Lightweight centrality measures in networks under attack *

Giorgos Georgiadis † Lefteris Kirousis ‡†

Abstract

In this paper we study deliberate attacks on the infrastructure of large scale-free networks. These attacks are based on the importance of individual vertices in the network in order to be successful, and the concept of centrality (originating from social science) has been already utilized in their study with success. Some measures of centrality however, as is betweenness, have disadvantages that do not facilitate the research in this area. We show that with the aid of scale-free network characteristics such as the clustering coefficient we can get results that balance the current centrality measures, but also gain insight in the workings of these networks.

Keywords:
dynamics of scale-free networks, centrality, clustering coefficient, network attack

1 Introduction

Recently there has been an increase of interest in many natural and artificial large scale networks. For example, it is estimated that the network of web pages currently consists of several billions of vertices [1]. Many companies owning a search engine would like to know the specific characteristics of this

*This research was partially supported by the EU within the 6th Framework Programme under contract 001907 "Dynamically Evolving, Large Scale Information Systems" (DELIS) and was also partially supported by the European Social Fund (ESF), Operational Program for Educational and Vocational Training II (EPEAEK II), and particularly the Program *Pythagoras*.

†Department of Computer Engineering and Informatics, University of Patras, 26504 Patras, Greece. Email address: {georgiad, kirousis}@ceid.upatras.gr

‡Research Academic Computer Technology Institute, P.O. Box 1122, 26110, Patras, Greece

network for purposes of page indexing and maybe to predict, up to a point, its future behavior. In general, a network consists of objects, of different kind in each area of interest, which are represented by vertices, and connections between them, represented by edges. Many theoretical results exist due to graph theory of discrete mathematics, which handles such objects. This kind of modelling is possible for a variety of large scale networks, both natural occurring and artificial, such as networks of acquaintances, citation, food chains, infections, proteins [2, 3, 4, 5, 6, 7] or networks of power grids, internet infrastructure, web pages, and so on [8, 9, 10, 11, 12, 13, 14].

Especially in the case of social networks of any kind, they have been studied by scientists in the social science for quite a while, with methods such as questionnaires and personal interviews. One persistent question was that of the centrality of an individual in such a network, or how well "connected" this person is in his environment. For example a measure of this connectivity is a vertice's degree, that is the number of its immediate neighbours. The size of these kind of networks is in the order of several tens or in some cases several hundreds of vertices and so the research is not directly applicable to the large networks arising, mainly technological. This is just one of the problems scientists face with the large scale networks.

Another problem is the nature of these networks. The most prominent model in graph theory until recently have been the random graph model introduced by Erdős and Rényi [15]. In this model, any two vertices have equal probability to be connected by an edge. This model is very well studied and many results exist on it, but unfortunately does not describe our observations in real world networks. In many real world networks there exist a percent of vertices that appear to be better connected to the rest of vertices. Furthermore, during the network growth, they gain easier connections and certainly not with equal probability to the other vertices. One of the unique characteristics of these networks, that distinguishes them from previous studied networks, is the power law form of the distribution of the vertex degrees.

In this study we address the problem of network attack. We are assuming the existence of an adversary that wants to harm a network, by directly attacking and removing the vertices comprising it. He has the ability to measure some variables of the network in order to make educated guesses as to what vertex should target next. In our experiments we measure the efficiency of strategies based on such measures as to the computational time needed to target a vertex on the specific network and the end result after the attack. We care about measures that produce most harm with little effort. The strategies we are using will be based on the centrality measures taken

from traditional social network research.

In this paper we take into account previous similar studies and we compare our findings with theirs. We propose an attack strategy that is a trade-off between worst and best strategy so far and has significant and unique advantages. We also offer insight to the workings in power-law graphs and indicate future research areas.

The paper is organized as follows. In Section 2 we review the fundamental concepts needed in our study, along with a separate discussion on the most widely used measures of centrality. In Section 3 we introduce the use of centrality measures as attack strategies. Our experiments can be found in Section 4 along with our analysis on the results.

2 Fundamental Concepts and Definitions

In this section we define the graph-related concepts that we will use in our experiments and analysis, along with the main notions of centrality that are of interest to us.

2.1 Graph-related concepts

Throughout this paper we represent a network as an undirected, unweighted graph $G(V, E)$, where V is the set of vertices (i.e. computers) and E is the set of edges (i.e. communication links). Their sizes are $|V| = n$ and $|E| = m$ respectively. The degree k_v of a vertex v is the number of edges originating from or ending to vertex v. We are interested only in graphs generated by the preferential attachment procedure, first proposed by Barabási and Albert [16], which we will briefly describe here. The iterative creation process consists of 4 steps:

Step 0: Initially the graph has n_0 vertices and no edges.

Step 1: Add a new vertex v to the graph.

Step 2: Create l edges, each time connecting the new vertex v to a vertex w, with probability proportional to this vertice's degree:

$$p_w = \frac{k_w}{\sum_{u \in V} k_u}$$

Step 3: Repeat steps 1 and 2 for $(n - n_0)$ rounds.

The end result of this procedure is a graph of n vertices and $m\,(n - n_0)$ edges, with vertex degree distribution $P\,(k)$ that follows a power law, with exponent $\gamma = 3$ ($P\,(k) \propto k^{-3}$). We will call such a graph a Barabási-Albert network, or BA network for short. Of course there are many models with creational procedures that generate graphs with power law degree distributions (like the Watt's "small worlds" [17, 18]), but we feel that the classical preferential attachment model describes complex networks generation in a more general way. Other than that, our specific results may differ in other models but the essence of our insights should still apply.

Other concepts that will be used are that of the 1-neighbourhood and 2-neighbourhood of a vertex. Having a vertex v as center, its 1-neighbourhood $\Gamma_1\,(v)$ consists of all vertices at distance 1, i.e. its direct neighbours. Such neighbours from now on will be called *first-neighbours* and of course it holds that $|\Gamma_1\,(v)| = k_v$. Similarly, the 2-neighbourhood of a vertex v consists of all vertices at distance exactly 2 (from now on *second-neighbours*) and it holds that $|\Gamma_2\,(v)| \leq \sum_{w \in \Gamma_1(v)} k_w$. The inequality in the above expression stands for the fact that some first-neighbours of v may have common neighbours, thus limiting the number of (unique) vertices in $\Gamma_2\,(v)$. This phenomenon is called *clustering* [19, 13, 20] and is not only possible but characteristic of power law graphs. This relation between a vertex and its first- and second-neighbours lead to the emergence of several structures in the graph, most common of which is the triangle. In a triangle, three vertices are joined by three edges, one for each pair of vertices. The existence of triangles is characteristic of a power law graph, and its the feature that makes them so popular in different disciples: for example, in social science, two of one's friends have greater probability of knowing each other than two random-picked strangers.

2.2 Standard Centrality Measures

2.2.1 Degree Centrality

The degree centrality measure gives the highest score of influence to the vertex with the largest number of first-neighbours. This agrees with the intuitive way to estimate someone's influence from the size of its immediate environment. The degree centrality is traditionally defined analogous to a vertice's degree, normalized over the maximum number of neighbours this vertex could have. Thus, in a network of n vertices, the degree centrality of

vertex i, C_i^D, is defined as:

$$C_i^D = \frac{k_i}{n-1} \tag{1}$$

The normalization in the region $[0, 1]$ is used here to make the centrality of different vertices comparable, and also independent of the size of the network.

2.2.2 Closeness Centrality

This notion of centrality focuses on the idea of communication between different vertices. The vertex which is "closer" to all vertices gets the highest score. In effect, this measure indicates which one of two vertices needs fewer steps in order to communicate with some other vertex. Because this measure is defined as "closeness", the inverse of a vertice's mean distance from all others is used. Hence, if C_i^C is the closeness centrality, and d_{ij} the shortest distance between vertices i and j in terms of edge steps:

$$C_i^C = \frac{n-1}{\sum\limits_{j \in V} d_{ij}} \tag{2}$$

Again, this measure is normalized in the region $[0, 1]$. Additionally, it should be stated that the distance between two disconnected vertices must be a predefined very large value and not infinite, if it is desirable to discern among low closeness scores.

2.2.3 Betweenness Centrality

Betweenness centrality refines the concept of communication, introduced in closeness centrality. Informally, betweenness centrality of a vertex can be defined as the percent of shortest paths connecting any two vertices that pass through that vertex. The normalized version divides this value with the maximum possible betweenness centrality, that is all possible shortest paths in a completely connected graph. If C_i^B is the betweenness centrality of vertex i, (u, i, v) is the set of all shortest paths between vertices u and v passing through vertex i and (u, v) is the set of all shortest paths between vertices u and v, then:

$$C_i^B = \frac{\sum\limits_{u \in V} \sum\limits_{v \neq u \in V} \frac{|(u,i,v)|}{|(u,v)|}}{(n-1)(n-2)} \tag{3}$$

This definition of centrality explores a vertice's ability to be "irreplaceable" in the communications of two random vertices. It is of particular interest in the study of network attacks, because at any given time the removal of the maximum betwenness vertex seems to cause maximum damage in terms of connectivity and mean distance in the network. Its main disadvantage is that the summation operator practically means that it needs global information about the network, in order to compute a single vertice's betweenness, and that is simply not possible in many contexts. For the same reason it is expensive in computing time to compute a vertice's score, although this disadvantage was significantly improved recently [21, 22]. The importance of betweenness centrality as an attack strategy is further discussed below.

3 Centrality Measures As Attack Strategies

It has been shown in the past that the "random vertex hit" strategy performs poorly [23, 24], due to the hierarchical effect these networks present, i.e. a random vertex has increased probability to be one of the less connected vertices, since there are so many of them. So it is desirable to use a strategy that achieves better results, and such strategies could be based upon a vertex measure that can profile each vertice's potential by its value only. A number of publications exist [23, 25, 26, 24, 27], addressing the question of which strategy is best in achieving maximum destructive result with less vertex hits, the most extensive of which, to our knowledge, is [23]. Summarizing the results, the comparison is based upon two axes: different strategies and recalculation of measures. The different strategies studied are vertex deletion based upon degree centrality scores and upon betweenness centrality scores, and it is clearly shown that betweenness produces better results. The recalculation of the involved measures refers to the recalculation of degree/betweenness centrality after each vertex deletion.

In this section we introduce a strategy that balances the advantages and disadvantages of the above mentioned strategies. We will not study the closeness centrality, as it has the same basic flaws as betweenness and none of its advantages. Furthermore we will focus on the recalculated versions, since the distribution of these measures may vary significantly between deletions.

3.1 How to measure destructive power?

We are interested in the destruction of the network under consideration. Ideally that would meant the isolation of each vertex, but it can be argued that it is enough to break the network to a sufficiently large number of connected

components. We choose to examine only the size of the largest component, as a particulary small largest component would mean that the network has degenerated into many small connected components. Additionally, we can measure directly the impact of vertex deletions in the hierarchical structure by examining what happens to the largest component: a successful attack would probably target this component and shrink its size dramatically. Another reason is that this technique has been used successfully during previous studies [23], and its use will make our results directly comparable. We use specifically a normalization over the largest component size with the initial network size, in order to produce a percentage comparable between different size networks.

Since we start with a connected network, it would take some time before it becomes disconnected, and during that time the size of the largest component would not carry significant information. Thus, in addition to the largest component size, we use the mean shortest path length of the network, and specifically its mean inverse. The mean shortest path length is the mean length of all shortest paths in the network, between all pairs of vertices. If by d_{uv} we denote the length of the shortest path between vertices u and v, then the mean shortest path length l, in a network of n vertices, is

$$l = \frac{\sum\limits_{u \in V} \sum\limits_{v \neq u \in V} d_{uv}}{n\,(n-1)}.$$

The mean inverse of shortest path length l^{-1} is defined as

$$l^{-1} = \frac{\sum\limits_{u \in V} \sum\limits_{v \neq u \in V} \frac{1}{d_{uv}}}{n\,(n-1)}$$

In practice we use the mean inverse of shortest path length because by doing that we nullify the effects of disconnected vertices and their "infinite" distance. Increasing mean value of this measure means that average distances in the network are increasing, and this subsequently means that the attack in the network produces quantifiable, destructive results. Clearly, since we use the mean inverse of this measure, we expect it to decrease with time.

3.2 Standard Centrality Measures Explained

As it was already mentioned, the random vertex hit strategy has practically no effect to the network's integrity, and that is because it cannot take

into consideration its hierarchical structure. This is exactly where degree based attack succeeds. By targeting the highest degree vertices first, it attacks directly the global network connectivity. It must be pointed out that not all properties of Barabási-Albert networks are known. Initially it was believed that high degree vertices were connected with other high degree vertices preferably over lower degree vertices (assortative mixing) [28]. Recent studies [29] show that Barabási-Albert networks are rather neutral on this property, and in some cases even show the opposite behavior (disassortative mixing), i.e. high degree vertices prefer lower degree vertices to connect to. We believe this observation can explain the success, albeit partial, of this strategy, as in the disassortative mixing the deletion of highest degree vertex would effect many vertices: but since this is not a predominant phenomenon the effectiveness of this strategy would be limited.

On the other hand, the betweenness based strategy seems ideal, especially with the performance metrics used (mean shortest path length, largest component size). By definition, the betweenness measures the ability of a vertex to be irreplaceable in shortest paths throughout the network. So when this vertex is removed, inevitably all shortest paths depended on it will be removed also, and equally long or longer paths would take their place. This has an obvious impact in the mean shortest path length, which is constantly non-decreasing, at least as long as a unique giant component exists. Such high betweenness vertices, which connect many others with shortest paths, would be initially located in the largest component as most vertices would be located there. Therefore, the failure of these paths effects also the largest component size, since multiple failures may produce disconnected vertices. Similar arguments can be used with the closeness centrality.

3.3 Proposed Strategies

We propose a family of strategies based, in part, on edge degree. Although a formal definition of edge degree does not exist, we experimented with several possible definitions, all based on vertex degree. Specifically, an edge's degree has some connection with the endpoint vertices of this edge. As was the case in [23] we settled with the edge degree being the product of the endpoint vertices' degrees, as it followed closely our intuition on the importance of edges. If $e = (w, u)$ an edge with endpoints w and u, having degrees k_w and k_u respectively, its edge degree $k_e^{\Gamma_1}$ is defined as:

$$k_e^{\Gamma_1} = k_w \cdot k_u$$

The first strategy which uses the edge degree to select vertices does so

by first selecting the edge with maximum degree, and then the vertex of this edge with maximum (vertex) degree. In case of multiple edges/vertices with same (maximum) degree, we choose uniformly at random. Note that this strategy examines the immediate neighbourhood of each endpoint vertex, and scores higher edges having endpoints with large 1-neighbourhoods. From now on we will refer to this strategy as "1-neighbourhood edge degree" strategy.

The second strategy defines edge degree as the product of the 2-neighbourhoods of its endpoint vertices. This 2-neighbourhood edge degree $k_e^{\Gamma_2}$ of an edge $e = (w, u)$ is defined formally as:

$$k_e^{\Gamma_2} = \sum_{i \in \Gamma_1(w)} k_i \cdot \sum_{j \in \Gamma_1(u)} k_j$$

The vertex selection is exactly the same as before: choose the edge with maximum degree and then the endpoint vertex with maximum degree. We will refer to this strategy as "2-neighbourhood edge degree" strategy.

The third proposed strategy is based on the "2-neighbourhood edge degree", as defined above. The main difference is that it penalizes the existence of triangles in which the edge is present. Specifically, it divides the above computed edge degree by the number of triangles that this edge participates in, plus one to avoid division by zero. Thus, if T is the number of triangles involving the edge in question as a side of the triangle, the formal definition of the alternative edge degree is:

$$K_e^{\Gamma_2} = \frac{\sum_{i \in \Gamma_2(w)} k_i \cdot \sum_{j \in \Gamma_2(u)} k_j}{T + 1}$$

We will refer to this strategy as "2-neighbourhood edge degree with penalty".

4 Experiments

For the experiments we used networks of 1500 vertices, created by the BA procedure mentioned in Section 2. The parameters of importance are the size of the initial network (before the procedure starts adding vertices) and the degree of each added vertex. We used degree 5 for each new vertex and we kept the initial network small, consisting of 5 vertices connected with random edges. Each edge between two vertices had 0.5 probability of existing, so as to differentiate the vertices for the growing procedure. We kept the initial network intentionally small because larger (initial) networks

create larger gaps between high degree and low degree vertices during the network growth. As a result, highly central vertices are fewer and more easily recognizable by any targeting strategy and are diminished quickly, leaving no time for the various strategies to produce different results.

The results are shown in figures 4, 5, 6 and 7, located at the appendix. The inverse mean shortest path length, the size of the largest component and the clustering coefficient are measured after each vertex deletion and shown in separate illustrations. For each of these parameters, five different data sets exist, corresponding to the five strategies under study (highest degree, betweenness, 1-neighbourhood edge degree, 2-neighbourhood edge degree and 2-neighbourhood edge degree with penalty). Their values at each deletion step are the average of 50 experiments with different networks of 1500 vertices. Of the monitoring parameters, the easiest to read is the size of the largest component and its transition is shown in magnification in fig. 6. It is easy to see the relation between the various strategies, as each one, having done a preliminary work, performs better or worse than the others during the transition.

4.1 Drilling into experimental results

The inverse mean shortest path length initially declines, meaning that distances inside the network begin to grow in general. At some point this trend is reversed because during the deletion process the connected components become quite small and the distances inside them are even smaller than in the initial network. Thus the inverse length continues to increase as connected components are cut into smaller pieces and this continues until they stop breaking up. At this point the inverse length is at its maximum value and almost all significant vertices are gone, as subsequent deletions leave the components at roughly the same size. From this maximum point on, the network continues to shrink with almost constant rate. The five strategies differ mainly in their ability to break the already small connected components to even smaller ones, leading to higher maximum points, as is shown in fig. 4 in appendix.

The size of the largest component is a more straightforward measure. After each deletion the size of a component is reduced by one and, at least initially, the deleted vertex is selected from the largest component. As more "central" vertices are deleted, critical paths collapse and the largest component breaks into smaller pieces. Figure 5 in appendix shows clearly that there is an early stage where the strategies built up tensions by deleting important vertices, a transition phase where very important vertices are

gone and each deletion breaks the largest component in small pieces, followed by a slow shrinking of the largest component. The transition phase is where the various strategies compete, and betweenness is the most successful in making the transition in fewer deletions. However, comparable to the degree strategy which performs poorly, our proposed strategies bridge the gap with betweenness by up to 23%, 29% and 55% for 1-neighbourhood, 2-neighbourhood and 2-neighbourhood with penalty edge degree respectively.

The clustering coefficient during the early stage is decreasing by orders of magnitude, meaning that the deleted vertices, tagged as central by the various strategies, contribute greatly to the global clustering coefficient (fig. 7 in appendix). During the transition phase it appears fluctuating due to the shrinking of the largest component and the increase in the number of components, and in the last phase it is completely wiped out as triangles do not practically exist. The betweenness stands out, since during the transition phase it creates a seesaw effect on the clustering coefficient, never destroying all triangles in the connected components. Although we have no solid evidence, we feel that this observation is the key to understanding the role and the success of betweenness, and to replicate its behaviour in other measures.

In order to understand why the proposed strategies work as they do, we focus on a high degree edge and examine its specific characteristics (fig. 1). Just by looking at the high degree edge alone, one can argue that it connects high degree vertices, therefore is important for the communication of $(k_w - 1)$ vertices (at the one endpoint) with another $(k_u - 1)$ vertices (at the other endpoint). So its deletion alone would probably effect many vertices and the distances between them. As for the highest degree endpoint (which will eventually be deleted), one must keep in mind that high degree vertices don't usually connect to other high degree vertices. On one hand, deleting high degree vertices is a successful enough strategy on its own (see degree centrality strategy), but with maximum edge degree we ensure that the high degree vertex to be deleted will be connected to the highest possible degree vertex (given its not a common phenomenon) and the deletion will effect a greater number of vertices. On the other hand, since high degree vertices don't connect often, this filter differentiates adequately otherwise equal vertices (i.e. when degree centrality is used).

The 2-neighbourhood edge degree strategy operates in an similar way. The same arguments as above are still valid here, i.e. a high degree edge connects more vertices than a low degree one. The main difference is that we are now talking about vertices that are part of the 2-neighbourhood of the one endpoint vertex which connect with the vertices of the 2-neighbourhood of

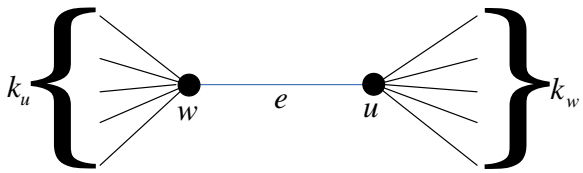

Figure 1: 1-neighbourhoods of two connected vertices

the other endpoint vertex. This may be more reliable than the 1-neighbourhood of the endpoints since there seems to exist a light disassortative mixing. This means that high degree vertices connect to lower degree, thus their influence dies out quickly as we move further from their center. By using 2-neighbourhoods we favor vertices that their influence two steps away is still strong. The downside of this strategy is its slightly larger computational load compared to the 1-neighbourhood edge degree, but this is still far from that of betweenness centrality. Furthermore it uses semi-local information for its computation, which, we estimate, should not be a problem in most practical uses.

The mechanism behind the alternative 2-neighbourhood edge degree strategy is somewhat different. Obviously the same arguments of the two previous strategies are still valid here. An instance that is handled differently is shown in fig. 2. Normally the 2-neighbourhood edge degree of this edge would be

$$\sum_{i \in \Gamma_2(w)} k_i \cdot \sum_{j \in \Gamma_2(u)} k_j$$

but since it participates in a triangle due to the common neighbour of both endpoint vertices, this edge degree is divided by 2. Thus edges that connect two "smaller" vertices in terms of 2-neighbourhoods can have larger edge degree and be selected instead. The situation is even worse if the edge participates in more triangles, as in fig. 3, for its edge degree would be even smaller. The edge degree can gradually increase, if vertices comprising the triangles become selected for the deletion process, and the triangles collapse.

This edge degree with penalty measures the size of two 2-neighbourhood connecting through one edge, as was the case in the previous 2-neighbourhood edge degree. But it also considers the importance of alternative paths

between these two 2-neighbourhoods. There is no doubt that selecting one endpoint of an edge participating in many triangles will destroy these triangles also, but vertices connecting same size 2-neighbourhoods with no triangles are more important to the whole network, and this is expressed by this measure and verified by our results.

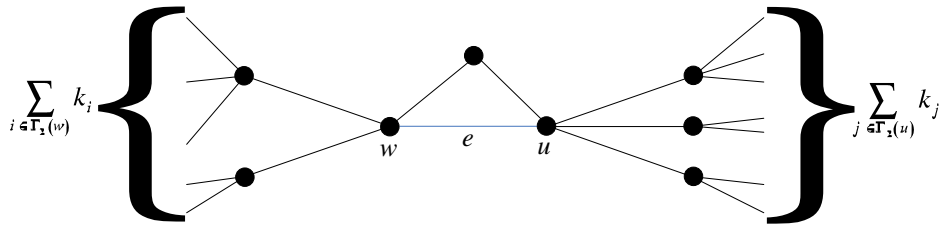

Figure 2: 2-neighbourhoods of two connected vertices

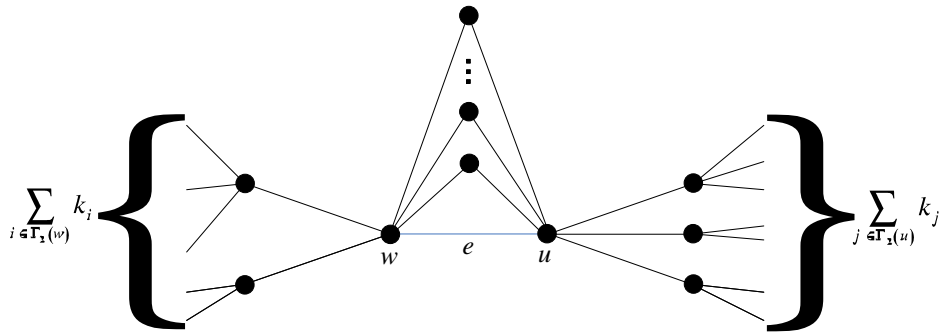

Figure 3: 2-neighbourhoods of two connected vertices with triangles

4.2 Algorithm Complexities

Theorem 1. *The worst case time complexities of the proposed strategies are* $O(m)$, $O(m\sqrt{n})$ *and* $O(mn)$ *for 1-neighbourhood, 2-neighbourhood and*

2-neighbourhood with penalty edge degrees, respectively. Furthermore the average case time complexity is $O(m)$ for all strategies.

Proof. The 1-neighbourhood edge degree just multiplies two integers, namely the vertex degrees of the edge endpoints, for all edges. The query of a vertice's degree is an $O(1)$ operation in the LEDA enviroment we are using [30], so the total cost of computing 1-neighbourhood edge degree for all edges is $O(m)$ in any case.

The 2-neighbourhood edge degree queries for each endpoint vertex, the vertex degree of all its neighbours and sums it, multipling the two endpoint sums, and does this for all edges. The iteration of all neighbouring vertices of a vertex has guaranteed asymptotic complexity on the number of actual neighbours. So in the worst case, this computation has $O(mk_{max})$ complexity, where k_{max} is the maximum degree in the network. Specifically for the Barabási-Albert network there exist an analytic solution [31] for the degree $k_i(t)$ of a vertex i at timestep t, as

$$k_i(t) = l\sqrt{\frac{t}{t_i}} \tag{4}$$

where l is the number of edges per new vertex and t_i is the timestep when vertex i was added to the network. After n timesteps the maximum degree in the network is

$$k_{max} = O\left(\sqrt{n}\right)$$

And so the worst case time complexity is $O(m\sqrt{n})$. But since the mean degree in the network is $\overline{k} = 2l$ (as can be easily seen), the average case time complexity is $O(m)$.

The third strategy is computed as above, but for one endpoint of the edge, we scan its neighbour's neighbour lists to find the other endpoint (indicating the existence of triangles). So its worst case complexity is $O\left(m{k_{max}}^2\right)$ and thus $O(mn)$. Similarly its average case complexity is $O(m)$. □

5 Conclusions

We have studied three novel strategies in network attack and compared them with two traditional approaches, degree and betweeness centrality, both with its own merits and flaws. These strategies have proven to be simple enough to implement, with low computational cost, and yet efficient compared to the best strategy. In addition to their value as attack strategies, they can help to shed light to the inner workings of a power law network. One of

the great difficulties in their study is our ignorance as to what measures are important to the behaviour of these networks. Our experiments link the degree-degree correlations among vertices with their centrality in the network. Furthermore, to the extent of our knowledge, it is the first time that the clustering effect is linked to the centrality of a vertex. Although we know this is responsible for the "denseness" of power law networks, its exact role remains unclear. Our third strategy indicates that it plays a major role in conjunction with other phenomena, such as the degree-degree correlations. It would be of interest to study several models of networks, other than the BA model, that show documented assortative or disassortative behavior and models that have known clustering coefficient distributions, in order to explore further these effects of our strategies. Furthermore, it is the subject of future research to determine whether the utilization of other network structures, similar to the triangles we are using in this study, will help bridge the gap between local strategies and global ones, as is betweenness. This development will not only help us to study larger networks but will also reveal the role of individuals in such a vast network.

References

[1] Baldi P, Frasconi P, and Smyth P. *Modeling the Internet and the Web: Probabilistic Methods and Algorithms*. Wiley, 2003.

[2] Jeong H, Tombor B, Albert R, Oltvai N, and Barabási A. The large-scale organization of metabolic networks. *Nature*, 407:651–654, 2000.

[3] Liljeros F, Edling R, Amaral N, Stanley E, and Aberg Y. The web of human sexual contacts. *Nature*, 411:907–908, 2001.

[4] Mariolis P. Interlocking directorates and control of corporations: The theory of bank control. *Social Science Quarterly*, 56:425–439, 1975.

[5] Pimm L. *Food Webs*. University of Chicago Press, 2nd edition, 2002.

[6] Podani J, Oltvai N, Jeong H, Tombor B, Barabási A, and Szathmáry E. Comparable system-level organization of archaea and eukaryotes. *Nature Genetics*, 29:54–56, 2001.

[7] Jones J and Handcock M. An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings Of The Royal Society Of London Series B-Biological Sciences*, 270(1520):1123–1128, June 2003.

[8] Albert R, Jeong H, and Barabási A. Diameter of the world wide web. *Nature*, (401), 1999.

[9] Chen Q, Chang H, Govindan R, Jamin S, Shenker S, and Willinge W. The origin of power laws in internet topologies revisited. In *IEEE Infocom 2002*, 2002.

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

[11] Huberman A. *The Laws of the Web: Patterns in the Ecology of Information*. The MIT Press, 2001.

[12] Lawrence S and Giles L. Searching the world wide web. *Science*, 280:98–100, 1998.

[13] Ravasz E and Barabási A. Hierarchical organization in complex networks. *Phys. Rev. E*, 67(026112), 2003.

[14] Watts J. A simple model of global cascades on random networks. In *Proc. Natl. Acad. Sci. USA*, volume 99, pages 5766–5771, 2002.

[15] Erdős P and Rényi P. On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–291, 1959.

[16] Barabási A and Albert R. Emergence of scaling in random networks. *Science*, 286:509–512, October 1999.

[17] Watts J and Strogatz H. Collective dynamics of 'small-world' networks. *Nature*, pages 393–393, 1998.

[18] Watts J. *Small Worlds*. Princeton Uni. Press, 1999.

[19] Newman J. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2003.

[20] Bornholdt S and Schuster G, editors. *Handbook of Graphs and Networks*. Wiley-VCH, 2002.

[21] Brandes U. A faster algorithm for betweenness centrality. *Journal of Mathematical Sociology*, 25(2):163–177, 2001.

[22] Newman J. Scientific collaboration networks. ii. shortest paths, weighted networks, and centrality. *Phys. Rev. E*, 64(016132), 2001.

[23] Holme P, Kim J, Yoon No, and Han K. Attack vulnerability of complex networks. *Phys. Rev. E*, 65(056109), 2002.

[24] Cohen R, Erez K, Ben-Avraham D, and Havlin S. Resilience of the internet to random breakdowns. *Phys. Rev. Let.*, 85:4626–4628, 2000.

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

[26] Callaway S, Newman J, Strogatz H, and Watts J. Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.*, 85:5468–5471, 2000.

[27] Cohen R, Erez K, Ben-Avraham D, and Havlin S. Breakdown of the internet under intentional attack. *Phys. Rev. Let.*, 86:3682–3685, 2001.

[28] Krapivsky L and Redner S. Organization of growing random networks. *Phys. Rev. E*, 63(066123), 2001.

[29] Zhuang-Xiong H, Xin-Ran W, and Han Z. Pair correlations in scale-free networks. *Chinese Physics*, 13(3):273–278, 2004.

[30] Mehlhorn K and Nahe S. *LEDA : A Platform for Combinatorial and Geometric Computing*. Cambridge University Press, Cambridge, 2000.

[31] Barabási A, Albert R, and Jeong H. Mean-field theory for scale-free random networks. *Physica A*, 272:173–187, 1999.

Appendix

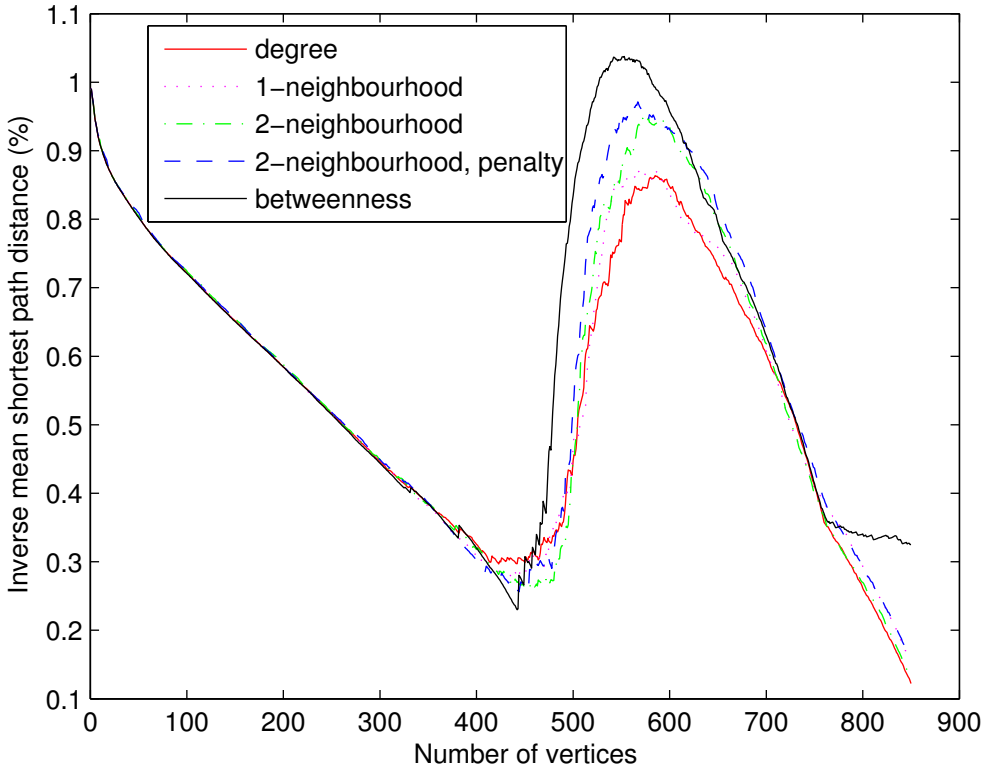

Figure 4: Inverse mean shortest path length shown as percent of the initial length, as vertices are sequentially deleted. The results are the average of 50 experiments with networks of 1500 vertices.

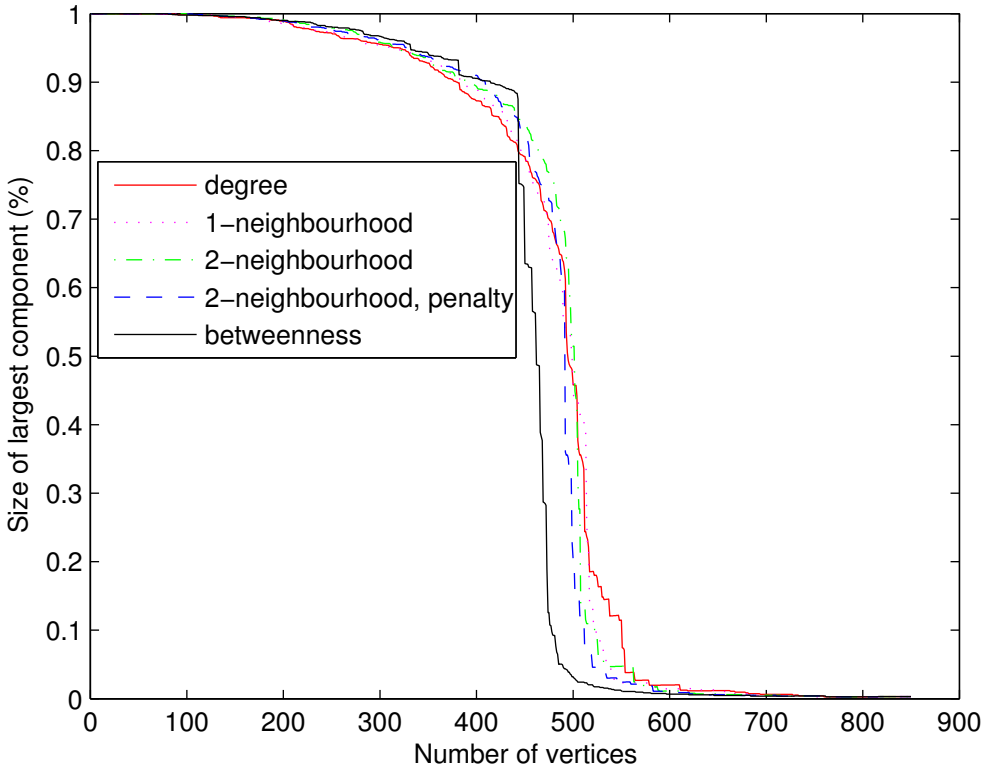

Figure 5: Size of largest connected component shown as percent of the initial size, as vertices are sequentially deleted. The results are the average of 50 experiments with networks of 1500 vertices.

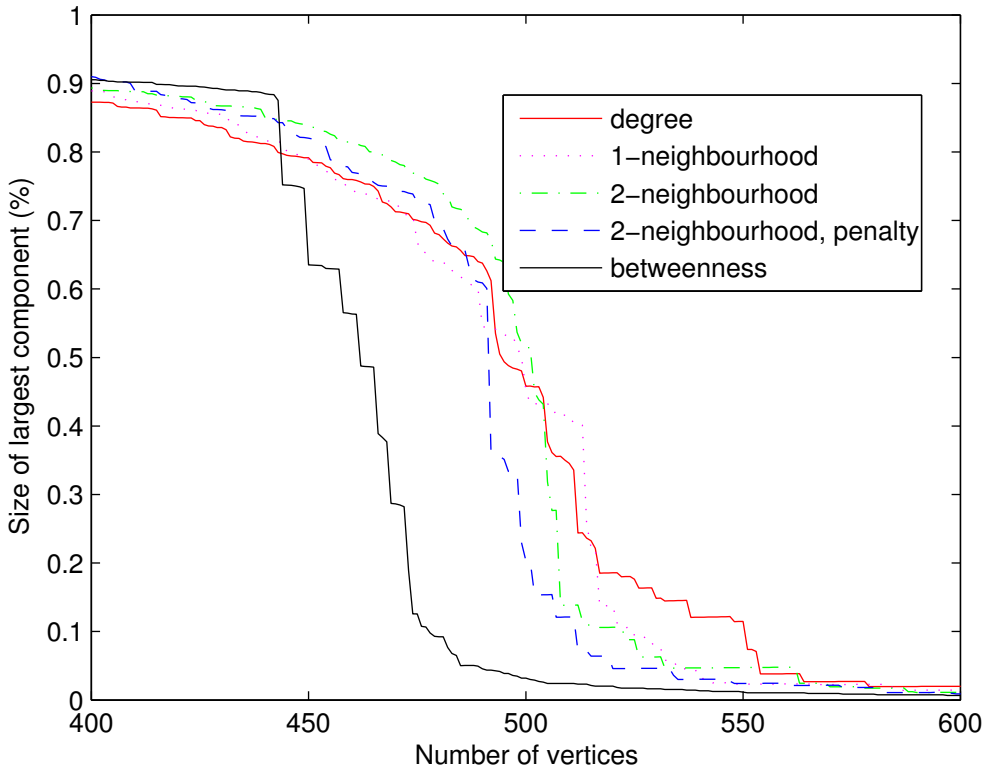

Figure 6: Size of largest connected component shown as percent of the initial size, as vertices are sequentially deleted. Detail of the transition.

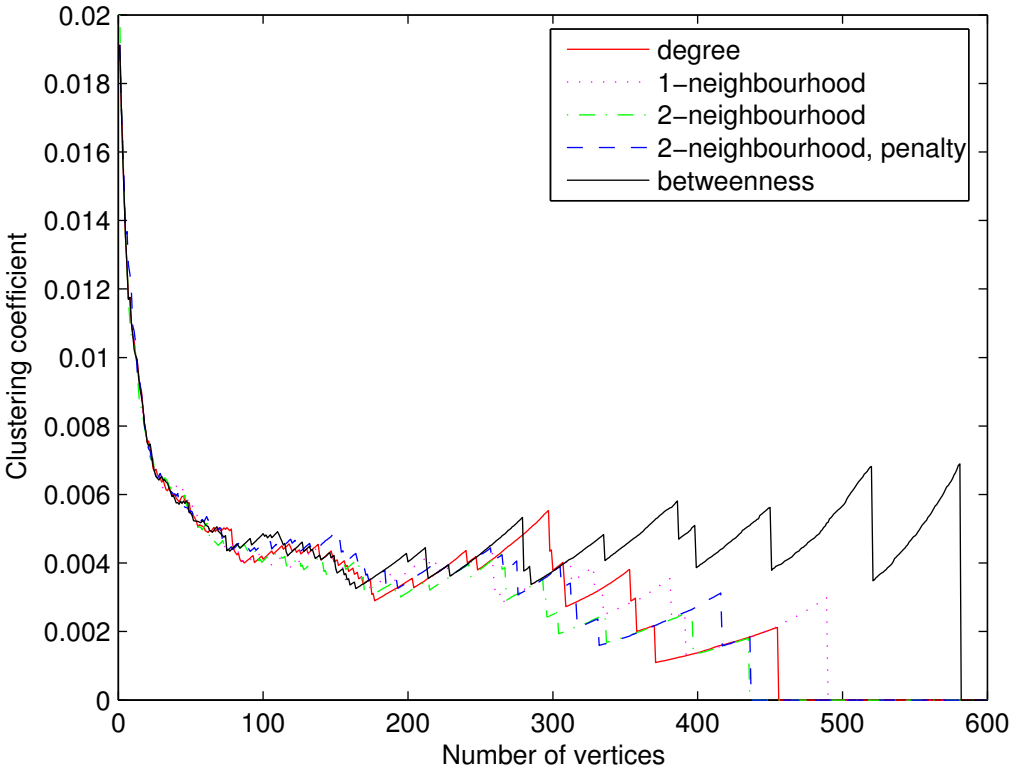

Figure 7: Clustering coefficient of network, as vertices are sequentially deleted. The results are the average of 50 experiments with networks of 1500 vertices.