Evaluating links through spectral decomposition

Gonzalo Travieso and Luciano da Fontoura Costa

Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, SP, PO Box 369, 13560-970, Brazil
E-mail: gonzalo@ifsc.usp.br and luciano@ifsc.usp.br

Received 19 September 2011
Accepted 12 December 2011
Published 25 January 2012

Online at stacks.iop.org/JSTAT/2012/P01015
doi:10.1088/1742-5468/2012/01/P01015

Abstract. Spectral decomposition has rarely been used to investigate complex networks. In this work we apply this concept in order to define two kinds of link-directed attacks while quantifying their respective effects on the topology. Several other kinds of more traditional attacks are also adopted and compared. These attacks had substantially diverse effects, depending on each specific network (models and real-world structures). It is also shown that the spectrally based attacks have special effects in affecting the transitivity of the networks.

Keywords: random graphs, networks

ArXiv ePrint: 1109.4900
The theory of graphs (see e.g. [1]) and networks (see e.g. [2, 3]) represents one of the most multidisciplinary, integrated and applicable areas of theoretical mathematics and computing. Although its origin is often traced back to Euler’s solution of the Königsberg bridge problem, graphs have been around for much longer, at least since the first map was drawn on sand. Because graphs and networks can represent most discrete structures possibly underlying dynamical systems [3, 4], they are particularly useful for modeling a vast range of problems. The identification of structured connections in growing graphs, especially the existence of hubs and their fundamental importance (see e.g. [2]), helped to catalyze a surge of interest which had already been sparked by random graphs and small world network studies, giving rise to the new theory of complex networks.

A good proportion of the investigations on complex networks have focused on relatively simple properties such as the node degree (i.e. the number of connections established by a node), the transitivity (i.e. the degree of interconnectivity among the immediate neighbors of a node) and the shortest path length between two nodes. These measurements [5] are particularly important because they correspond to the distinguishing features of the main complex network models. For instance, small world networks are characterized by low mean shortest path length together with high transitivity, and scale-free networks exhibit power law degree distributions. However, as these measurements are not sufficient to provide a complete, invertible, representation of the complex network of interest, they will not be sufficient to directly express many important connectivity properties.

There are so many possible measurements of complex networks that it is useful to organize them into categories (see e.g. [5]). A particularly interesting and useful category of measurements are those called spectral (see e.g. [6]–[8]), in the sense of involving the eigenvalues of the adjacency matrices of the graphs analyzed. Spectral measurements are special kinds of graphs called geographical graphs, which are characterized by the fact that the links have well-defined positions in an embedding space.
approaches to graphs and networks are particularly important for many reasons, including the relationship between eigenvalues and the dynamics of the network, connectedness, cuts, modularity and cycles, among others. Such concepts and methods have progressively attracted attention from the complex networks community, to the extent that some of the best community finding algorithms in this area are now based on spectral methods (see e.g. [9]).

Spectral approaches for graphs have many relationships with theoretical and applied physics; they may consider the adjacency (see e.g. [6, 8]) or Laplacian matrices (see e.g. [7]) of graphs. In this work we concentrate attention on the former kinds of approaches. More specifically, because the spectrum of a graph does not provide a complete representation, we focus our attention on the possibility of using the eigenspaces of graphs [8] in order to derive more powerful features for characterizing the graph connectivity. Complete representations are important in graph studies because they allow a one to one mapping from any graph (including its isomorphisms) into a feature space which can be used for unambiguous graph classification (see e.g. [5]), avoiding degenerate mappings\(^2\). While any graph can be precise and completely represented in terms of its spectrum and eigenspaces, such a formulation ultimately depends on the node labeling for the correct identification of the eigenspaces. Therefore, such a representation is not invariant under node label permutations and graph isomorphisms. While the existence of a complete and invariant representation of graphs does not seem to be likely (see e.g. [8]), it is still interesting to consider additional features rather than just the graph spectrum. One of the most natural such additions can be achieved by considering also the eigenspaces of the graphs.

We consider here the problem of quantifying the importance of links in networks using the spectral decomposition of the adjacency matrix. On the basis of link spectral measurements that are described below, a fraction of the links are removed and the effect of this removal on the network topology is quantified for some specific model or real networks. For the sake of comparison, the same procedure is also applied using other, non-spectral, link measurements. There are many works dealing with the vulnerability of model and real networks to attacks on nodes or links [10]–[16]. Those works do not use spectral measurements. Spectral techniques are often used to express centrality measures of nodes [17] and for community detection [9, 18, 19], but were also used for addressing the network vulnerability problem [20]–[22]. Also in the study of expander graphs [23], spectral techniques are used with respect to resilience against attacks or faults. A graph is an expander if any subgraph is well connected with the rest of the graph. More specifically, the number of connections from the subgraph to the rest of the graph grows with the size of the subgraph. It is known that random regular graphs (graphs where all nodes have the same degree) are expanders with high probability. For regular graphs, a theorem says that graphs with high spectral gap are good expanders [23] and are resilient against link faults [24]–[27]. The spectral gap is the difference between the first and the second eigenvalue. The expander property is also related to subgraph centrality [28]. Because of this results, there are many applications for expander graphs where high resilience against faults and attacks is required, as in computer networks [29]–[31]. None of these works used spectral decomposition. Spectral decomposition is used in [32], where the authors consider the problem of reconstructing a network after perturbation.

\(^2\) In a degenerate mapping, two or more different graphs can be mapped into the same representation, precluding map inversion. Degenerate mappings are frequently used for network characterization and classification.
This paper is organized as follows. First, the basic concepts from complex networks (see e.g. [2]–[5]) and eigenspace (see e.g. [6,8]) theories are presented in an introductory and self-contained way. Then the experimental methodology is explained as regards the generation of the synthetic complex network models and measurements used for the evaluation, and this is followed by the presentation and discussion of the results.

1. Basic concepts

A graph (or complex network) $\Gamma = (V, E)$ is a discrete structure composed of a set of nodes $V$ and a set of links or links $E$, with $N = |V|$ and $L = |E|$.³ We henceforth assume that the nodes of such a graph are labeled with successive positive integer values, i.e. 1, 2, ..., $N$, and that multiple or self-connections are not present. The existence of a link extending from node $i$ to node $j$ is indicated, in the case of undirected graphs considered here, by the unordered pair $(i, j)$. A graph can be completely specified in terms of its adjacency matrix $A$ of dimension $N \times N$. The presence of the link $(i, j)$ is indicated as $A_{ij} = A_{ji} = 1$; otherwise $A_{ij} = A_{ji} = 0$. Note that the trace of $A$ is zero. A graph is said to be connected if any node can be reached from any other node while moving through the links of the graph. The sequence of nodes (or links) visited during such movements is called a path of the graph, with length equal to the number of links involved.

The characteristic polynomial of the adjacency matrix $A$ is given as $P_\Gamma(\lambda) = \det(\lambda I - A)$. The eigenvalues of $A$, assumed to be sorted and represented as $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_N$, correspond to the zeros of $P_\Gamma(\lambda)$. Because the graph is undirected, $A$ is symmetric, implying real eigenvalues. Each of these eigenvalues is a solution of the equation $Av_i = \lambda_i v_i$ for a non-zero vector $v_i$ called the eigenvector associated with $\lambda_i$. The set of the $N$ eigenvalues is called the spectrum of $\Gamma$. The largest eigenvalue $\lambda_1$ of a graph $\Gamma$ is the index of $\Gamma$. In a connected graph, the eigenvector associated with the index has all its elements positive and is called the principal eigenvector of $\Gamma$.

Because the adjacency matrices obtained by changes of the labels of the nodes are similar (and the graphs isomorphic), the spectrum of a graph does not depend on the node labeling and is therefore an invariant under label permutations. Note that, unlike eigenvalues, eigenvectors are not invariant under label permutations.

The original adjacency matrix can be expressed in terms of its spectral decomposition given as

$$A = \sum_{i=1}^{N} \lambda_i S_i,$$

(1)

where $S_i = v_i v_i^T$, $i = 1, 2, \ldots, N$.

It follows immediately from equation (1) that a graph can be completely specified in terms of its spectrum and eigenvectors.

2. Evaluation methodology

General insight. We use the entry in line $i$, column $j$ of some components of the spectral decomposition of the adjacency matrix as an indication of the ‘importance’ of...
the link between nodes \( i \) and \( j \). If all components are included, the adjacency matrix is recovered, and the importance is equal to 1 for all the links. We need to choose some components, with two possibilities presenting themselves: specifying a fixed number of the most important eigencomponents or specifying a threshold, including all components with eigenvalue above the threshold. We consider a simple case of each of these possibilities:

- The component corresponding to the largest eigenvalue, here called the \textit{largest eigencomponent}, that is, the matrix given by
  \[
  \lambda_1 S_1
  \]
  as in equation (1). Note that in this case the value associated with link \((i, j)\) is given by \(v_1(i)v_1(j)\), or the product of the eigenvector centralities of the two linked nodes.

- The sum of the contribution of all positive eigenvalues in equation (1), here called \textit{positive eigencomponent}:
  \[
  \sum_{i, \lambda_i > 0} \lambda_i S_i.
  \]

\textit{Evaluation.} To compare different ‘importance’ measurements we compute them for the same network and subsequently attack a fraction of the most important links in the network according to each measurement. The link ‘importance’ measurement whose attack yields the most significant impact on the network structure (as evaluated by some network measurements) is considered the most effective.

\textit{Networks used.} We used three network models and two real-world networks. The models used are:

- \textbf{Erdős–Rényi (ER)} Used as a null model. Here we use the ER model with fixed number of nodes and links, and random choice of the node pair to be connected by each link.

- \textbf{Watts–Strogatz (WS)} A network model with high transitivity. The model starts with a ring of nodes, each one connected to a fixed number of neighbors in each direction on the ring. Afterward, each link has one of its endpoints rewired with a fixed probability.

- \textbf{Barabási–Albert (BA)} Chosen because of the presence of hubs. Starting with a small number of seed nodes, new nodes are added one at a time. Each new node makes a fixed number of connections with existing nodes, randomly chosen with probability proportional to the number of connections (preferential attachment).

- \textbf{Holme–Kim (HK)} The growing scale-free network model with finite transitivity proposed by Holme and Kim \cite{33}. Similar to the BA model, but adding a \textit{triad formation} step where with some probability an additional link is added between the new node and a random neighbor of the preferentially chosen node.

The real-world networks used are the US Airports (USAir97) and the US Power Grid (USpowerGrid) networks, both from the Pajek data sets \cite{34}. Those networks were chosen because attacks on their links have a clear meaning, respectively deactivating a connection between two airports and cutting the power connection between two grid nodes.

\textit{Measurements.} We use two kinds of measurements: the ones used to rank the links and the ones used to evaluate the network topology.

\textit{Link ranking.} We propose that the ‘importance’ of link \((i, j)\) be taken as the entry in row \(i\), column \(j\) of the matrices defined by equations (2) and (3).

For comparison, we evaluate also the following possibilities:
Betweenness centrality of the link.

A degree product corresponding to the product of the degrees of the nodes at the two ends of the link. This measurement was also used in [35].

Random order, that is, nodes are attacked at random. This is used for comparison, and corresponds to the case where no information about links are used for the attacks.

Network topology. To evaluate the network topology, we consider the following measurements:

Transitivity Sometimes called also the clustering coefficient, this is a simple example of a measurement of local network connectivity.

Average path length This measurement, also called distance, helps quantify how the accessibility of nodes is being affected by the attack.

Largest degree Looking at the largest degree in the network we can evaluate the effect of the attack in the most important hubs.

Number of clusters As the links are removed, the network breaks into independent clusters. The number of cluster gives a measure of the fractioning of the network.

Largest cluster The size of the largest cluster, measured as the fraction of nodes in this cluster.

Attack. An attack using a given link measurement is simulated by the following procedure. First the measurement is computed for each link; the links are then ranked from largest to smallest value, and a given fraction of the links with the largest values are removed from the network. Finally, network measurements are computed for the attacked network.

3. Results and discussion

For the model networks, the experiments were run using networks of 1000 nodes and average degree 4. For Watts–Strogatz networks a rewiring probability of 0.05 was used; for the Holme–Kim networks, the probability of a triad formation step is 0.8. The results are averages over 50 networks for each model. Error bars for the standard deviation are not shown in the figures in order to increase readability. The statistical variation of the results is in most cases not important for the discussion. Where they are important, this is explicitly stated in the text.

3.1. Clusters

With the exception of ER networks, all networks considered here are connected, i.e. they consist of a single cluster of nodes reachable from each node. The ER networks can have multiple clusters, but almost all nodes are in the largest component.

To study the effect of attack on links in the clustering structure of the network, we evaluate the fraction of nodes in the largest cluster (figure 1) and the number of clusters (figure 2) as a function of the fraction of removed links.

Looking first at random attacks (continuous lines) we see that all networks, with the exception of WS networks, have a large fraction of nodes (more than 70%) in the largest component even after removing 50% of the links. The resilience of scale-free (BA, HK
Evaluating links through spectral decomposition

Figure 1. Fraction of nodes in the largest cluster as a function of the fraction of removed links, for different networks.

doi:10.1088/1742-5468/2012/01/P01015
Figure 2. Number of clusters in different networks as a function of the fraction of removed links.
and USAir97) networks under random attack is well known. Random removal of links will slightly reduce the degree of the hubs, but the overall connectivity determined by them will only be affected with the removal of a large fraction of the links. For ER networks, randomly removing links is equivalent to reducing the average degree (probability of connection of pairs). As the average degree of 4 used here is significantly above the percolation threshold, the removal of many links is needed to break the largest component. The behavior of the USpowerGrid network under random attack is similar to that of the previously discussed networks. WS networks have a sharp drop in the number of nodes of the largest cluster after removal of about 30% of the links. The rewiring process used in the construction of these networks breaks a small number of local connections to connect previously distant nodes. As the local connectivity is high, nodes tend to stay connected after removing links. But sections of the original ring can become disconnected if cut away from other sections of the ring, which is possible when the number of removed links is high. In that case, sets of nodes are disconnected from the largest component, decreasing its size sharply. Looking at the results in figure 2 for the number of clusters when randomly removing links, we see that all networks have similar behavior, with the number of clusters increasing faster as the fraction of removed links is increased.

For networks created following the ER model, the other kinds of attacks have similar results to random attacks. This is expected, as nodes and links are basically homogeneous in these networks. There is a small tendency of attacks by degree product to preserve the clustering structure. This is due to the fact that these attacks remove links from well connected nodes, having small impact on the overall connectivity. Attacks based on the largest positive eigencomponent break the network slightly faster than random attacks. This is possibly due to the increase emphasis given by this attack type to links that are important in the connection of two regions with many nodes.

In the case of WS networks, there are strong differences among the effects of different attack types. Attacks by degree product, largest eigencomponent, and, especially, betweenness centrality reduce the largest cluster much faster than random attacks, while attacks by positive eigencomponents preserve this cluster. On the other hand, the number of clusters increases faster for attacks by largest eigencomponent (which is the case also for almost all other networks except BA networks and the US airport network). In WS networks, all nodes start with the same number of links. Nodes with a larger number of links are those that received rewired links. Therefore, a long range link certainly has a high degree product. All other links that connect to the node that received the rewiring also have high degree product. Therefore, attacking links by degree product destroys these important connections, cutting the ring near the high degree node and contributing to a rapid decline in the size of the largest cluster. The same happens with attacks based on betweenness centrality, as links with high betweenness centrality are those that connect large sets of nodes on both sides, as the betweenness centrality of a link increases with the number of pairs of nodes that are connected by it. Further, the nodes that lost links during rewiring have fewer redundant connections, such that their remaining links have increased betweenness centrality. Therefore, attacking links by betweenness centrality tends to break small sets of nodes or large chunks of nodes away from the largest cluster. For a small fraction of links removed (below 20%), the largest eigencomponent attack is the most effective. A link has a high value for this metric if it connects two nodes for which the corresponding entries in the eigenvector associated with the largest eigenvalue are large.
Evaluating links through spectral decomposition

Links in the neighborhood of nodes that received rewired links have increased value for the largest eigencomponent, as this new link increases the eigenvector centrality of the nodes. For this reason, attacks by largest eigencomponent are very effective for disconnecting the neighborhood of those nodes. Attacks following the positive eigencomponents only affect the size of the largest cluster or the number of clusters for large fractions of removed links in the WS networks. The reason for this is that this measurement is large for links that take part in the formation of triangles (as will be clear when discussing the results of section 3.4), which in the WS model are local redundant links.

For BA networks, as for the ER networks, all kinds of attacks have behavior similar to that of random attacks. But because there is significant heterogeneity among the nodes, the differences are more marked in this case. To understand the results, please note that the first nodes added to the network in this model are highly connected, including among themselves, while nodes added later are more peripheral. The network has low transitivity, and the nodes with higher transitivity are the old ones or small groups in the periphery, where few choices exist for connection. This means that, when attacking using degree product or betweenness centrality, both measurements that favor links that connect hubs, the effect on the size of the largest cluster is small, as these links connect high degree nodes in a region of the network where many alternative paths exist. When attacking by betweenness centrality, hubs are not specifically targeted, such that there is higher probability that some of the connections among hubs will survive, preserving the connectivity of the network and explaining why such attacks are less effective than degree product attacks. If there is a relatively well connected subset of nodes with few connections to the hubs, all the links to the hubs high value for the largest eigencomponent. This possibly explains why attacks by largest value are slightly more effective than random attacks, even though this measure also prioritizes hubs. It is interesting that for this network model, the most effective attack is based on the positive eigencomponent, in contrast to what happens for the WS model. The explanation is that in this case links that take part in local transitive connections can be part of small local groups of nodes in the periphery, such that the removal of the links disconnects some of the nodes.

With the exception of attacks by largest eigencomponent, all attacks to links of the HK networks are similar to random attacks until a large fraction of links are removed (about 40%). At this point attacks by betweenness centrality and degree product are able to significantly reduce the size of the largest cluster. It is interesting to note that the number of clusters does not show the same sharp turn for this fraction of removed links. This can only be explained by the fact that at this point, the largest cluster is being split into two large parts, instead of having some small pieces separated. Because of the triad formation step present in this model, many links are used for local connectivity. This reduces the number of links used for global connectivity. In a scale-free network, the links that are responsible for the global connectivity are associated with hubs. This means that attacks by degree product and betweenness centrality are very effective for disrupting the largest cluster, but only if enough of the links are removed. This effect is more prominent here than in BA networks because in HK networks, some links that would be used in BA networks for global connectivity are used here (because of the triad formation step) for local connectivity. Attacks by largest eigencomponent are effective because they give more emphasis to links that connect large peripheral sets of nodes with the core of hubs than the other measurements.
The USAir97 network displays an interesting behavior under attacks: most kinds of attacks are less effective than random attacks, while betweenness centrality is highly effective. The removal of a small fraction (5%) of the links with high betweenness centrality is enough to break the network into about 60 disconnected components. This is due to the existence of isolated sets of airports with few connections to others. These few connections have high betweenness centrality, and their removal disconnects the network. The preservation of connectivity with other attacks is due to the fact that there are a huge number of connections among most of the airports, with many airports having high degree and a large amount of redundant paths and triangles. Attacks by degree product, largest eigencomponent and positive eigencomponents thus remove links from the well connected hubs, without affecting the largest cluster.

The USpowerGrid network has a response to attacks similar to that of the WS networks, with the exception of attacks by positive eigencomponents, which is more effective for reducing the size of the largest cluster and increasing the number of clusters than random attacks. To understand this result we need to know that this network has a large number of structures known as border trees (tree structures in the periphery of the network) and chains (sequences of nodes of degree 2 connected in a chain) [36, 37]. These long, linear structures imply that the links where they connect with the rest of the network have high betweenness centrality, and also a high value of the largest eigencomponent. When those links are removed, all the linear structure is separated into a small cluster. The nodes where these structures connect have higher degree, making attacks by degree product also effective. Attacks by positive eigencomponent also have a strong impact on the largest cluster and the number of clusters, as the nodes where the linear structures connect have higher local redundancy than the linear structures themselves.

3.2. Distance

The average distances computed only consider pairs of nodes that are in the same connected cluster. This means that, as links are removed, the average distance increases, until the largest cluster is broken into many small clusters. After that, the average distances decrease with further removal of links. Therefore, with respect to distances, an attack is more effective if it increases the average distances faster and reaches a maximum of average distances for a smaller fraction of removed links. Results for average distances are shown in figure 3.

To understand the results, please consider that the removal of a link will increase the average distances if no node is disconnected by the removal, but the average distances will decrease if nodes are disconnected.

First note that random link attacks are not able to reach a maximum of average distances. This is expected given that the results of section 3.1 show that the largest cluster is not broken by such attacks.

For all network types, we see that attacks by degree product and also betweenness centrality are the most effective for increasing average distances. Links that connect nodes with high degree and links with high betweenness centrality are important contributors to a small average distance, as they take part in shortest paths for many pairs of nodes. This explains why these two kinds of attacks are always effective for increasing the average distance for all network types.
Figure 3. Average distance as a function of the fraction of removed links for different networks.
In ER networks, there is small heterogeneity in the nodes, and therefore in the links that connect the nodes. But the heterogeneity of links is somewhat stronger than that of nodes when considering degree product and betweenness centrality. As connections of highly connected nodes are more important for average distances than for the connectivity of the largest cluster, attacks using these measurements are in this case significantly more effective than random attacks. But the same does not happen for attacks based on the largest eigencomponent, as this measurement breaks the network into a larger number of components than the other attacks, therefore not increasing the average distance more than random attacks.

For WS networks, the results for attacks using degree product, betweenness centrality, and positive eigencomponent are as expected considering the preceding discussion and the results of section 3. For attacks using the largest eigencomponent, the average distance grows more slowly than for random attacks, while the size of the largest cluster drops faster. This is possible because of the large number of small clusters being taken away from the largest cluster.

The effect of attacks on the average distance of BA networks is as expected considering the previous discussion about the degree product and betweenness centrality (very heterogeneous measurements in these kinds of networks), that have small effects on the size of the largest cluster, such that all removals will increase the distances. An interesting change is that, with respect to average distances, attacks using the largest eigencomponent are more effective than using positive eigencomponents. This is also reasonable, as positive eigencomponents break the BA network into a larger number of components.

Results for HK networks reveal no surprises considering the results of section 3.1 and the discussion above about the relation of breaking the cluster and average distance.

Because of the sharp breakup of the USAir97 network under betweenness centrality attack, the distances, that should increase fast under this attack, have a moderate increase. Attacks by degree product always increase distances in the largest cluster, and are therefore more effective. Attacks by largest eigenvalue are similar to product degree ones for this network, but this metric does not have such sharp differences among nodes, and the results are not so striking. Attacks by positive eigenvalues are similar to random attacks, because there is local redundancy overall in the core of this network.

As explained in section 3.1, attacks on the USpowerGrid network using betweenness centrality, degree product, and positive eigenvalues have a strong effect of disconnecting the nodes where the linear structures of this network connect, resulting in well-defined effects on distance even for a small fraction of removed links.

### 3.3. Maximum degree

The effect of attacks on the maximum degree is basically a result of how much a particular attack favors the elimination of links attached to hubs. The results are shown in figure 4. These results involve a large amount of statistical variability. Therefore, close results should be considered statistically equivalent.

For all network types, attacks by degree product are the most effective, as expected. The exception is for the USAir97 network in the region that starts at a fraction of removed nodes of 0.1 and continues until shortly after 0.4. In this region, attacks by betweenness centrality are more effective. We are not sure of the reason for this anomaly, but it is possibly related to a specific situation where some links with a high degree product do not...
Figure 4. Maximum degree in different networks as a function of the fraction of removed links.

doi:10.1088/1742-5468/2012/01/P01015
connect to the most connected nodes, i.e. hubs connect with low degree nodes. Attacks by betweenness centrality also have high impact, due to the known correlation between the two measurements. An exception is the case for the USPowerGrid network, where attacks by betweenness centrality are less effective than random attacks. This means that for this specific network the connectivity depends on a set of nodes that have low degree, which are the nodes where the border trees and chains connect.

With respect to the spectral attacks, we see that the attacks using positive eigencomponents are similar to random attacks (within statistical variability), as this measurement is related to local redundant connections, and not to the degree of the hubs. As discussed, in BA networks there are some triangles connecting hubs and in HK networks there is more locality in the periphery, explaining the slight differences from the random attack cases. The behavior of attacks by largest eigencomponent is more dependent on the network topology. For ER and WS networks, they behave similarly to random attacks, as these networks do not have hubs, and small variations in the node degree of the link endpoints is not correlated with the value associated with the link by the largest eigencomponent. For BA and HK networks, where the hubs are important for the overall connectivity, they also have an impact on the value associated by the largest eigencomponent with the link. Therefore this measurement behaves like the degree product and betweenness centrality. The USAir97 network is also a scale-free network, and has similar behavior to the BA and HK networks, with the exception that attacks by largest eigencomponent are similar to attacks by degree product for small fractions of removed links, but similar to random attacks for large fractions. The behaviors of the attacks by largest eigencomponent and by betweenness centrality on the USpowerGrid also have the interesting property that they do not affect the maximum degree until about 30% of the links are removed. The long linear structures in this network are important in terms of eigenvector centrality, but not related to the high degree nodes.

3.4. Transitivity

From the model networks used, only the WS and the HK networks have finite transitivity. We therefore show the results of effects of the attacks on these two models and the two real-world networks in figure 5.

Note that, when you remove a link from a high degree node, the expected result is an increase in the transitivity. This is due to the fact that in the computation of the transitivity, the numerator is the number of triangles, and the denominator is the number of connected triples. If you remove a link from a node, all triples that counted this link with all other links in the node are removed from the denominator, and the number of triangles is only correspondingly reduced if many triangles used this link. Also, links with high betweenness centrality are in regions where there are few optional paths, and therefore involve a smaller number of triangles. Thus, removing links with high betweenness centrality increases (in general terms) the transitivity.

For all network types, random attacks decrease the transitivity linearly, as in a random removal we are not preserving links in triangles, and if only one of the links is removed, there is one triangle less, but there is still a connected triple.

We also see, for all networks, that attacks using positive eigencomponents reduce the transitivity fast. This is a clear and consistent result, showing that links with high values of positive eigencomponents are involved in redundant local connectivity.
Results for the WS network follow the general trends discussed above. It must be noted that in this network links with high degree product and betweenness centrality are the ones close to rewired links. After removing these links, there are only the other links that were part of the original ring, and therefore belong to triangles. The result is that after some fraction of the links are removed, the transitivity starts to decrease. The results for the largest eigencomponent are similar, because the links involved are the same, but in this case, the number of links affected by the rewiring is larger than for the other measurements, such that we do not see the transitivity being reduced in our results.

For the HK networks, the results are similar to the ones for WS networks, with the difference that the betweenness centrality increases the transitivity less. This is because in...
Evaluating links through spectral decomposition

In this network the betweenness centrality is larger in some cases for peripheral links (links that are connected to smaller degree nodes), such that their removal does not decrease the denominator of the transitivity formula as much as attacks by degree product.

In the case of the USAir97 and USpowerGrid networks, we see a decrease in the transitivity with attacks by degree product. In the first case, this is due to the high number of connections among airports in the core of the network, resulting in the presence of many triangles connected to hubs. In the second case, the nodes where the linear structures are connected have higher degrees, and are the ones with triangles. For the largest eigencomponent, the behavior is the same as for the degree product for the airport network, because links that connect nodes with high eigenvector centrality are the same as the ones that connect the hubs. In the power grid network, the largest eigencomponent attack is close to a random attack, showing that transitivity is not associated with eigenvector centrality in this network.

4. Concluding remarks

Despite the increasing attention paid to the effects of complex network structure on vulnerability to attacks, relatively little attention has been paid to spectral approaches. In this work we report an investigation where spectral decomposition of the adjacency matrix is applied in order to direct the attacks, while the respective effects on the network topology are identified. More specifically, we used the eigencomponent associated with the highest eigenvalue and the sum of all positive eigenvalues. Several interesting results were identified. The most definite verified effect is the sharp decrease of the transitivity of the network implied by attacks guided by the positive eigencomponent. Not only does the transitivity decrease fast with the removal of links, but also the effect is consistent in all network topologies studied. The effect of the diverse kinds of attacks on network measurements depend largely on the specific network types. For instance, the effects on the cluster structure of ER and BA networks are all similar to random attacks, but WS and HK networks are more sensitive to attacks by node degree and betweenness centrality. On the other hand, attacks by largest eigencomponent tend to break all kinds of networks (with the exception of the US airport network) into many small clusters. It is clear also that, although there are correlations among the degree, betweenness centrality and eigenvector centrality, attacks directed by these measurements can have different impacts depending on the network structure, showing that the measurements quantify different centrality aspects. Though spectral decomposition approaches have seldom been used in complex networks research, our results show that they are capable of emphasizing specific aspects of the network topology. It is therefore expected that additional applications could benefit from considering spectral decomposition.

Acknowledgments

Luciano da Fontoura Costa is grateful to FAPESP (2005/00587-5) and CNPq (308231/03-1) for financial sponsorship.

References

[1] Tutte W T, 2001 Graph Theory (Cambridge: Cambridge University Press)
[2] Albert R and Barabási A L, Statistical mechanics of complex network, 2002 Rev. Mod. Phys. 74 47

doi:10.1088/1742-5468/2012/01/P01015
Evaluating links through spectral decomposition

[3] Newman M E J, The structure and function of complex networks, 2003 SIAM Rev. 45 167 arXiv:cond-mat/0303516

[4] Boccaletti S, Latora V, Moreno Y, Chavez M and Hwang D-U, Complex networks: structure and dynamics, 2006 Phys. Rep. 424 175–308

[5] da Fontoura Costa L, Rodrigues F A, Travieso G and Villas Boas P, Characterization of complex networks: a survey of measurements, 2007 Adv. Phys. 56 167

[6] Biggs N, 1993 Algebraic Graph Theory (Cambridge: Cambridge University Press)

[7] Chang F R, 1997 Spectral Graph Theory (Providence, RI: American Mathematical Society)

[8] Cvetkovic D, Rowlinson P and Simic S, 1997 Eigenspaces of Graphs (Cambridge: Cambridge University Press)

[9] Newman M E J, Finding community structure in networks using the eigenvectors of matrices, 2006 Phys. Rev. E 74 036104

[10] Latora V and Marchiori M, Vulnerability and protection of infrastructure networks, 2005 Phys. Rev. E 71 056109

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

[12] Estrada E, Graph spectra and the structure of complex networks, 2008 Technical Report 26 University of Strathclyde

[13] Ma X, Gao L and Yong X, Eigenspaces of networks reveal the overlapping and hierarchical community structure more precisely, 2010 J. Stat. Mech. P08012

[14] Qin S and Dai G, Community finding of scale-free network: algorithm and evaluation criterion, 2009 Mining Complex Data vol 165 Studies in Computational Intelligence ed D Zighed, S Tsumoto, Z Ras and H Hacid (Berlin: Springer) pp 223–42

[15] Goerdt A and Molloy M, Analysis of edge deletion processes on faulty random regular graphs, 2000 Lect. Notes Comput. Sci. 1776 38–47

[16] Goerdt A, The giant component threshold for random regular graphs with edge faults, 2001 Theor. Comput. Sci. 259 307

[17] Goerdt A, Random regular graphs with edge faults: expansion through cores, 2001 Theor. Comput. Sci. 264 91

[18] Estrada E, Spectral scaling and good expansion properties in complex networks, 2006 Europhys. Lett. 73 649

[19] Song D, Zuckerman D and Tygar J D, Expander graphs for digital stream authentication and robust overlay networks, 2002 Proc. 2002 Symp. on Security and Privacy IEEE Computer Society, pp 258–70

[20] Loguinov D, Casas J and Wang X, Graph-theoretic analysis of structured peer-to-peer systems: routing distances and fault resilience, 2005 Trans. Netw. 13 1107

[21] Gkantsidis C, Mihail M and Saberi A, Random walks in peer-to-peer networks: algorithms and evaluation, 2006 Perform. Eval. 63 241

doi:10.1088/1742-5468/2012/01/P01015
Evaluating links through spectral decomposition

[32] Liu D, Wang H and Van Mieghem P, *Spectral perturbation and reconstructability of complex networks*, 2010 Phys. Rev. E 81 016101

[33] Holme P and Kim B J, *Growing scale-free networks with tunable clustering*, 2002 Phys. Rev. E 65 026107

[34] Batagelj V and Mrvar A, 2006 Pajek datasets http://vlado.fmf.uni-lj.si/pub/networks/data/

[35] Kaiser M and Hilgetag C C, *Edge vulnerability in neural and metabolic networks*, 2004 Biol. Cybern. 90 311

[36] Villas Boas P R, Rodrigues F A, Travieso G and da Fontoura Costa L, *Border trees of complex networks*, 2008 J. Phys. A: Math. Theor. 41 224005

[37] Villas Boas P R, Rodrigues F A, Travieso G and da Fontoura Costa L, *Chain motifs: the tails and handles of complex networks*, 2008 Phys. Rev. E 77 026106

doi:10.1088/1742-5468/2012/01/P01015