TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 6232 ~ 6239
ISSN: 2302-4046

Efficiency Analysis of Scale-free Network Cascading Failures under Different Types of Attacks

Yuanni Liu*, Hong Tang, Guofeng Zhao, Yunpeng Xiao, Chuan Xu
The School of Communication and Information Engineering of ChongQing University of Posts and Telecommunications
No. 2, Chongwen Road,Nanan district, 400065, ChongQing, China
*Corresponding author, e-mail: liuyn@cqupt.edu.cn

Abstract

Network cascading failure can result in a congestion regime with degradation in the network performance. When cascading failure occurring, the netwok traffic will be rerouted to bypass malfunctioning routers, eventually leading to an avalanche of overloads on other routers that are not equipped to handle extra traffic, which Lots of failure models have been constructed to investigate how a small shock can trigger avalanches mechanisms affecting a considerable fraction of the network. In this paper, based on our AHP network cascading model, we have estimated how the efficiency will be affected when coefficients of K, S, T changed, we find the fact that the network efficiency of BA network is determined by its attacked types, and the efficiency is largely influenced by the attacked types of K and S, and under the same number of failure node, the efficiency under attacks of types T and I are relatively higher than that of the efficiency under attacks of types S and K , and the importance I is largely determined by the proportion of T, when the node failure number is equal, the higher proportion of T, the higher efficiency it is under I type attacks. We also tabled some proposals for reducing the damage that the networks suffered from the cascading failures.

Keywords: node degree distribution, complex network, attack, network efficiency

1. Introduction

In many real networks, one or a few nodes or edges failures, caused by random failures or deliberate attacks, will eventually leading to a considerable number of nodes or network crashes, which is called cascading failures, and there have been a lot of research about network attacks and the protection approaches[1]. The typical example is the North American Blackout. A great deal of efforts had been devoted into the improvement of network reliability, but large-scale cascading network failures occur constantly. Therefore, it is necessary to prevent and control the cascading failures in the networks, and explore the efficiency of the network under different types of attacks.

Current networks such as the World Wide Wed [1-2], the Internet , airplanes connection networks , and some biological systems, are different from random networks and all share the same property of having a power-law degree distribution $P(k) \sim k^{-\lambda}$ with an exponent λ that ranges between 2 and 3. Networks with power-law degree distribution have been named scale-free networks, which carry with them a well-recognized strength-tolerance of random failures. But they are particularly susceptible to failure of specific nodes that are highly connected and if such removals occur, the network will disintegrate rapidly[3-5]. In fact, although most failures emergence and dissolve locally, largely unnoticed by the rest of the world, a few trigger avalanche mechanisms that can have effects over the entire networks.

In this paper, we analyzed the efficiency of the BA network under different types of attacks as the alteration of node failure rate. Based on the initial experiment, we explored how the efficiency of BA network will be influenced by different types of attacks when the node importance *I* was calculated by *K, S, T* under four cases of coefficient.

The remainder of this paper is organized as follows: In the next section, we will introduce the related work; in section 3, we will show how to set up the elements coefficients in our AHP model [6]; in section 4, we will present the simulation results. Finally, the conclusion is drawn in section 5.

Received April 20, 2013; Revised July 16, 2013; Accepted July 25, 2013

2. Related work

Lots of interest has been focused on the studies of the consequences of different types of failures both on scale-free models and on real-world networks [7–9] to protect existing networks, and to locate the most critical nodes in order to reduce their criticality. The robust of networks to the removal of nodes or arcs, due either to random breakdowns or to intentional attacks, has been studied in [10–13], which have focused only on the static properties of the network showing that the removal of a group of nodes altogether can have important consequences. In [11], Paolo proposed a cascading network model to show how the network is influenced by network failures.

In Ref. [11], Crucitti proposed a cascading network model to show how the network is influenced by network failures. In his model each node is characterized by a given capacity C_i according to the tolerance parameter α , and every node has the same tolerance parameter α which is determined by the node importance. He also applied the model to the Italian electric power grid [6].

In Crucitti's cascading model, a generic communication/ transport network can be represented by a weighted undirected graph *G*, with *N* nodes and *M* arcs, and *G* is described by an $N \times N$ adjacency matrix $[e_{ij}]$. If there is an arc between node *i* and node *j*, the entry e_{ij} is the value, ranging in (0,1], attached to the arc; otherwise $e_{ij}=0$. The e_{ij} is a value of the path along the arc, and the smaller e_{ij} is, the longer it takes to exchange a unitary packet of information along the arc between *i* and *j*. Initially, at time $t=0$, for all the existing arcs, the $e_{ij}=1$, meaning that all the transmission lines are functioning equally. The model consists of a rule for the time evolution of $[e_{ij}]$ that mimics the dynamics of flow redistribution following the breakdown of a node.

In order to define the network efficiency [10], it assumes that any couple of nodes takes the most efficient path to communicate with each other, and the efficiency of a path is the so-called harmonic composition of the efficiencies of the component arcs. The ε_{ij} is defined as the efficiency of the most efficient path between *i* and *j*, and matrix $[\varepsilon_{ij}]$ is calculated by means of the algorithms used in Ref. [14]. With the knowledge of the path efficiency between any couple of the nodes *i* and *j*, we can calculate the average efficiency of the network by $E(G)=\sum_{i\neq j\in G}\varepsilon_{ij}/[N(N-1)]$, which is a measure of the performance of *G* at a given time.

C_i: the capacity defined as the maximum load that node can handle.

$L_i(t)$: the load on node *i* at time *t*, which is the total number of most efficient paths passing through *i* at time *t* [11].

The capacity C_i of node *i* is proportional to its initial load: *L(0)*, $C_i=\alpha L_i(0), i=1,2,...,N$, where $\alpha>1$ is the tolerance parameter of the network. The initial removal of a node, simulating the breakdown of an Internet router, starts the dynamics of redistribution of flows on the network. The initial removal of a node, simulating the breakdown of an Internet router, starts the dynamics of redistribution of flows on the network. The removal of a node will change the most efficient paths between the node pairs and consequently the redistribution of the loads, resulting overloads on some nodes. At each time *t* the efficiency of an arc is changed by the following iterative rule:

$$e_{ij}(t+1)=\begin{cases} e_{ij}(0)\dfrac{L_i(t)}{C_i}; L_i(t)>C_i \\ e_{ij}(0); L_i\le C_i \end{cases} \quad (1)$$

where *j* is all the first neighbors of *i*. Following rule (1), if at time *t* a node *i* was congested, the efficiency of all the edges passing through it will be reduced, as a result, the traffic (information) will take the new most efficient paths as the alternative one, which is a softer, and in some degree, a more realistic situation. But his model has some limits:

1) It will be a waste to assign every node with the same tolerance parameter α , when the network resource is finite.

2) Node importance should take account to multi-element instead of the single element of node degree.

In [6], we proposed an improved resource allocation model based on AHP[15] to analysis the efficiency of the network when it is under cascading failures caused by different types of attacks. The contributions of our model are mostly as follows:

1) We allocated nodes with different tolerance parameter α based on the node importance I to allocate the Ci
2) The importance I of a node is determined by three elements: node degree K; the number of the shortest paths S through a node; the number of the shortest paths T though the neighbor of a node.
3) We fixed every element a weight to compute I by AHP theory.

In our improved model, first we will assume that the network resource C_{all} is finite, which is a realistic assumption in the design of an infrastructure network, since the capacity cannot be infinitely large because it is limited by the cost. Secondly, we will assign every node with the capacity by its importance I_i determined by three elements, and in this way, different node will be characterized by different tolerance parameter α. The more important of the node, the more resource is allocated, which is consistent with the actual situation in the network. The correctness of the model has been proved in Crucitti's paper. Since we only altered the parameter α by changing the way of the resource allocation in this model, which will not affect the validity of the original cascading model.

In our model, we have considered three different elements to determine the importance I_i of a node *i*:

1) The degree K_i, i.e., the number of edges the node has.
2) S_i: the number of the *shortest* paths (over all pairs of nodes of the network) that pass through node *i*.
3) T_i: the number of the shortest paths that pass through all the neighbors of node *i*.

Weights of the three elements took account to determine the *Ii* are calculate by AHP. With the preference of alternative on each criterion, AHP can derive the appropriate weight for every element, and calculate the overall importance I_i of node *i* in Eq.(2).

$$I_i = w_1 K_i + w_2 S_i + w_3 T_i \tag{2}$$

Then the C_i of every node can be calculated by *Eq.* (2) and (3), and the tolerance parameter of a node is *Eq.* (4)

$$C_i = L_i(0) + \frac{I_i}{\sum_{j \in N} I_j}(C_{all} - \sum_j L_j(0)) \tag{3}$$

$$\alpha_i = \frac{C_i}{L_i(0)} \tag{4}$$

Based on our AHP network cascading model, we have investigated how the tolerance parameter α influenced the efficiency in BA and ER networks, and draw conclusions as follows:

(1) In order to protect network from random attack, the nodes should be distributed with different tolerance parameter α.
(2) As the real network follows the power-law in BA network, assign different tolerance parameter α to the node may improve the efficiency of the network.
(3) The importance of node *i* in Ref. [6] is calculated by $I_i = 0.1047K_i + 0.6370S_i + 0.2583T_i$, and the simulation results shows that the maximum damage to the network is the deliberately attack of node removal based on type *K*, which is larger than the types of *T* and *S*. In addition, under the node removal based on *I*, the network efficiency will be higher than that of Crucitti's model if only the weight of *K* will not be equal to 1. Therefore, in our model, the network efficiency affected by the change of the elements' weights will be always higher than that of the original model.

In this paper, we will find how the efficiency will be affected when coefficients of *K, S, T* changed

3. Computing the coefficients

We have described the AHP configuration process in [6], in this section , first we will give a short description again, and then we will set up four groups of different parameters for the correspondent *K*, *S*, *T* to calculate the node importance *I*.

3.1 The AHP configuration process

AHP is a well-studied, widely-applied technique in multi-criteria decision analysis [15], a field in decision theory. According to the decision maker's preferences with regard to individual element [15], the AHP can provide a simple, yet systematic way to find the overall best weight for all elements.

First, AHP will model the decision problem as a decision hierarchy. Then, it will perform pair-wise comparisons of all elements, and it will specific the preference of one element over the other using a number for each comparison. The scale from 1 to 9 has proven to be the most appropriate [15], in which, when comparing criteria *r* to *q*, 1 means *r* and *q* are equally preferred, 3 means weak preference for *r* over *q*, 5means strong preference, 7 means demonstrated (very strong)preference, 9 means extreme preference. The inverse values1/3, 1/5, 1/7 and 1/9 are used in the reverse order of the comparison (*q* vs. *r*). Intermediate values (2, 4, 6, 8) may be used when compromise is in order. As we do in Table 1, where elements are compared based on their importance.

Although the entire matrix contains of 9 preferences, to computed the *I*, we only need to specify 3 of them–'*K vs. S*', '*K vs. T* ', and '*S vs. T* '. the weights of all elements, which are computed from the principal eigenvector of the preference matrix. With the preference of alternative on each criterion, AHP can derive the appropriate weight for every element, and calculate the overall importance I_i of node *i* in Eq.(2).

Then the *Ci* of every node can be calculated by Eqs. (2) and (3), and the tolerance parameter of a node is Eq. (4).

3.2 Computing the coefficients

In order to find out how the alteration of the coefficients will affect the network efficiency under different types of attacks, first, we will set *K*, *S*, *T* with equal importance to determine the node importance *I* as in comparison matrix 1.

Table 1. Comparison matrix 1

Elements	*K*	*S*	*T*	weights
K	1	1	1	0.3333
S	1	1	1	0.3333
T	1	1	1	0.3333

In comparison matrix 1, we assume that all the elements are equally preferred, so the correspondent value of *I* can be calculated as I=0.3333K+0.3333S+0.3333T.

Second, we make *K* is the most important element to determine the node importance *I*, and *S* is the less important element, while *T* is the lest important element. The correspondent value are set in comparison matrix 2.

In comparison matrix 2, first, we weakly prefer *K* over *S*, which means the value of *K* vs. *S* is 3, and *S* vs. *K* is 1/3. In addition, we strongly prefers *S* over *T*, which means *S* vs. *T* is 5, and *T* vs. *S* is 1/5. Further more we extremely prefer *K* over *T*, which means *K* vs. *T* is 9, and *T* vs. *K* is 1/9. Finally, the node importance can be calculated as

I=0.6716K+0.2654S+0.0629T.

Table 2. Comparison matrix 2

Elements	*K*	*S*	*T*	weights
K	1	3	9	0.6716
S	1/3	1	5	0.2654
T	1/9	1/5	1	0.0629

Similarly, we make *S* is the most important element to determine the node importance *I*, and *T* is less important element, while *K* is the lest important element. The correspondent values are set in comparison matrix 3.

In comparison matrix 3, first, we extremely prefer *S* over *K*, which means the value of *S* vs. *K* is 9, and *K* vs. *S* is 1/9. In addition, we strongly prefers *S* over *T*, which means *S* vs. *T* is 5, and *T* vs. *K* is 1/5. Further more we weakly prefer *T* over *K*, which means *T* vs. *K* is 3, and *K* vs. *T* is 1/3. Finally, the node importance can be calculated as I=0.0627K+0.7596S+0.1777T.

Table3. Comparison matrix 3

Elements	*K*	*S*	*T*	weights
K	1	1/9	1/3	0.0704
S	9	1	5	0.7514
T	3	1/5	1	0.1782

Finally, we make *T* is the most important element to determine the node importance *I*, and *K* is the less important element, while *S* is the lest important element. The correspondent value are set in comparison matrix 4.

In comparison matrix 4, first, we weakly prefer *K* over *S*, which means the value of *K* vs. *S* is 3, and *S* vs. *K* is 1/3. In addition, we strongly prefers *T* over *K*, which means *T* vs. *K* is 5, and *S* vs. *T* is 1/5. Further more we strongly prefer *T* over *S*, which means *T* vs. *S* is 7, and *S* vs. *T* is 1/7. Finally, the node importance can be calculated as I=0.1884K+0.0810S+0.7306T

Table 4. Comparison matrix 4

Elements	*K*	*S*	*T*	weights
K	1	3	1/5	0.1884
S	1/3	1	1/7	0.0810
T	5	7	1	0.7306

The reason we set the element values in comparison matrix 1 to 4 is that , in comparison matrix 1 all the element are equal importance, which is a reference to the values in comparisons of 2, 3, 4. In comparisons 2, 3, 4, the importance of *K*, *S*, *T* are all changed in different ranges from small values to large values.

In this way, we set four different groups of parameters to determine the weights in Eq.2 to compute four different cases of the node importance I_i*s* as follows:

case1: I=0.3333K+0.3333S+0.3333T, case2: I=0.6716K+0.2654S+0.0629T,
case3: I=0.0627K+0.7596S+0.1777T, case4: I=0.1884K+0.0810S+0.7306T,

In order to find out how the efficiency will be affected when coefficients of *K*, *S*, *T* changed, we will estimate the network efficiencies under different types of attacks correspondent to the four cases of I_i in section IV.

4. Simulation results

In this section, we will find out how the efficiency will be affected when coefficients of *K, S, T* altered. Our Scale-free network topologies, i.e., graphs with an algebraic distribution of degree $P(k) \sim k^{-\lambda}$, with $\lambda = 3$ [16], is generated artificially according to the BA model [2]. In both cases we have constructed networks with N=500, and M=1494. In our experiments, we set $C_{all} = 1.3\sum_i L_i(0)$. In the simulation we will observe the ratio of E/E(0) to explore the efficiency influenced by different types of failure, where *E(0)* is the initial efficiency *E(G)* of the network, and *E* is the efficiency of the network on a certain point. We will find out how the efficiency will be affected when coefficients of *K*, *S*, *T* changed.

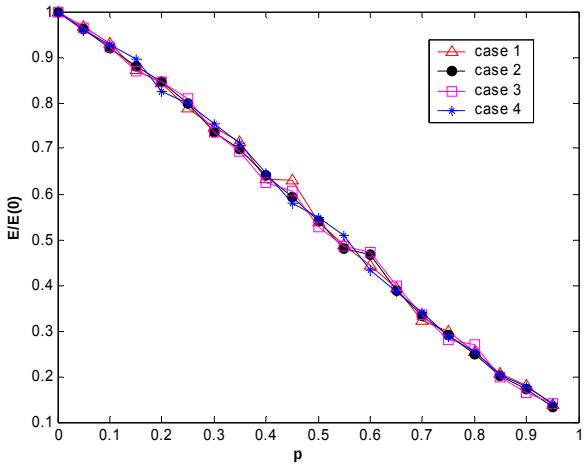
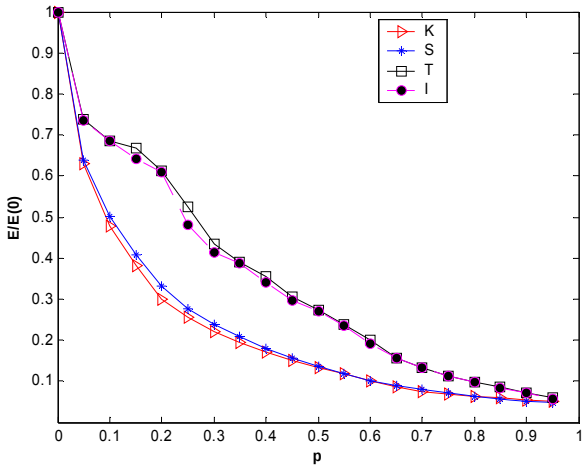


Figure.1 The efficiency of BA network under random failures in four cases

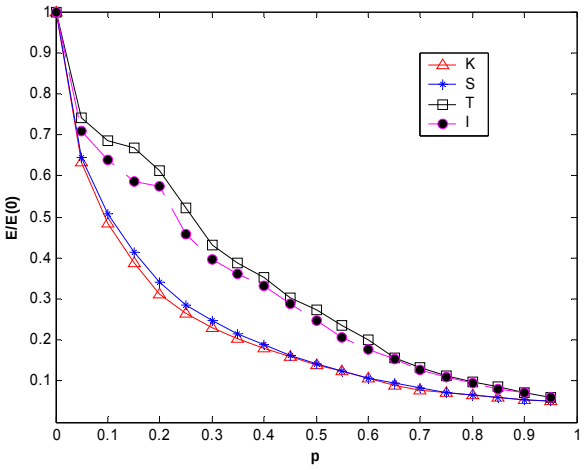
In Figure. 1, we show the results of the simulations of *E/E(0)* as the functions of *p* under random failures in four cases. In different cases, the *I* is calculated by different coefficients of the elements *K, S, T* in the four cases that we have set in section 3. In Figure.1, the four curves are overlapped, which means that the efficiency of the BA network is less influenced by the way of the resource allocation under random failures.

Figure.2 (a), (b), (c), (d) are the efficiency of BA network under different types of deliberate attacks in cases 1, case 2, case 3, and case 4 respectively. From Fig.2 we can draw the conclusion that:

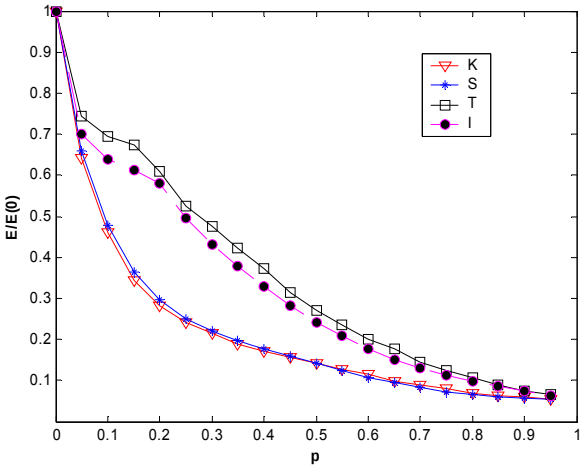
1) The efficiency of the network is reduced as the increment of the failure node number. When the failure node number is equal, the efficiency of the network under the types of *K* and *S* attacks are lower than that of the types of *T* and *I*, and the type of *I* is lower than that of the type *T*.
2) In the case of the same number of failure node, the greater proportion of *I* decided by *T*, the higher efficiency of the network than the same situation under *I* type attacks, and the closer of the curves from *T* type attack to *I* type attack.
3) Efficiency of the BA network with different tolerance parameter is determined by its attacked types, in which, the efficiency is largely influenced by the attacked types of *K* and *S*, and under the same proportion of the failure node, is lower the attacked types of *T* and *I*. While, *I* is mostly determined by the proportion of *T*, in the same situation, the higher proportion of *T*, the higher efficiency under *I* type attacks.



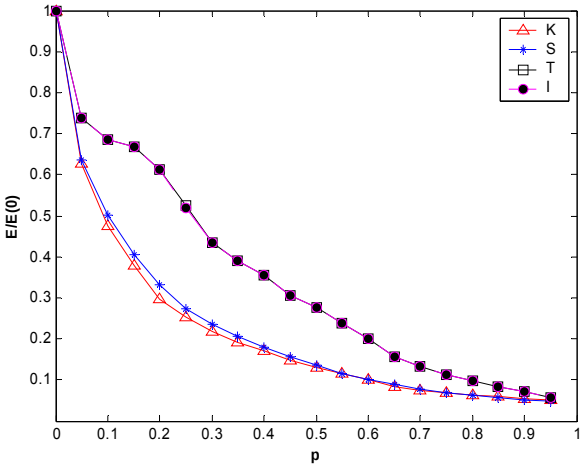
(a) The efficiency of BA network under different types of deliberate attacks in case 1



(b) The efficiency of BA network under different types of deliberate attacks in case 2



(c) The efficiency of BA network under different types of deliberate attack in case 3



(d) The efficiency of BA network under different types of deliberate attack in case 4

Figure.2 The efficiency of BA network under different types of deliberate attacks in four cases

5. Conclusion

In this paper, we have estimated how the efficiency will be affected when coefficients of *K*, *S*, *T* changed. The conclusions can be derived as:

1) The removals based on *K* and *S* influenced the network efficiency are larger than the case, where the removals based on the *T* and *I*. As a result, in order to improve network efficiency, we can reduce the weights of these two elements took account in determining the importance of a node.
2) The efficiency of BA network is determined by its attacked types. The efficiency is largely influenced by the attacked types of *K* and *S*, and under the same number of failure node, the efficiency under attacks of types *T* and *I* are relatively higher. The importance *I* is largely determined by the proportion of *T*, when the node failure number is equal, the higher proportion of *T*, the higher efficiency it is under *I* type attacks.

Acknowledgements

This work was supported by the CQUPT Dr. Start-up Fund Research (A2011-48), Natural ScienceFoundation of CQUPT(A2012-83), National Program on Key Basic Project(973 program: 2012CB315803), National Natural Science Foundation of China(641040044, 60873079), the Natural Science Foundation of ChongQing(CSTC2009BA2089).

References

[1] Zhang Yu. Computer Network Attack Detection Based On Quantum Pso And Relevance Vector Machine. *AISS*.2012; 4(5): 268-273.
[2] Réka Albert, Hawoong Jeong & Albert-László Barabási. Diameter of the world-wide-web.*Nature*.1999; 401(6749):130-131.
[3] Albert-László Barabási, Réka Albert. Emergence of scaling in random networks.*Science.*1999; 286(5439): 509-512
[4] Musirin Serwan. *Cascading collapse assessment considering hidden failure*. IEEE Proceedings on Informatics and Computational Intelligence(ICI) , Bandung, 2011: 318-323.
[5] Wang Guo-hong, Xing Rui, Tang Liyan.*The risk of cascading breakdown in industrial cluster innovation networks: a complex networks perspective*. IEEE Proceeding on Management Science and Engineering. Melbourne, VIC, 2010: 1445-1455.
[6] Yuanni Liu, XinLi, Shanzhi Chen, Zhen Qin. Model for Cascading Network Failures Based on the Nodes with Different Tolerance Parameter. *The journal of China universities of posts and telecommunications*. 2011; 18,(5):95-101.
[7] Holme Petter, Beom Jun Kim. Edge overload breakdown in evolving networks. *Phys. Rev.E*. 2002; 66(3) :036119.
[8] Adilson E. Motter, Ying cheng Lai. Cascade-based attacks on complex networks. *Phys. Rev.E.* 2002; 66(6):065102.
[9] Crucitti Paolo, Latora, Vito, Marchiori. Model for cascading failures in complex networks. *Phys. Rev.E*. 2002; 69(4), pp.045104.
[10] Holme Petter, Beom Jun Kim, Chang No Yoon, Seung Kee Han. Attack vulnerability of complex networks. *Phys. Rev.E*. 2002; 65(5):056109.
[11] Paolo Crucitti, Vito Latora, Massimo Marchiori, Andrea Rapisarda. Efficiency of Scale-Free Networks: Error and Attack Tolerance. *Physica A*. 2003, 320(11):622-642.
[12] Girvan Michelle, M.E.J Newman. *Community structure in social and biological networks*. Proceedings of the National Academy of Science. 2002; 99:7821-7826.
[13] James E. Smith. Characterizing computer performance with a single number.*Commun. ACM*. 1988; 31(10):1202-1206.
[14] Erdo¨s Paul, Alfréd Rényi.On random graphs I. *Mathematics*.1959; 11(6):290-297.
[15] Thomas L. Saaty.The Fundamentals of Decision Making and Priority Theory with the Analytic Hierarchy Process. AHP Series.Fourth Edition. New York: RWS Publications, 2000.
[16] Georgos Siganos, Faloutsos Michalis, Faloutsos Petros, Faloutsos Christors. Power laws and the AS-level Internet topology, Networking. *IEEE/ACM Transactions on Networking*.2003; 11(4): 514-524.