Game Theoretic Approach for Discovering Vulnerable Links in Complex Networks

Mishkovski Igor, Sonja Filiposka, Sasho Gramatikov, Dimitar Trajanov and Ljupco Kocarev

Dept. of Computer Sciences
Faculty of Electrical Engineering and Information Technology
University Ss. Cyril and Methodious Skopje
Skopje, R. Macedonia
{igorm, filipos, saso.gramatikov, mite, lkocarev}@feit.ukim.edu.mk

***Abstract* – Complex networks have been an up-and-coming exciting field in the realm of interactions. With their widespread use appearing on the horizon it is ever more vital to be able to measure their vulnerability as a function of their topology. Precisely, discovering vulnerable links, disposed to attacks, can help in hardening these links and by that providing more secure and reliable network structure. This paper addresses the link vulnerability of different topologies of complex networks such as: random networks, geographic random networks, small world networks and scale-free networks. We introduce measure for vulnerability of complex networks, and prove by simulations that network vulnerability heavily depends on the network topology.**

***Index Terms* – Complex Networks, Vulnerability, Game Theory, Network Topology.**

I. Introduction

Shielding a link from malicious attacks is a key challenge to network security and management. Identifying and hardening the key links in a certain network will increase the network reliability but also it will decrease the amount of time needed to wield a reliable network. The emergence of terrorist attacks opened a new direction in the vulnerability analysis. Now the engineers must also be aware of intentional network attacks by the terrorists. The impact of these intentional link failures on the performances of the network depends on the routing strategy and the topology of the network. One way to deal with these intelligent attacks is to make the network more robust, i.e. to have more alternative routes. Thus, the original concept behind the Internet was that of a network that would withstand a nuclear attack [1]. But, no matter how much the network is robust there is always an open hole for the intelligent attackers.

In order to analyze these intelligent attacks on complex and man-made networks we use the game theoretic approach, proposed in [1].

Game theory introduces mathematical background for different analysis of the interactive processes for decision making. This theory enables tools that can leverage the prediction of what might happen in an environment where there is interaction between agents with conflict interests, i.e. non-cooperative environment. The traditional applications of the Game Theory try to find out the equilibrium point, i.e. set of strategies in which it is almost impossible for the individuals to change the current strategy.

This theory was introduced in [2] and its further development was due to the Nash Equilibrium concept in [3]. The games that were studied during the evolution of this theory were well defined mathematical objects. The games are consisted of players, a set of strategies, and specification of the profits for every combination of the strategies.

In the game presented in this study the players are the router, which seeks minimum cost paths for the packets, and a tester, whose aim is to maximize the trip cost. The solution of the game is the mixed strategy Nash equilibrium where the path selection probabilities are optimal for the router and the link failure probabilities are optimal for the tester. The overall vulnerability of the complex network is measured by the statistically expected trip-cost and the critical links for the network performance are indicated by the link failure probabilities.

There have been many uses of the proposed game theoretic approach by Bell in [1]. In [4] Bell quantifies the risk in transporting hazardous materials across a road network. In [5] using this approach the authors quantified the reliability of communication in mobile ad hoc network (MANET). In [6] authors propose a new vulnerability identification method in multicommodity stochastic networks.

This paper is an extension of the work by Bell [1] in the way that it analyzes the vulnerability of different topologies of complex networks. Thus, four generators were implemented for the different network topologies: random (Erdos Renyi - ER), geographic random (GER), small world (SW) and scale-free (SF). We introduce a new measure for vulnerability of the networks, and prove by simulations that the vulnerability of a network largely depends on its topology. Additionally the game theoretic approach was used in order to seek out the most vulnerable links in these topologies and to compare the topologies in terms of vulnerability.

The rest of the paper is as follows. In Section 2 we give a survey of the complex networks. In this survey we analyzed the: random networks, geographic random networks, small world networks and scale free networks. Section 3 presents the game theoretic tool which we used to measure the vulnerabilities of these types of networks. In Section 4 we give the topology-dependent properties of the observed networks. Section 5 presents the results obtained from the vulnerability analysis. Section 6 concludes the paper.

II. Complex Networks

Complex network is a complex graph-based structure made of nodes (which can be individuals, computers, web pages, power grid plants, organizations, cities, proteins in

the human body, etc.) that are connected by one or multiple types of interdependence (i.e. friendship, network links, power transport network, trade, roads, chemical reactions, etc.) These graphs or networks have certain properties which limit or enhance the ability to do things with them [22]. For example, small changes in the topology, shutting down only small number of links between the nodes, may lead to serious damage to the network capabilities.

A. Random Graphs

The Euler's introduction of the graph theory, was the initial step to uncover the properties of large, but ordered graphs.

Major breakthroughs are eight papers authored by Erdos and Renyi [7] laying down the foundation of the theory of random networks. They took on the challenge of explaining a very complex phenomenon by proposing an elegant mathematical answer to describe complex graphs within a single framework. By deliberately discarding the fact that different systems follow disparate rules in building their own networks, they follow the simplest solution: connect the nodes randomly.

Although Erdos and Renyi say that we need only one link per node to stay connected, real networks (like the worldwide net) are not only connected but are well beyond the threshold of one. Consequently, the networks in nature are very dense networks within which every node is navigable.

Start with a large number of isolated nodes. Then randomly add links between the nodes. If this continues, inevitably pairs of connected nodes will connect together forming clusters of several nodes. When enough links are added such that each node has an average of one link, a unique cluster emerges. That is, most of the nodes will be part of a single cluster such that, starting from any node, any other node can be reached navigating along the links between the nodes.

Although Erdos and Renyi say that we need only one link per node to stay linked, actual networks (like the worldwide net) are not only connected but are well beyond the threshold of one. Consequently, the networks in nature are very dense networks within which every node is navigable.

If the network is large, despite the links' completely random placement, almost all nodes will have approximately the same number of links. The result shows that the distribution of the number of links in a random graph is according to the Poisson distribution, which predicts that it is exponentially rare to find a node which deviates from the average by having considerably more or fewer links.

B. Geographic Random Networks

A geographic random network consists of set of points randomly scattered over a region according to some probability distribution, and these nodes are connected by an edge only if the distance between the nodes is less than specified value [16].

These types of networks are different from random networks in a way that they do not follow the property of independence or near-independence between the status of different edges. In geographic random networks triangular property is more realistic, which means if X_i is close to X_j, and X_j is close to X_k, then X_i will be fairly close to X_k.

With the advances in wireless communication technology geographic random topology is more and more present in the real complex and man-made networks. The ad hoc networks and mesh networks follow the properties of the geographic random graphs. This model can also represent a network of randomly placed sensors, each equipped with a limited communication capability.

In geographic random networks the average number of neighbors the node has, or the average node degree of the node depends on the transmission range of the node and the density of the nodes in the terrain, it can be calculated as:

$$k = \frac{N \cdot r^2 \cdot \pi}{a^2} \qquad (1)$$

where N is the number of nodes in the terrain, r is the transmission range of the nodes and a is the size of a square terrain.

C. Small World Networks

The random network theory has dominated network thinking since its introduction in 1959. In 1967, Stanley Milgram [8] turned the concept of "six degrees of separation" into a much celebrated, groundbreaking study on interconnectivity.

A repeated characteristic of complex networks is the small-world phenomenon, defined by the co-existence of two apparently contrary conditions:

(i) the number of intermediaries between any pair of nodes in the network is quite small - i.e. six-degrees of separation phenomenon and

(ii) the large local "cliquishness" or redundancy of the network - i.e., the large overlap of the circles of neighbors of two network neighbors. The latter property is typical of ordered lattices, while the former is typical of random graphs [9].

When one says that the network has "small world" topology, it means that almost every pair of nodes is connected by a path with an extremely small number of steps.

This kind of topology can be mostly seen in the social networks but there are also some technology, man-made and complex networks that have these characteristics. The Web falls in the same class of networks, where it has been shown that any document is on average only nineteen clicks away from any other [10]. Taken together these two networks suggest that behind the short observed distances of the enormous networks there is a fundamental property. This suspicion was later confirmed by subsequent discoveries which demonstrated that small separations are common in just about every network scientists have had a chance to study. The Internet, a network of hundreds of thousand of routers, has a separation of ten. The networks composed of proteins [11] with connections that indicate the physical interaction of the proteins exhibit small-world properties. Other examples are the road maps, electric power grids,

neural networks etc. The highly interconnected nature of these networks is the reason for this small separation.

If you consider a network in which the nodes have on average k links, there are however k^2 nodes two links away and roughly k^d nodes d links away. So if k is large, for even small values of d, the number of reachable nodes can become very large. If you have N nodes in the network, k^d must not exceed N. Thus, using $k^d=N$, a simple formula is obtained that works well for random networks, showing that the average separation follows the equation:

$$d = \frac{\log N}{\log k} \tag{2}$$

"Small worlds" are a generic property of networks in general. Most networks obey it since it is rooted in their structure. In Granovetter's paper [12] a new proposition for the structure of complex network emerges. The structure of the complex network around an arbitrary node is rather generic. In his view the graph is structured into highly connected clusters, or close-knit circles of nodes, in which every node has link to everybody else.

Watts [13] answered the question concerning the likelihood of forming clusters of nodes. To achieve this Watts and Strogatz introduced a quantity called the clustering coefficient. This coefficient tells how closely knit the circle of neighboring nodes is. A number close to 1.0 means that all neighbor nodes of one node are also neighbors with each other. Working on available networks, it has been shown that real networks like the network of mathematicians' co-authorship, or the collaboration graph of scientists are showing evidence of high clustering.

D. Scale-free Networks

Malcolm Gladwell's [14] conclusion has shown an altogether new property of complex networks: Connectors – nodes with an anomalously large number of links – are present in very diverse complex systems, ranging from the Internet to the cell. They dominate the structure of all networks in which they are present, making them look like small worlds. Indeed, with links to an unusually large number of nodes, hubs create short paths between any two nodes in the system.

Power laws mathematically formulate the fact that in most real networks the majority of nodes have only a few links and that these numerous tiny nodes coexist with a few big hubs, nodes with an anomalously high number of links. In a random network the peak of the node degree distribution implies that the vast majority of nodes have the same number of links and that nodes deviating from the average are extremely rare. Therefore, a random network has a characteristic scale in its node connectivity, embodied by the average node and fixed by the peak of the degree distribution. In contrast, the absence of a peak in a power law degree distribution implies that in a real network there is no such thing as a characteristic node. There is no intrinsic scale in these networks. This is why Albert Barabasi and his group described the networks with power law distribution as scale-free networks [10]. For scale-free networks the number of nodes with exactly k links follows a power law, each with a unique degree exponent that for most systems varies between two and three:

$$N(k) \sim k^{-\gamma} \quad \gamma=(2,3) \tag{3}$$

The first proposal for generation of scale-free networks is the Albert Barabasi model [15] which draws from the fact that scale-free topology is a natural consequence of the ever-expanding nature of real networking. Starting from two connected nodes, every time a new node is added to the network, it prefers to attach to the more connected nodes. The expansion of the network means that the early nodes have more time than the latecomers to acquire links. Thus growth offers a clear advantage to the senior nodes, making them richest in links.

After the first model appears making it possible to create a scale-free network using growth and preferential attachment, several additions to the model follow. An important addition to the model is the possibility for creating a competitive environment [10]. Here each node has certain fitness η, a quantitative measure of a node's ability to stay in front of the competition. The introduction of fitness changes what is considered attractive in a competitive environment. In the original model it is assumed that node's attractiveness is determined solely by its number of links. In a competitive environment, nodes with higher fitness are linked to more frequently. A simple way to incorporate fitness is to assume that preferential attachment is driven by the product of the node's fitness and the number of links it has. Later it is shown that the fitness distribution can lead to two cases: the rich get richer scale-free topology, and the-winner-takes-all network where only one huge hub exists and all nodes are connected to it.

There are many networks that obey the scale-free characteristics, such as: protein-protein interaction networks., the World Wide Web, semantic networks etc.

III. Link Vulnerability Analysis Using Game Theory

Game theoretic approach for measuring the vulnerability of stochastic networks was introduced in [1]. The players in the game are a "router" which seeks minimum cost paths throughout the network and a virtual tester which aim is to maximize the cost of the trip. The game is with mixed strategies, where the path selection probabilities are optimal for the router and the link failure probabilities are optimal for the tester. Also an overall measure, statistical – trip cost, for the vulnerability of the network is introduced. By using this approach and one can identify the critical links for the network performance.

The objective of the game is to seek links whose failure would damage the performance of a complex network the most.

In this game it assumed that one link can fail at a given time and each failure scenario corresponds to 1 failed link.

When a link fails, some penalty must be introduced. The queue on each link is assumed to be a (random arrivals/random service times/single server). The degree of saturation on link i equals p_i, giving an s-expected delay of:

$$d_i = \frac{p_i}{1 - p_i} \tag{4}$$

Initially, the failure penalty is assumed to be the same for all links and equal to 10 units. Thus, cost of the link *i* under failure scenario *j* is equal to:

$$c_{i,j} = \begin{cases} 10, if\ i = j, \\ d_i, otherwise \end{cases} \quad (5)$$

Because the cost of the link in the complex networks is traffic dependent MSA (Method of Successive Averages) algorithm, can be used [1]. The procedure is as follows.

Step 1: Link failure probabilities (q_j) for all scenarios are initialized to 1/nLinks.

Step 2: Link use probabilities (p_i) for all links are initializes to 0 and n ← 1.

Step 3: s-Expected link-costs are calculated as a function of link-use probabilities and the path with the minimum s-expected cost is sought (we used the Dijkstra algorithm [17]); x_i ← 1 if link *i* is on the shortest path, 0 otherwise.

Step 4: Update link use probabilities:

$p_i \leftarrow (1/n) \cdot x_i + (1-(1/n)) \cdot p_i$ for all links *i*.

Step 5: Find the *j* which maximizes $\sum_{i,j} p_i \cdot c_{i,j}(p_i)$, and y_j ← 1; for all scenarios, $k \neq j$, y_k ← 0.

Step 6: Update link failure probabilities:

$q_j \leftarrow (1/n) \cdot y_j + (1-(1/n)) \cdot q_j$, for all scenarios *j*.

Step 7: n ← n + 1; return to *Step 3* until satisfactory convergence is reached.

The vulnerability measure for the whole network, s-expected trip cost, can be calculated as:

$$C = \sum_{i,j} p_i \cdot c_{i,j} \cdot q_j \quad (6)$$

This algorithm is a heuristic one and there is no guarantee for convergence. And when it occurs is one of possibly many solutions to the problem. The convergence can be ensured using convergence criteria which measures the distance to a solution explicitly.

IV. Topology Dependent Properties of the Observed Complex Networks

The game theoretic approach used in [1] and also explained here was tested on four topologies of complex networks: random topology, geographic random topology, small world topology, and scale free topology. For the purpose of the simulations we implemented network generators, in Matlab [18], for each network topology. The number of nodes (*N*) in the network was 500 and the average node degrees (k_{av}) of the networks were close to 6.

The generator for the small world network was based on the Watts-Strogatz model with probability of reconnection of 0.1.

The generator for the Scale-Free networks was based on the BA model were at the beginning the network was consisted of 4 entirely connected nodes.

The generator for random networks was based on the Erdos-Renyi model. The nodes were randomly connected where the probability for a link between nodes *i* and *j* was:

$$p_{i,j} = \frac{k_{av}}{N-1} \quad (7)$$

The geographic random generator was invented by us. The algorithm is as follows. At first the nodes were randomly scattered along a $1m^2$ square terrain and their connectivity radius was calculated using (2) as:

$$r = \sqrt{\frac{k_{av}}{N \cdot \pi}} \quad (8)$$

All of the networks need to be connected, i.e. starting from any node; any other node can be reached navigating along the links between the nodes. For this purpose we used eigenanalysis [19]. We checked the connectivity of the networks by finding that the second eigen value of the Laplacian matrix was bigger than 0 [19]. In the case of the geographic random network generator if this is not the case the giant component was found and the nodes from the islands were again randomly scattered, this process last until the giant component is consisted of all the nodes in the network.

Table 1 gives the topology dependent properties of the observed network topologies, such as the average node degree [21], clustering coefficient [13] and normalized betweenness centrality [21], obtained by Ucinet [20].

Table 1 Topology dependent properties of the observed complex networks

Measure/Network	**ER**	**GER**	**SW**	**SF**
Average node degree	6.27	6.26	6.00	5.98
Clustering coefficient	0.014	0.627	0.447	0.055
Average normalized node betweenness	0.521	3.352	0.894	0.445

V. Vulnerability Analysis

The MSA algorithm and network generators were implemented in Matlab. In the simulation there were 1500 OD (origin-destination) commodity pairs. The nodes in commodity pairs were chosen randomly. The game specified used in the simulation allows the virtual network tester to fail only 1 link at a time. This is suitable where the probability of 2 or more concurrent failures is very small.

The first analysis was related to the vulnerability of the whole network. Figure 1 shows the convergence of the MSA method for the four topologies of complex networks. One can see that this method converges very fast despite the network topology.

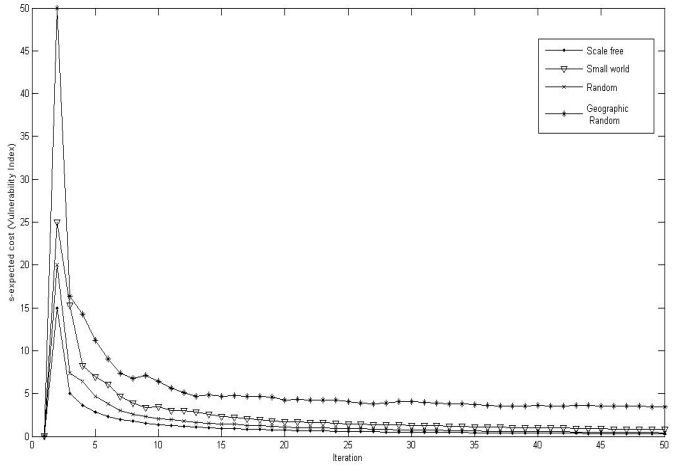

Fig. 1 Convergence of the Vulnerability index for the four topologies of complex networks

The value to which it converges as a function of the network topology is shown in fig. 2. As we can see the poorest performance gives the geographic random topology with index of 3.4189, then small world network (index: 0.8001), then the random network topology (index: 0.4279) and then scale free network topology (index: 0.3163).

In reality, these results mean that the scale free topology (for example the Internet network) is more resistant to link failure then network with geographic random topology, i.e. ad hoc network.

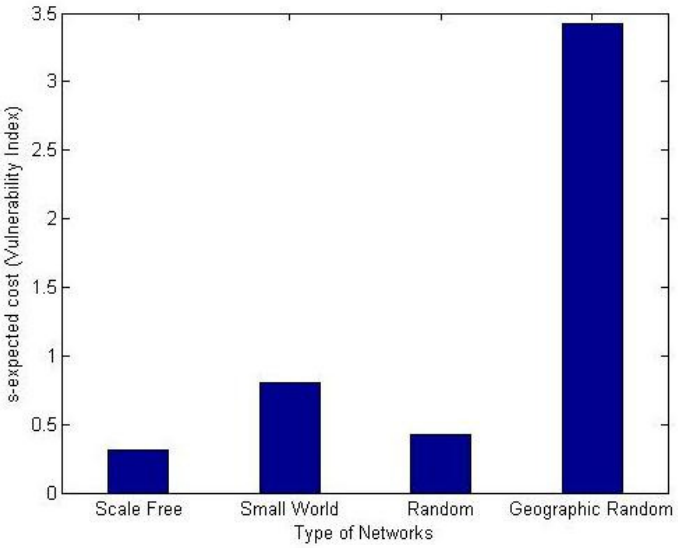

Fig. 2 Vulnerability index for the four topologies of complex networks

The second analysis consisted of finding the most vulnerable links in the networks, i.e. find the link with the greatest link failure probability. We seek out the weakest link for the most resistant topology, scale free, and the least resistant topology, geographic random. In fig. 3 the scale free network is shown with the weakest link (with bold line) between nodes 1 and 2. This is the link that connects the two biggest hubs in the network. The link failure probability of this link is around 0.95 and obviously it must be the most protected and robust link in the scale free network. Thus, its successful attack and removal will dramatically degrade the network performances.

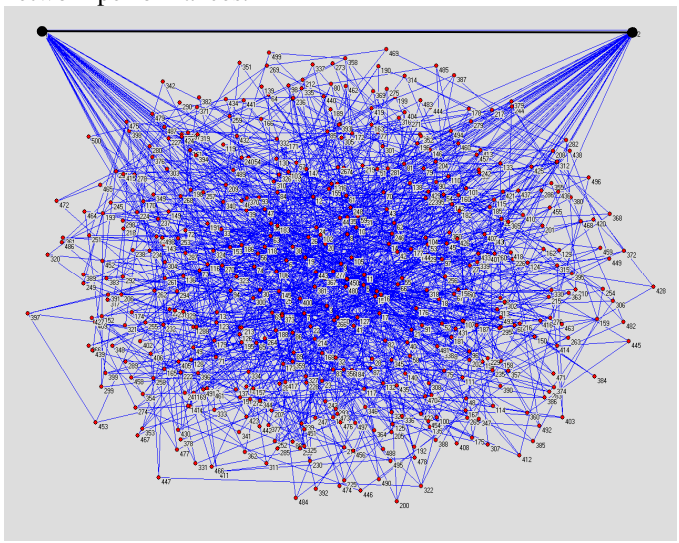

Fig. 3 Scale free network. The weakest link is the link connecting the two biggest hubs, node 1 and node 2

In fig. 4 the geographic random network is shown with the weakest link (with bold line) between the nodes 80 and 371. From the figure one can see that if this link is attacked and destroyed then the path length of its neighbors will increase dramatically. The link failure probability of this link is around 0.80 and it is much bigger than any other link in the geographic random network.

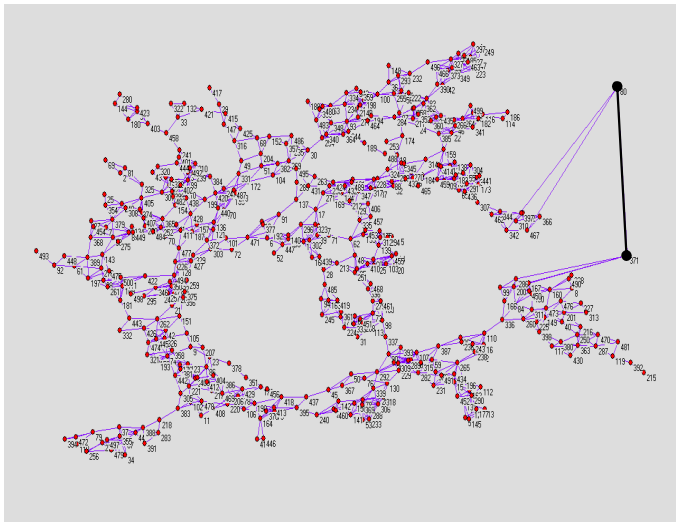

Fig. 4 Geographic random network. The weakest link is the link connecting the nodes 80 and 371

VI. Conclusion

The contribution of this work is twofold. Firstly, we use game theoretic approach to measure the vulnerability of complex networks with different topologies. We have studied vulnerability index in networks with four network topologies: random network, geographic random network, small-world network and scale free network. Our results show that the vulnerability of a network heavily depends on its topology. Concretely, we show that the scale free topology is the most resistant network topology to

intelligent link attacks and geographic random is the most vulnerable network to this kind of attacks. Secondly, using this approach we identify the weakest links in complex networks.

Our future work will be focused on using this approach for identifying vulnerability of different kind of real complex and other types of networks. Furthermore, we want to measure how the vulnerability of different network topologies changes after failure of certain nodes. These nodes can be chosen randomly or using some algorithm for choosing the most influent node, i.e. pageRank algorithm.

Another direction is to measure the vulnerability of the network by using the graph theory and network analysis to measure centrality of an edge, i.e. the edge betweenness.

REFERENCES

[1] M.G.H. Bell, *The use of game theory to measure the vulnerability of stochastic networks*. Reliability, IEEE Transactions on Volume 52, Issue 1, March 2003 Page(s): 63 – 68.

[2] John von Neumann, Oskar Morgenstern, *Theory of Games and Economic Behavior*. Princeton University Press, 1944.

[3] J. Nash, *Equilibrium point in n-person games*. Proceeding of the National Academy of Science, 36, 1950.

[4] M.G.H. Bell, *Mixed Route Strategies for the Risk-Averse Shipment of Hazardous Materials*, Netw. and Spat. Econ., vol. 6, no. 3, pp. 253-265, 2006.

[5] H. Karaa, and J.Y. Lau, *Game Theory Applications in Network Reliability* in Proc. Communications, 23rd Biennial Symposium, 2006, pp. 236-239.

[6] Satayapiwat, P.; Suksomboon, K.; Aswakul, C, *Vulnerability analysis in multicommodity stochastic networks by game theory*, ECTI-CON 2008, pp. 357-360.

[7] Michael Karonski and Adrzej Rucinski: The Origins of the Theory of Random Graphs, The Mathematics of Paul Erdos, Berlin, Springer, 1997

[8] Miligram S.: The small world problem, Psychology today 2, pp. 60-67,1967.

[9] L.A.N. Amaral and J.M. Ottino, Complex Networks, Augmenting the framework for the study of complex systems, Eur. Phys. J. B 28, 147-162, 2004.

[10] Albert Laszlo Barabasi: Linked, Penguin Group, London, May, 2003

[11] P. Bork et al. "Protein interaction networks from yeast to human", Current Opinion in Structural Biology, 14(3):292-299, 2004.

[12] Mark S. Granovetter: The Strength of Weak Ties: A Network Theory Revisited, Sociological Theory 1, 1983

[13] D. J. Watts: Small Worlds: The Dynamics of Networks between Order and Randomness, Princeton University Press, 2003

[14] Malcom Gladwell: The Tipping Point, New York, Little, Brown, 2000

[15] Albert B., Barabasi A.L.: Statistical mechanics of complex networks, Reviews of modern physics, Vol. 74, January 2002.

[16] Mathew Penrose: Random Geometric Graphs, Oxford University Press, New York, 2004

[17] E. W. Dijkstra: *A note on two problems in connexion with graphs*. In Numerische Mathematik, 1 (1959), S. 269–271.

[18] THE MATHWORKS, INC. 1997. MATLAB: *The Language of Technical Computing*. The MathWorks, Inc. Using MATLAB (Version 7).

[19] Keith Briggs, *Graph eigenvalues and connectivity,* july 2003.

[20] Borgatti, S., Everett, M. & Freeman, L. (1996a). *UCINET IV Version 1*.64. Natick, MA: Analytic Technologies.

[21] Freeman, L. C. (1979). *Centrality in social networks: Conceptual clarification. Social Networks*, 1(3), 215-239.

[22] Sonja Filiposka, Dimitar Trajanov and Aksenti Grnarov, Survey of Social Networking and Applications in Ad Hoc Networks, ETAI 2007, Ohrid, R.Macedonia, 2007.