science Division
São José dos Campos – SP – Brazil

cinara@ita.br, carlos@comp.ita.br

Abstract. *Complex networks model real networks found in a wide range of domains from biological and social to technological environments. Their nature inherently dynamical emerges from ubiquitous and autonomic characteristics. Thus, assessing their behavior when dealing with failure and attacks is relevant to define mechanisms of control and recovery. However, practical aspects makes global metrics not measurable. Moreover, standard local metrics are not suitable to deal with global properties. In this paper, we present a new local metric to detect global changes in complex networks using local neighborhood information. The results show that the proposed metric accomplished satisfactory results for scale-free, small-world and random networks.*

Resumo. *Modelos de redes complexas são encontradas em uma ampla gama de aplicações, desde redes em ambientes biológicos e sociais até ambientes tecnológicos. A sua natureza intrinsecamente dinâmica é resultado de características de autonomia e ubiquidade. Desta forma, avaliar o seu comportamento face a falhas e ataques é relevante para definir mecanismos de controle e recuperação. No entanto, aspectos práticos tornam métricas globais não mensuráveis. Além disso, as métricas locais não são adequadas para lidar com propriedades globais. Neste trabalho, apresentamos uma nova métrica local para detectar mudanças globais em redes complexas considerando o conceito de informação da vizinhança. Os resultados mostram que a métrica proposta obteve resultados satisfatórios para redes livres-de-escala, mundo-pequeno e aleatória.*

1. Introduction

We are surrounded by environments and systems which can be classified or modeled as complex networks. They may be found in nature, for example in ecological food webs [Almaas and Barabasi 2004], *Escherichia coli* [Wuchty 2003], neurons in *Caenorhabditis Elegans* worms [Watts 2003]. Complex networks may also be models for social [Dodds et al. 2003] [Holl and H 2003] and technological networks, such as P2P, overlay, sensor, and communication networks [Broder et al. 2000] [Faloutsos et al. 1999] [Lua et al. 2005] [Biskupski et al. 2007]. These networks are formed and evolve in an adhoc manner. They are naturally self-organized and the interactions among their elements take into account only local information which is a cause for many systemic properties to emerge.

As complex networks are present in several domains under environments that are usually large, complex, highly dynamic and heterogeneous many studies, in different contexts, have been developed to understand how they are naturally organized, how they are shaped and how they evolve [Watts 2003] [Albert and Barabasi 2002] [Vega-Redondo 2007] [Newman 2003]. The main shared statement is that although complex networks formations are targeted to achieve specific objectives and functionalities, all of them show similar organizational principles. The understanding of these mechanisms supports the design of computational models and metrics, which can further foster a better understanding about their dynamical aspects and evolution process, as well as their characteristics.

Regarding to structural properties of the complex networks both global and local characteristic are relevant. Global assessment is often performed by computing the average of the shortest distance between any two nodes in a network, namely characteristic path length (L) [Newman 2003]:

$$L = \frac{1}{\frac{1}{2}n(n+1)} \sum_{i \leq j} d_{ij}, \tag{1}$$

where d_{ij} is the geodesic distance from vertex i to vertex j. In contrast, local analysis provide a mechanism to quantify the existence of tightly linked subgraphs, that is, it measures the cliquishness of a typical neighborhood and then, can expresses the configuration of the cluster in which a node is taking part. Consider a node i which has k_i connections to other nodes and l_i edges between these k_i nodes. If the nearest neighborhood of i was part of a clique, there would be $k_i(k_i-1)/2$ edges between them [Albert and Barabasi 2002]:

$$C_i = \frac{2l_i}{k_i(k_i-1)} \tag{2}$$

The average over all network elements gives the clustering coefficient (C) and expresses the cohesion of its elements. Another crucial statistic is the degree distribution, that describes the pattern which originates from the formation of links among nodes in the network and can determine some particular characteristics in the network dynamic.

The evaluation of real networks based on these properties concern that most pairs of nodes seem to be connected by short path lengths, being very efficient for global communication and information spreading, a phenomenon known as small-world. It is important to note that Erdős and Rényi [Erdős and Rényi 1960] have demonstrated the same characteristic for random networks. But, on the other hand they exhibit a much higher cluster coefficient than random graphs and a degree distribution which deviates from the typical distribution of random graph (Poisson). These characteristics have established two major classes of network topologies: small-world (SW) and scale-free (SF). Figure 1 show examples of SW, SF and random networks to point out their main topological differences.

Scale-free topologies are present in many real networks [Newman 2003] [Almaas and Barabasi 2004] and exhibit the same SW pattern for L, as well as a high cluster coefficient C, but the distribution of node degrees follows a power-law $P(k) \sim k^{-\gamma}$,

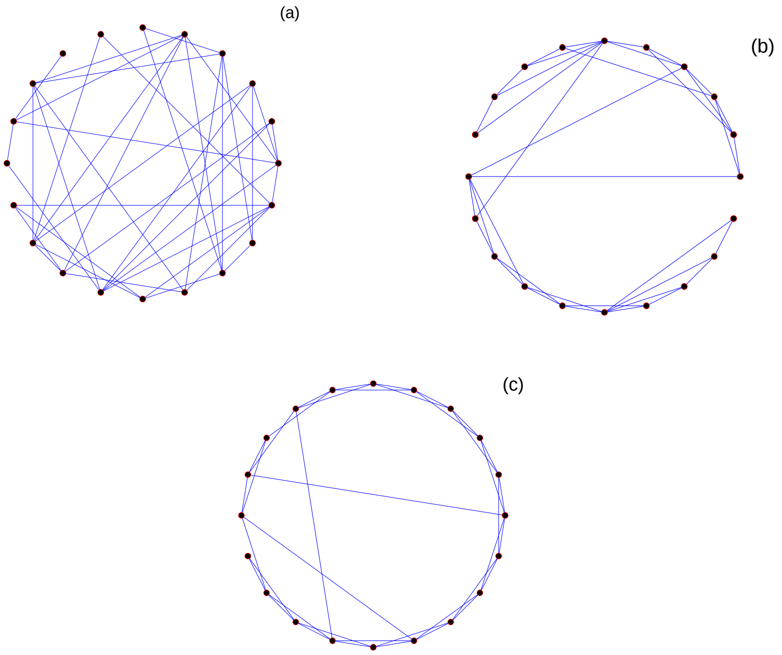

Figure 1. Example of network topologies random (a), SF (b) and SW (c) with $N = 20$, $\langle k \rangle = 4$.

where $P(k)$ gives the probability that a randomly selected node has exactly k edges (Figure 2 (b)). Consequently, it means that a few nodes have many connections, in contrast with random networks (Figure 2 (a)) that exhibit degree distribution approaching a Poisson distribution (line). As can be seen in Figure 2 (c) the degree distribution presented in small-world networks model proposed by Watts [Watts 2003] [Watts and Strogatz 1998] is almost uniform.

Despite these well-defined properties, realistic models of networks such as the so called *ad-hoc* networks do not rely on any fixed or predefined infrastructure, on the contrary their intrinsic features such as mobility, with nodes leaving or joining the network, often cause frequent changes in their topologies. In addition, these networks are subject to technical problems (failures), or even intentional attacks.

Thus, nodes rely on each other to keep the network connected, and one of the main goals for network applications is to ensure that the network services will be available and working efficiently regardless of frequent topological changes. In fact, several researchers have studied the impact of failures and attacks into the network structure [Dall'Asta et al. 2006] [Crucitti et al. 2004] [Albert et al. 2000], considering it an important issue to define mechanisms to support efficiency improvement or system stability in such adverse conditions.

A rather studied aspect is the behavior of the topology when nodes are dropped, in particular what fraction of nodes must be removed before the network splits or breaks into isolated clusters, and what are the properties of such groups. It is known that complex networks are very robust against failures, but very vulnerable against attacks [Latora and Marchiori 2002] [Albert et al. 2000]. Thus, to evaluate the network effi-

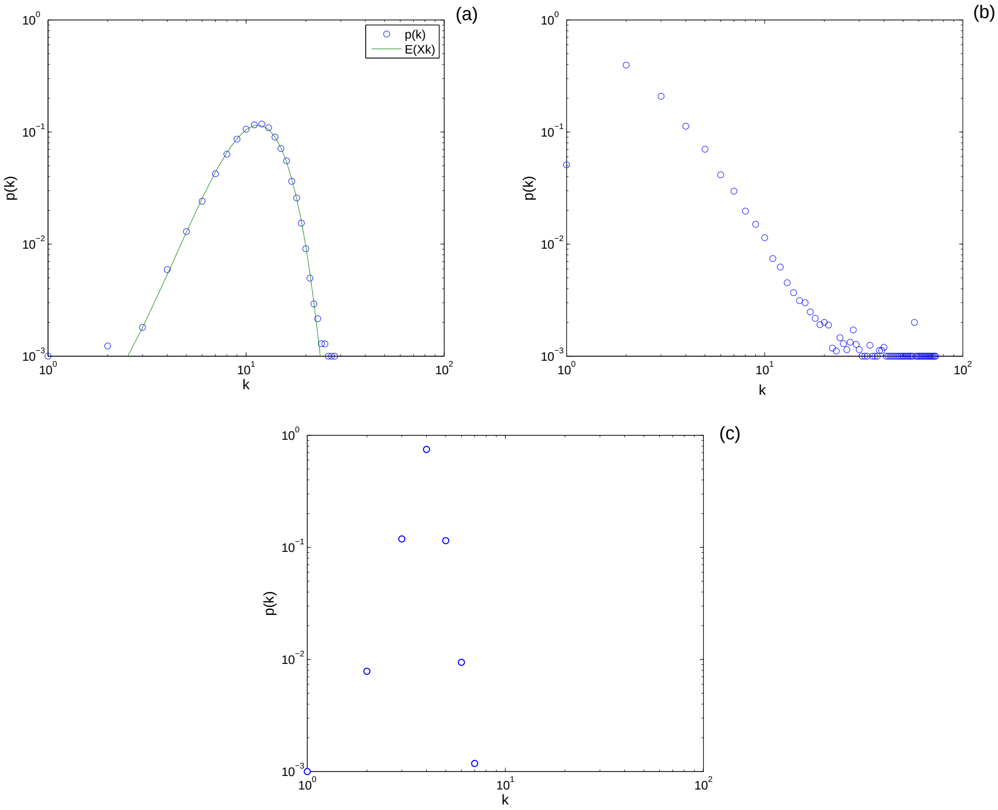

Figure 2. The degree distribution averaged over 30 random (a), SF (b) and SW (c) graphs with $n = 1000$ and $\langle k \rangle$ = 6

ciency, an appropriate measure must be used.

At first glance, the intuitive measure to be used is the characteristic path length L, because it can express how far apart the nodes are from each other, in other words, how efficient the network is with respect to information dissemination through its elements. When the focus of analysis (as in this paper) is the network behavior against attacks and failures, the network has always a chance of becoming disconnected. In such case, L is not appropriate because it considers the average of the shortest path lengths over all nodes in the network, and some of them would not be reachable. This means that adjustments are necessary to perform efficiency computation, especially when considering that isolated clusters can continue to perform tasks and exchange information. Another issue is that both L and C are well-defined only for topological/unweighted graphs, and real networks usually exhibit information about node relationships.

Lattora *et al* [Latora and Marchiori 2002] [Crucitti and Latora 2001] introduced the efficiency E which measures how efficiently the nodes exchange information in a local or global scope, independently of whether the network is weighted or topological. It can also be applied to disconnected graphs. Consider a graph G where d_{ij} is the smallest sum of the physical distances throughout all the possible paths between nodes i and j. The efficiency E_{ij} is inversely proportional to the shortest distance: $E_{ij} = \frac{1}{d_{ij}}$. If there is no path between them, the distance d_{ij} is $+\infty$, and therefore $E_{ij} = 0$. Thus, the global efficiency of a graph G can be defined as:

$$
\begin{gathered}
E_{glob}(G) = \frac{\sum\limits_{i \neq \mathrm{j} \in G} E_{ij}}{N(N-1)} = \\
\frac{1}{N(N-1)} \sum_{i \neq \mathrm{j} \in G} \frac{1}{d_{ij}}
\end{gathered}
\tag{3}
$$

The efficiency E_{glob} defined in 3 ranges in $[0, \infty]$. To normalize it, consider the ideal case G_{ideal} where all the possible $N(N-1)/2$ edges are in the graph. This is the case when E_{glob} assumes its maximum value. Then, the normalized efficiency is $\frac{E_{glob}(G)}{E_{glob}(G_{ideal})}$. On the other hand, the local efficiency 4 represents the average efficiency of local subgraphs. Note that it is equivalent to the cluster coefficient:

$$E_{loc}(G) = \frac{1}{N} \sum_{i \neq \mathrm{j} \in G} E(G_i) \quad where \tag{4}$$

$$E(G_i) = \frac{1}{k_i(k_i-1)} \sum_{l \neq m \in G_i} \frac{1}{d_{lj}} \tag{5}$$

and G_i is the subgraph containing all nodes directly connected to i (k_i is its degree).

These concepts of efficiency are appropriate to the context of failures and attacks in both topological and weighted networks, and are adopted in this work to evaluate the network efficiency. However, in general the sheer size of the networks, the operational constraints (e.g., node mobility) and the need to make decisions quickly make the computation of global measurements unfeasible, and local metrics are required for network evaluation.

Implicitly, the requirement for locality allows each node to be autonomous and able to make decisions about its configuration, considering local network behavior in order to achieve global objectives, or even to perform cooperative tasks when necessary. This implies that the nodes of the network must be ”aware” of the network status (albeit in a localized manner), as well as have mechanisms to recognize and react to changes.

The main issue is then how a node or a cluster can be ”aware” to predict and take advantages of the global status of the network. For instance, when the global efficiency is affected, node or group mechanisms may be triggered to reduce its impact. Also, when efficiency increases, the nodes can take advantages of this to perform tasks that require high reliability or performance. Mechanisms to learn the pattern of network operations in various scenarios are also useful.

Thus, the main goal here is to evaluate the measurements that are potentially useful for capturing global network changes considering only local information. As the network topology is not previously known, three of the major classes of topologies found in real complex networks (scale-free, small-world and random networks) were considered for the analysis of both purely topological and weighted relations.

The rest of the paper is organized as follows. In section 2 the behavior of complex networks against attacks and failures is analyzed. Section 3 discusses the use of classical measures to estimate global efficiency. In Section 4, a new local metric to capture the global state of the network using local neighborhood information is presented, together with the obtained results. Finally, Section 5 presents the main conclusions and some directions for future work.

2. Attacks and Failures in Complex Networks

Evaluating the behavior against failures and attacks requires a methodological approach that is briefly sketched in what follows. As noted above, the network topology is assumed not known in advance and its behavior is therefore unpredictable. However, several studies [Crucitti et al. 2004] [Albert and Barabasi 2002] [Newman 2003] pointed out that it is possible to find similar patterns in its formation, such as small characteristic path length and high cluster coefficient. This means that real networks are very efficient to disseminate information and the structure of their groups is very cohesive. Then, the likelihood of groups of nodes remain connected when a node is lost is large.

These two important properties (stable patterns and group cohesiveness) are present in SF and SW networks. As far as the degree distribution is concerned, most real networks show power-law distribution, but different patterns can be found. As our intent is to assess both network behavior and metrics performance in unknown topologies, we used three different topological models for the sake of generalizing the results: SF, SW and random networks. All show a small characteristic path length, differing however in degree distribution and cluster coefficient. SF networks have a power-law tail degree distribution, random networks hold Poisson degree distributions, and in SW topologies all nodes have almost the same number of connections. Regarding the clustering coefficient, SF and SW exhibit high C, but random networks exhibit poor clustering.

In these sense, a specific predefined procedure is adopted to build such networks. Each one has a particular method and a set of parameters that are not specified here because this is not the focus. However, we make sure that all generated networks show the same number of nodes and an equivalent number of edges, as well as the expected features. Small-world networks were generated by Watts and Strogatz β-model (WS) [Watts 2003]. Considering scale-free as networks which present small-world properties plus preferential attachment, the model created by Klemm-Euguluz (KE) [Klemm and Eguíluz 2002] was used. The model of random graphs used to generate random networks is $G(n, p)$ where n is the number of nodes and p is the independent probability $0 < p < 1$ that exists an edge between every possible pairs of nodes in G [Erdős and Rényi 1959].

In order to evaluate the impact of failures and attacks global (3) and local efficiency (4) were used to compute network efficiency. To change a network global status, two distinct strategies were used: one to decrease and another to increase network efficiency. For the first one, three strategies were considered, one for failures and two for attacks. To mimick failures, the nodes were chosen randomly with uniform distribution. To perform attacks, the better positioned nodes must be dropped. There are several centrality measures which allow ranking nodes ordered by its importance, each of them underlying some useful characteristic for network analysis. In the cases we consider here, the probability that a node can disconnect the network or be harmful to its performance is relevant to evaluate network robustness, survival and resilience. For this, two centrality metrics were adopted. The first one is Degree Centrality (DC), defined in (6) and where the central nodes are simply the ones that hold more ties to other nodes in the network [Wasserman et al. 1994]:

$$C_D(n_i) = d(n_i) \tag{6}$$

where $d(n_i)$ is the degree of node i.

The second centrality metric is Betweeness Centrality (BC) — defined in (7) — which establishes as high scoring the nodes which play an important role in the interactions between other nodes, that is, the node which is contained in most of the shortest paths between all the pairs of nodes in the network shows the largest score. The BC of node i is computed considering the number of geodesic paths linking all pairs of nodes (not including i) presented in the network $(g_{jk}(n_i))$ in which i is included. As, in such case, all geodesics are equally likely to be chosen, the probability of a link is calculated by $g_{jk}(n_i)/g_{jk}$, where g_{jk} is the number of geodesics paths between actors j and k. Then, the BC of node i is the sum of these estimated probability over all pairs of nodes not including i [Wasserman et al. 1994]:

$$C_B(n_i) = \sum_{j<k}(g_{jk}(n_i)/g_{jk}) \tag{7}$$

The evaluation of variability in the network efficiency must take into account the efficiency increasing. in order to achieve this some experiments using BC and DC to rank the network nodes and creating links between nodes under different combinations were performed. For instance, we created links between nodes with lowest DC and highest DC. Note that this strategy uses global information, but now the point is only to evaluate metrics that can capture the increase of the network efficiency. It was observed that creating links between nodes with high degree of centrality and low betweeness centrality the global network efficiency increased, because nodes which were (likely) at the network boundaries were connected to others with high degree, adding shortcuts to the network.

Figure 3 shows the initial network scores for local and global efficiency, as well as their oscillations when exposed to continuous attacks and failures. All results were averaged over 30 networks containing 100 nodes and 300 edges ($M \sim 300$). In each iteration a node was chosen to be dropped from the network in accordance to the priorities previously defined: highest DC, highest BC and randomly. Then the global and local efficiency were computed. As it may be noticed, initial global efficiency is fairly similar for all networks, but local efficiency is higher in SW and SF networks. However, they react differently against failures and attacks.

In general, the global efficiency and local efficiency were lightly affected by failures in every network topology. In fact, global and local efficiency are hardly affected at all in random networks. Indeed, this happens because the choice for node removal is random: sometimes nodes at the border, and sometimes well-centralized nodes are chosen for removal. In contrast, the defined strategies of attack (BC or DC) affect SW and SF network efficiencies in different ways. This is understandable, because a) the strategy to compute them is not the same and b) the degree distribution in both topologies is different.

SW topologies generated by β-model inherit characteristics of regular lattice topologies, where the degree distribution is uniform and just a few nodes play the role of connecting to other nodes far apart in the network, thus supporting shortcuts in node communication. This implies that node degree centrality in these networks is uniform, but node BC is more heterogeneous (nodes playing the role of shortcuts have higher BC). Moreover, those nodes with high BC have links with other nodes which are not in the same neighborhood, then decreasing local efficiency. Regarding to attacks when they are directed to high BC nodes, the network looses an important channel of communications, affecting the dissemination of information, even though the neighborhood becomes more robust locally. Also, the network can potentially be split.

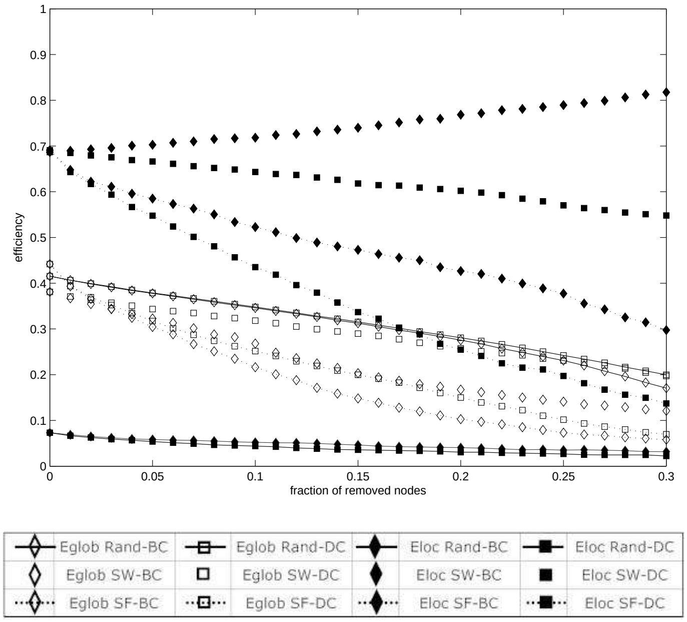

Figure 3. Efficiency - Failures/Attacks.

Attacks targeted to high DC nodes are less harmful because the degree distribution is uniform, thus the probability to disconnect shortcuts is also uniform. When these nodes are chosen the efficiency decreases faster, but on the contrary the global efficiency may be less affected. Global efficiency is more affected at the beginning, but along time the values approach those of random graphs. Notice that at this point, in spite of attacks to nodes with high degree, the global efficiency is not much affected as SF networks, the local efficiency remains high and the curve slop drops slightly.

The main reason for this might be a larger number of nodes playing the role of shortcuts. The SF networks are more affected by attacks than SW ones. The node degree in SF networks is more heterogeneous, as the probability that any node links to another node is proportional to its degree. Global efficiency decreases a little faster for BC than DC attacks. On the contrary, local efficiency decreases faster for DC attacks. This happens because nodes with high degree are linking nodes with low degree, making local connectivity more affected.

3. Evaluation of Classical Metrics

To evaluate the adequacy of the metrics we considered, for each network change, the global efficiency $E_{glob}(G)$ (3) and the score obtained by each node Sc_i according to the target metric. Each pair $(E_{glob}(G), Sc_i)$ was classified as:

- True Positive (TP): if both $E_{glob}(G)$ and Sc_i increase, that is, the node i might detect the global efficiency increase.
- True Negative (TN): if both $E_{glob}(G)$ and Sc_i decrease, that is, the node i might detect the global efficiency decrease.
- False Negative (FN): if $E_{glob}(G)$ increases and Sc_i decreases, but the node i detected that global efficiency decreases.
- False Positive (FP): if $E_{glob}(G)$ decreases and Sc_i increases, that is, the global efficiency decreases, but the node i detected that global efficiency increases.

After that, the number of TP ($\sharp TP$), TN ($\sharp TN$), FN ($\sharp FN$) and FP ($\sharp FP$) were counted. To evaluate metrics performance, normalized measures known as recall and precision were used. Recall means the percentage of matches of the recovered pairs $(eglob, Sc_i)$. Precision means the percentage of pairs $(eglob, Sc_i)$ which correctly detected the state of the network. Thus, it was possible to evaluate the metric precision and recall in both scenarios, when global efficiency increases (↑)(8) (9) and when it decreases (↓) (10) (11).

$$recall \uparrow = \frac{\sharp TP}{\sharp TP + \sharp FN} \tag{8}$$

$$precision \uparrow = \frac{\sharp TP}{\sharp TP + \sharp FP} \tag{9}$$

$$recall \downarrow = \frac{\sharp TN}{\sharp TN + \sharp FP} \tag{10}$$

$$precision \downarrow = \frac{\sharp TN}{\sharp TN + \sharp FN} \tag{11}$$

3.1. Centrality measures

The initial studies were focused on the evaluation of centrality metrics to detect global efficiency. The results (not shown here due to space limitations) were significant to understand the overall network behavior, but not to detect global efficiency variation. They revealed that usually node centrality and global efficiency change in different ways. For example, a node which shows high centrality and low degree probably connects two clusters which are apart. This node affects the centrality of other nodes in both clusters. If this node is dropped, nodes which were in the border of the cluster before will get better centralized, but the global efficiency is likely to be heavily affected.

3.2. Local efficiency

As expected, experiments have demonstrated the poor performance of the local efficiency measure (4) to detect global changes. Indeed, most of the nodes could not detect global changes, as might be inferred from the previous discussions and results (Section 2). The results show that recall and precision for detecting both decrease and increase of global efficiency were below 0.03. In fact, most of nodes (≈95%) did not alter significantly their values of local efficiency when global efficiency changed.

4. Using Neighborhood Information

The observed behavior of complex networks when exposed to attacks and failures and the results obtained with centrality and local efficiency measures make new approaches necessary. The analysis of the results previously reported, makes clear that is more important to capture the range of nodes that can be reached before and after network changes. The most intuitive and simple metric is the nodal degree, but considering a range greater than closer (direct) neighbors, comprising the neighbors at some distance greater than 1. Neighborhood connectivity r^{th} is defined as the number of nodes connected to node i at a shortest distance exactly equal to r [Vega-Redondo 2007]:

$$\begin{aligned} N_r^i = j \in N\{i\}\colon \exists K_0 \cdots, K_r \in N s.t \\ k_{s-1}k_s \in L(S = 1, 2\ldots, r), \\ k_0 = i, k_r = j \backslash \cup_s = 1^{r-1} N_i^s \end{aligned} \tag{12}$$

Once neighborhood connectivity of node i is computed using local information, the idea is to use it as a measure to detect global changes. At first it was considered a database containing networks with 100 nodes ($N = 100$), 300 edges ($M \sim 300$) and an average degree of 6 ($\langle k \rangle \sim 6$) for all considered topologies. Links have values in the range $0.01 - 1$ for weighted and 1 for topological networks. Taking into account these values the neighborhood connectivity r^{th} was set to 2 ($r = 2$). The results presented in Figures 4, 5 and 6 encompasses averages over 30 networks where 30 percent(30%) of the nodes were dropped.

The left half shows the results for global efficiency increase ($eglobal\ \uparrow$), and the right half for global efficiency decrease ($eglobal\ \downarrow$) when either attacks (BC, DC) or failures are present. However, in all evaluated scenarios (Figure 4), recall and precision were not significant for attacks, and despite the precision score for errors being between 0.49 and 0.39, the performance of the N_r^i measure was not compelling.

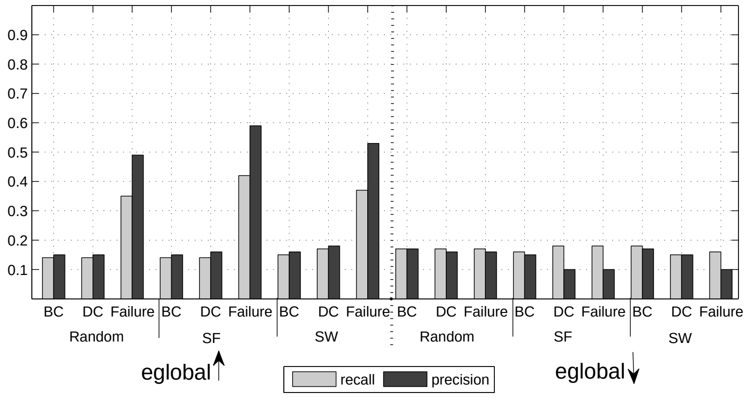

Figure 4. N_r^i recall and precision - weighted networks.

Intuitively, when network efficiency changes the amount of information flowing through the network also changes. Then the concept of neighborhood information can be extended from neighborhood connectivity to neighborhood efficiency with respect to the exchange information. In such case, the r^{th} neighborhood of node i ($\Gamma_r(i)$) is the subgraph S_r^i containing all vertices in the network at a shortest distance less or equal to r from i inclusive:

$$S_r^i = \Gamma_r(i) \tag{13}$$

In this direction, Stephenson and Zelen [Stephenson 1989] introduced a measure to compute the information centrality which considers the information contained in all possible combined paths between pairs of points. In addition, they developed a straightforward way to calculate this information for large networks. Furthermore the measure can be used for topological and weighted graphs without extra effort. The focus is not in centrality measures, but the computation of combined path information, which is integrated to computation of the information centrality, can be quite interesting to be applied in order to detect network changes.

Consider a network as a graph $G = (N, L)$ given by a set of nodes $N = 1, 2, \ldots, n$ and a set of links $L \subset N * N$, represented by using two $n \times n$ matrices. The first is the adjacency matrix, denoted by M, such that $m_{ij} = 1 \quad if(i, j) \in L$ and otherwise, 0. The second is the weight matrix denoted by W, such that w_{ij} is the weight associated with each link. The value of the information in the combined path from node i to j is given by [Stephenson 1989]:

$$I_{ij} = (C_{ii} + C_{jj} + 2C_{ij})^{-}1 \tag{14}$$

Where C is B^{-1} and B is a matrix $n \times n$ defined by:

$$B = \begin{cases} b_{ij} = 1 & \text{if there is no path between i and j} \\ b_{ij} = 0 & \text{if there is a path between i and j} \\ b_{ii} = 1 + \sum\limits_{j=1, j \neq i}^{n} (w_{ij} * m_{ij}) & \end{cases}$$

Therefore, the matrix I obtained from (14) contains the combined path between all the nodes in the graph. If the scale change from global to local, the concept of information flow may be applied to neighborhood information, in other words, to support the computation of information that is flowing in the neighborhood of a node. Using the previously concept of subgraph S_r^i to represent the neighborhood of node i, the neighborhood information of node i is the sum of the information flow in all possible paths:

$$N_{I_r^i} = \sum_{l} \sum_{j>l} I_{ij} \quad \text{wh} \quad \{l, j : l, j \in S_r^i\} \tag{15}$$

Figures 5 and 6 summarize the performance of $N_{I_r^i}$. The results show a high recall and a high precision in all network topologies when changes in network were caused by attacks. In these cases, all the scores were almost 1 for precision ($precision \downarrow\ and \uparrow \cong$ 1). This ensures the reliability of the obtained information. Even if not all nodes do capture the global efficiency, the information reliability regarding the detection of the efficiency

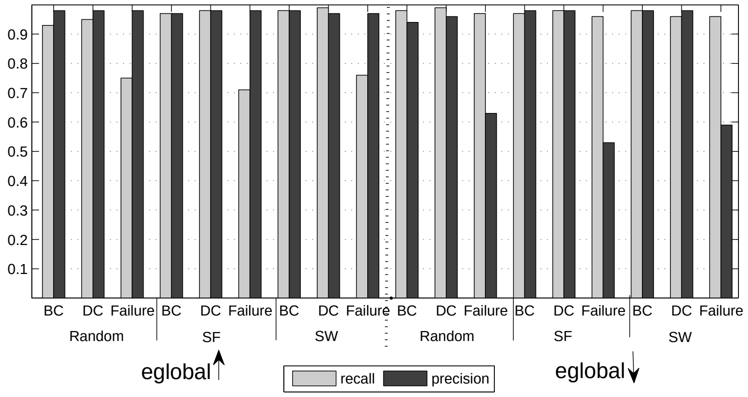

Figure 5. $N_{I_r^i}$ recall and precision - weighted networks.

variation might allow for the development of mechanisms to support network elements to take advantage of network status.

In case of changes driven by failure, when global efficiency increases the precision was high ($precision \downarrow \cong 1$), because of a path was created between nodes with high degree and low centrality, that is, the change could be better disseminated. On the other hand, the precision regarding detection of global efficiency is not as high ($precision \downarrow\approx 0.6$) for all topologies. This happens because failures may occur at the edge of the network. Even so, the obtained scores are significant since they were much better than the obtained from local efficiency (4), information based on degree (12) or global metrics. Notice that the results obtained with topological networks (Figure 6) were very similar to those obtained in weighted networks (Figure 5). This means that the neighborhood information is efficient to capture the topological changes, independent of the node connection intensity.

The results obtained by this measure are clearly dependent on the parameter r. However, despite the average characteristic path length (L) of the used networks being 2.3 (SF), 2.4 (SW) and 2.6 (Random), to detect network changes it would be necessary to compute the global efficiency over the entire network. Whether we consider that the size of subgraph S_r^i (neighborhood r of node i) is between 8 and 62 nodes for SF networks (an average of 30) or 13 and 28 (an average of 17) for SW, less effort is required to estimate global changes.

Figure 7 shows the results obtained from the average of five networks with $N = 1000$, $M \sim 3000$ and $\langle k \rangle \sim 6$, with average characteristic path length (L) 4 for random networks, 5 for SW networks and 6.74 for SF networks. As can be noted, despite the network size and L having increased and the value of parameter r remaining the same ($r = 2$) from previous experiments, the obtained results were also significant for all topologies. For example, in SF networks for both BC and DC attacks the precision is almost the same, while the recall decreases from 0.96 to 0.87 for BC and from 0.98 to 0.92 for DC. On the other hand, for failures precision decreases from 0.99 to 0.89, but recall increases 0.05. Concerning the size of subgraphs S_r^i, the of number of nodes decreased from the entire network to a range of 11 to 122 (and an average of 40) for SF and a range of 11 to 34 (and an average of 16) for SW, the relative radius of the subgraph decreased.

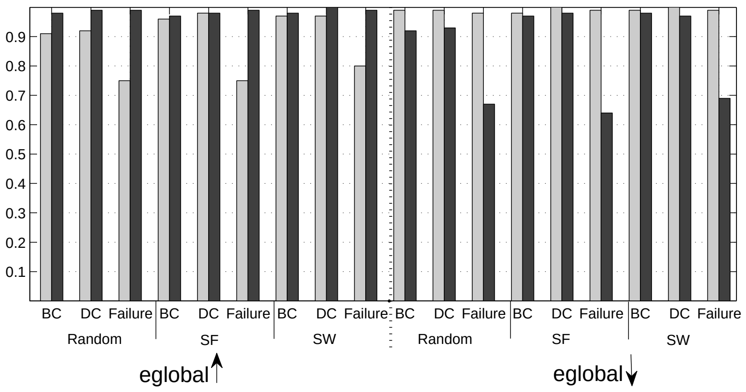

Figure 6. $N_{I_r^i}$ **recall and precision - topological networks.**

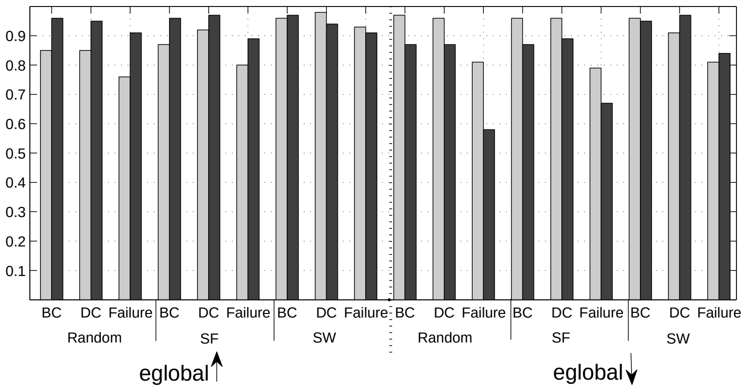

Figure 7. $N_{I_r^i}$ **recall and precision - topological networks.**

The results show that a node can be ”aware” of fluctuations in network performance. Changes in patterns may indicate problems in network dynamic, when the efficiency drops consecutively, or even opportunities to perform specific tasks. They also demonstrate that if most nodes could be ”aware” of changes, the computation might only be performed in some specific nodes. In this sense, preliminaries analysis shown that specific nodes are more sensitive to capture not only the tendency of the network efficiency, but also are able to estimate of how much was this change, more specifically ranges of changes. However, the approach emphasizes the use of local metrics which makes more complex the characterization of such nodes.

5. Conclusions

The obtained results from $N_{I_r^i}$ encourage good prospects. This metric showed better results when considering local efficiency, centrality measures and neighborhood connectivity. Similar results for topological and weighted networks showed that the neighborhood information captures not only the topological changes, but also the real efficiency of paths in the network. Furthermore, the neighborhood information precision is higher than what is obtained from other classical metrics, and the high number of nodes that can properly capture global changes allows the development of methods considering only a few nodes to perform this task. Thus, the metric can evolve or be combined to others in order to reach a robust method to detect global changes using local information.

New experiments must be done in order to evaluate real behavior for different network sizes. However, it is important to note that complex real networks are naturally organized in order to maintain a small characteristic path length, which usually grows at logarithmical rate, rather than exponential.

We are currently working on two main issues. The first one is investigating whether it is possible to classify nodes with respect to its accuracy in detecting global changes considering $N_{I_r^i}$, as well as the number of nodes that must be considered in order to obtain the best performance. The second is to combine $N_{I_r^i}$ and density measures to detect the probability that some node gets disconnected from the network, in other words, if the current network state can be harmful to network stability.

6. Acknowledgments

The authors thank CNPQ and FAPESP for the financial support to carry out the research reported herein.

References

Albert, R. and Barabasi, A. L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1).

Albert, R., Jeong, H., and Barabasi, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–382.

Almaas, E. and Barabasi, A. (2004). Power laws in biological networks. *q-bio*.

Biskupski, B., Dowling, J., and Sacha, J. (2007). Properties and mechanisms of self-organizing manet and p2p systems. *ACM Trans. Auton. Adapt. Syst.*, 2(1).

Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., Tomkins, A., and Wiener, J. (2000). Graph structure in the web. *Computer Networks*, 33(1):309–320.

Crucitti, P. and Latora, V. (2001). Efficient behavior of small-world networks. *Physical Review Letters*, 87(19):198701+.

Crucitti, P., Latora, V., Marchiori, M., and Rapisarda, A. (2004). Error and attack tolerance of complex networks. *Physica A: Statistical Mechanics and its Applications*, 340:Pages 388–394.

Dall'Asta, L., Barrat, A., Barthelemy, M., and Vespignani, A. (2006). Vulnerability of weighted networks. *Theory and Experiment*, 2006:04006.

Dodds, P. S., Muhamad, R., and Watts, D. J. (2003). An experimental study of search in global social networks. *Science*, 301(5634):827–829.

Erdős, P. and Rényi, A. (1959). On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297.

Erdős, P. and Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci*, 5:17–61.

Faloutsos, M., Faloutsos, P., and Faloutsos, C. (1999). On power-law relationships of the internet topology. In *In SIGCOMM*, pages 251–262.

Holl, J. and H, M. S. (2003). An assessment of preferential attachment as a mechanism for human sexual network formation.

Klemm, K. and Eguíluz, V. (2002). Growing scale-free networks with small world behavior. *Phys Rev E*, (65).

Latora, V. and Marchiori, M. (2002). Economic small-world behavior in weighted networks.

Lua, K., Crowcroft, J., Pias, M., Sharma, R., and Lim, S. (2005). A survey and comparison of peer-to-peer overlay network schemes. *Communications Surveys & Tutorials, IEEE*, pages 72–93.

Newman, M. E. J. (2003). The structure and function of complex networks.

Stephenson (1989). Rethinking centrality: Methods and applications. *Social Networks*, 11:1–37.

Vega-Redondo, F. (2007). *Complex Social Network*. Cambridge University Press.

Wasserman, S., Faust, K., and Iacobucci, D. (1994). *Social Network Analysis : Methods and Applications (Structural Analysis in the Social Sciences)*. Cambridge University Press.

Watts, D. J. (2003). *Small Worlds : The Dynamics of Networks between Order and Randomness (Princeton Studies in Complexity)*. Princeton University Press.

Watts, D. J. and Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature*, 393(6684):440–442.

Wuchty, S. (2003). Small worlds in rna structures. *Nucl. Acids Res.*, 31:1108–1117.