Degree based attacks and defense strategies in complex networks

Aviv Yehezkel[1] and Reuven Cohen[1,*]

[1] *Department of Mathematics, Bar-Ilan University, Ramat-Gan 52900, Israel*

(Dated: November 25, 2012)

We study the stability of random scale-free networks to degree dependent attacks. We present analytical and numerical results to compute the critical fraction p_c of nodes that need to be removed for destroying the network under this attack for different attack parameters. We study the effect of different defense strategies, based on the addition of a constant number of links on network robustness. We test defense strategies based on adding links to either low degree, mid degree or high degree nodes. We find using analytical results and simulations that the mid degree nodes defense strategy leads to the largest improvement to the network robustness against degree based attacks. We also test these defense strategies on an Internet AS map and obtain similar results.

PACS numbers: 89.75.Hc, 87.23.Ge

In the past few years the study of large connected network systems has become more and more popular [1–5]. The understanding that many real world networks, affecting almost every part of the modern life, from brain and biological systems to social systems, power grids and the Internet, the WWW and etc., are represented by new classes of random networks has led to breakdown of standard theoretical models and inspired a new area of research [1–15, 17, 20]. Many of these networks are characterized by a power law distribution in their nodes degrees. Such networks are constructed by nodes connected with links where the degree distribution $P(k)$ which is the probability that a node will have k links is:

$$P(k) \sim k^{-\gamma} \,. \tag{1}$$

where γ is usually between $2 < \gamma < 3$.

The scale-free character of these networks, represented by having no characteristic degree per node, has led to numerous unexpected results in many different properties that are very different from the results in lattice models and even ER random graphs [1–7]. One important property that was studied is the robustness and resilience of such networks under external attack or random failure. In other words, the stability of such networks and their ability to withstand progressive damage caused by successive removals or failures of nodes [11]. This research is aimed at getting a better understanding of the vulnerability of such networks in order to make them more secure and robust.

We denote p_c to be the critical fraction of removed nodes needed for destroying the integrity of the network and its topology. In general it has been observed that complex networks show great stability even against large number of repeated attacks or failures. For example, it was shown that $p_c = 1$ for random failures, i.e., in order to destroy the network we have to practically remove all the nodes [12]. On the other hand, one of the most efficient attacks that were studied is an intentional attack where only the nodes with the highest degrees need to be removed in order to destroy the network. In such attacks, removing only a small fraction p of the nodes is sufficient to destroy the network.

Although the intentional attack strategy is one of the most efficient attack strategies, in most of the cases we cannot use this attack as it requires complete knowledge and understanding of the network topology in order to be able to identify the highest connected nodes to be removed. In many realistic cases this required information is not available, does not exist or is not accessible. Sometimes, only partial knowledge of the network topology and its nodes is available. Therefore we want to find other attack strategies whose implementation does not rely on full knowledge of the network topology and its nodes connectivity, thus we will present attack strategy that only require partial knowledge of the network topology.

In this letter we consider scale-free networks where the robustness of each node depends on its degree, so the probability of damaging a node by some attack or failure depends on the degree k of the node. We study the stability of those networks to degree dependent attacks where the probability to remove node depends on its degree. This method demands that nodes with higher degrees will have higher probability to be removed.

Examples for such networks where higher degree nodes are less robust can be found in a variety of fields. In social networks, the most connected members of the group that have more links to other members are more visible and therefore have a higher probability to be attacked and removed from the group. Another example is of network traffic. High loads on highly connected nodes [13], which make them more vulnerable to attacks or failures. In some cases breakdowns are due to cascades of failures caused by the dynamics of damage spreading [14]. In computer networks many breakdowns are caused by congestion building [15].

We will also study the effect of different defense strategies, based on the addition of a constant number of links on network robustness against this attack and we will find that the mid degree nodes are very important to the

* reuven@math.biu.ac.il

robustness of the network and in order to make the network more robust under such attacks we have to defend the mid degree nodes.

This work applies for the attack of many real world networks where there is no full knowledge of the nodes degrees that is required for intentional attack, and especially to immunization strategies where the high degree nodes are usually not known [17].

We assign to each node a value $\pi(k_i)$ which represents the probability that a node i with degree k_i in network with N nodes will be removed from the network and become inactive.

$$\pi(k_i) = \frac{\alpha k_i}{K_{max}}\,, \quad 0 < \alpha \leq 1\,. \tag{2}$$

where $K_{max} \sim mN^{1/(\gamma-1)}$ is the maximal degree of a node in the network [12]. The parameter α can be thought to represent either the level of vulnerability of nodes, or the level of knowledge of the node degrees possessed by the attacker. We study this new attack with both numerical simulations and analytical treatment. First we will determine the critical percolation threshold p_c as a function of the network parameters γ, N and m and the attack parameter α.

In the numerical treatment we first build the network for a given γ. We fix the size of the network N, and the minimal degree m and assign the degree k for each node by using the power law distribution $P(k) \sim k^{-\gamma}$. In the numerical simulations and the analytical treatment we use the configuration model for the network construction where no correlations exist between the degrees of neighboring nodes. We then link random pairs of nodes that have not been directly connected to each other already and have not reached their given degree. We repeat this process until the entire network is built.

To find the percolation threshold p_c, we go over the network and remove each node with probability $\pi(k_i)$. For each removed node, all its links are cut and removed. After each removal we calculate the two moments $\langle k^2\rangle$ and $\langle k\rangle$ and divide them to get $\kappa = \frac{\langle k^2\rangle}{\langle k\rangle}$. if $\kappa \geq 2$, then a giant component still exist in the network [16]. Once we finished performing this procedure over the entire network we start again scanning the network implementing the same procedure until the network has no longer any giant component left. We continue implementing this procedure over the network until $\kappa < 2$. We then repeat this procedure for different γ values and for different N values and obtain p_c.

In order to analytically find the percolation threshold, let us denote by K_i the maximal degree of the network at the ith scan. We assume to know the maximal degree of the network at each scan. After d scans the probability that a node of initial degree k will still be functional is:

$$\rho(k,d) = \prod_{i=1}^{d}\left(1 - \frac{\alpha k}{K_i}\right). \tag{3}$$

Let us assume that we had to perform d_c scans until we managed to destroy the network. That is, until the percolation threshold is reached, and the giant component disappears. Lets also assume that during the last scan we went over only a fraction c of the network until it was destroyed and the process ended. In order to compute the probability that a node of degree k will still be functional at the end of this process we have to consider the two different cases stating whether this node was in the c fraction or not. We get the following result:

$$\begin{aligned}\rho(k) &= c\prod_{i=1}^{d_c}\left(1 - \frac{\alpha k}{K_i}\right) + (1-c)\prod_{i=1}^{d_c-1}\left(1 - \frac{\alpha k}{K_i}\right)\\ &= c\rho(k,d_c) + (1-c)\rho(k,d_c-1)\,.\end{aligned} \tag{4}$$

The condition for the existence of a giant component after d scans is as follows: if a node is reached through a link with probability $kP(k)/\langle k\rangle$ and has not been removed with probability $\rho(k,d)$ and the average number of outgoing links per site is larger than 1, then a giant component will exist [5]. Since a node is reached through a link, it then has $k-1$ other links that can be traversed.

$$\sum_{k=m}^{K}\frac{P(k)k(k-1)}{\langle k\rangle}\rho(k,d) = 1\,. \tag{5}$$

Now we can numerically solve Eq. (5) to calculate d_c and substitutes this for calculating the critical threshold of removed nodes:

$$p_c = \sum_{k=m}^{K} P(k)(1-\rho(k))\,. \tag{6}$$

From Eq. (6) we can compute the percolation threshold p_c for a given network with size N and exponent γ as a function of the attack parameter α. The lines in Fig. 1 represent the solutions of Eqs. (5) and (6), and we can see that they are in good agreement with the simulations. We can also see that for $\gamma < 3$, p_c becomes smaller than 1 already for very small α values and decays with growing α. According to Fig. 1 only a very small fraction of the network has to be removed for destroying the network, even without full knowledge of the network topology and connectivity.

We now turn to study the best way to defend against such attacks. To study the effect of different defense strategies on the robustness and stability of a network we consider different defense strategies based on addition of a constant number of links to different groups of nodes in the network: low degree, mid degree and high degree nodes, as well as very high degree nodes.

We test these strategies and find which defense strategy is more efficient and leads to the highest network robustness improvement against degree based attacks. In other words, we want to be able to identify the crucial elements of the network that determine the robustness of the network under degree based attacks.

Let us define "cut of the network" $C(i,j)$ to include all the nodes with initial degree $i \leq k \leq j$. We want to

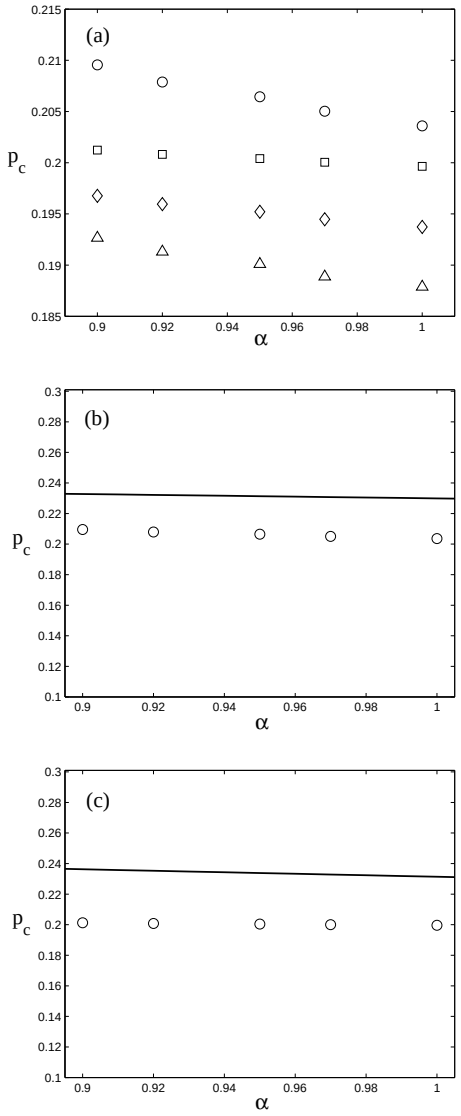

FIG. 1. (a) Values of p_c vs α for different γ values: $\gamma = 2.5$ ($\circ$), $\gamma = 2.3$ ($\Box$), $\gamma = 2.7$ ($\diamond$) and $\gamma = 2.1$ ($\triangle$). Symbols represent simulation data for $N = 10^5$ nodes and lower cutoff $m = 1$ averaged over 100 different network realizations. (b) Values of p_c vs α for $\gamma = 2.5$. Symbols represent simulation data for $N = 10^5$ nodes and lower cutoff $m = 1$ averaged over 100 different network realizations. Solid line is the theoretical prediction for the given parameters (Eqs. (5) and (6)). (c) Same as (b), with $\gamma = 2.3$.

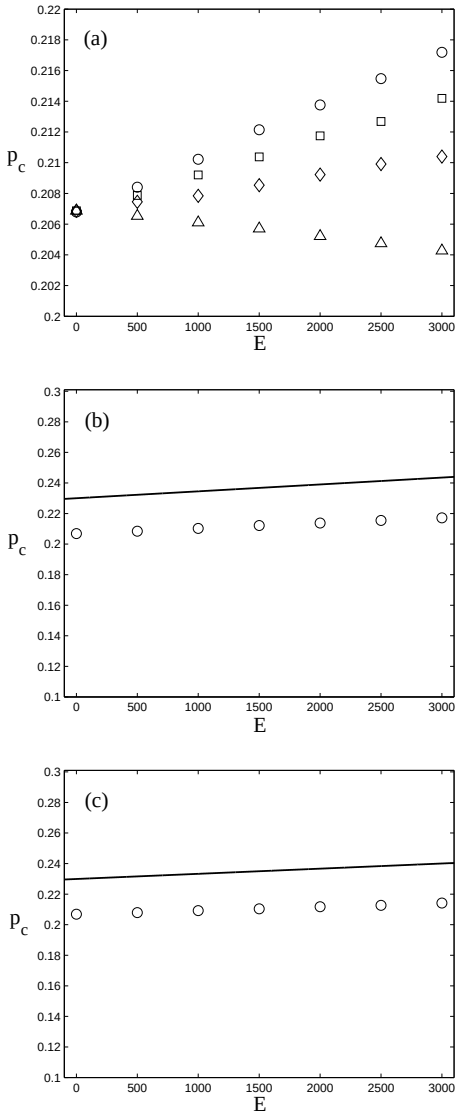

FIG. 2. (a) Values of p_c vs E for different defense strategies according to simulation data where $N = 10^5$ nodes with $\gamma = 2.5$, attack parameter $\alpha = 1$ and lower cutoff $m = 1$ from 100 different network realizations. The symbols represent low degree ($1 \leq k \leq 2$, $\Box$), mid degree ($3 \leq k \leq 6$, $\circ$), high degree ($7 \leq k \leq 10$, $\diamond$) and very high degree ($10 \leq k \leq K_{max}$, $\triangle$) defense strategies . (b) Values of p_c vs E for mid degree nodes defense strategy for nodes with initial degree $3 \leq k \leq 6$. Symbols represent simulation data for $N = 10^5$ nodes, $\gamma = 2.5$, attack parameter $\alpha = 1$ and lower cutoff $m = 1$ averaged over 100 different network realizations. Solid line is the theoretical prediction for the given parameters (Eqs. (5) and (6)). (c) Same as (b), for low degree nodes defense strategy for nodes with initial degree $1 \leq k \leq 2$.

defend the network with nodes from this cut by inserting E new links between nodes from $C(i,j)$ only. We use the same procedure as before and randomly select pairs of nodes and link them. Let us define $P(i,j)$ to be the proportional size of the cut $C(i,j)$ in the network:

$$P(i,j) = \sum_{k=i}^{j} P(k)\,. \tag{7}$$

The probability that a node from $C(i,j)$ will be chosen and get an additional link is $\frac{2E}{P(i,j)N}$ so after d scans the probability that a node of initial degree $i \leq k \leq j$ from the cut $C(i,j)$ will still be functional is now:

$$\rho(k,d) = \frac{2E\rho(k+1,d)}{P(i,j)N} + \left(1 - \frac{2E}{P(i,j)N}\right)\rho(k,d)\,. \tag{8}$$

We can now use the same method that was described before to find the number of scans d_c needed to attack and destroy the network and then find the fraction of removed nodes p_c. Notice that by the method suggested here some small correlation is induced between neighboring nodes' degrees, due to favoring connections between similar degree nodes. Thus, the criterion of $\kappa > 2$ is only a close approximation to the actual percolation threshold. A more exact, but more complicated solution can be obtained using the methods of [18, 19]. The results are also verified using direct evaluation of the giant component size (See Fig. 5).

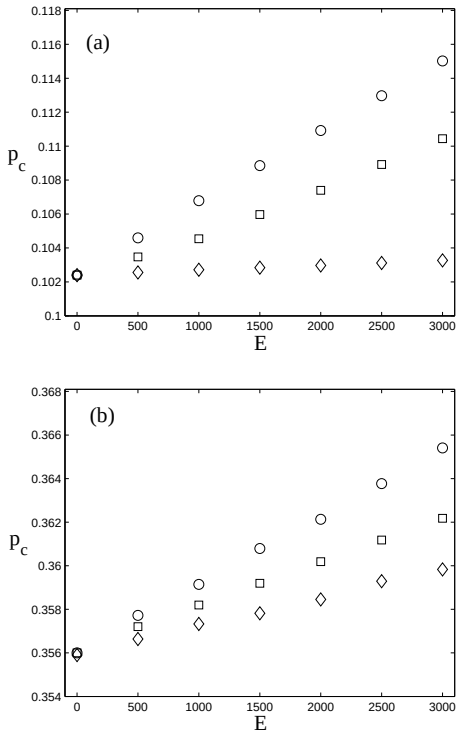

FIG. 3. Values of p_c vs E for different defense strategies according to simulation data where $N = 10^5$ nodes with $\gamma = 2.5$, attack parameter $\alpha = 1$ and lower cutoff $m = 1$ from 100 different network realizations. The symbols represent low degree ($1 \leq k \leq 2$, $\square$), mid degree ($3 \leq k \leq 6$, $\circ$) and high degree ($7 \leq k \leq 10$, $\diamond$) defense strategies for (a) degree based attack with $\pi(k) = \alpha\,(k/K_{max})^2$ and (b) degree based attack with $\pi(k) = \alpha\sqrt{k/K_{max}}$.

We consider 4 such cuts: low degree nodes $C_1(1,2)$, mid degree nodes $C_2(3,6)$, high degree nodes $C_3(7,10)$ and very high degree nodes $C_4(10,K_{max})$. We can use the above procedure to obtain p_c for the different cases by using Eq. (8) with Eqs. (5) and (6). Fig. 2 shows the simulation results for p_c as a function of the number of additional links E for the different defense strategies. The lines in the figures represent the solutions of Eq. (6) and show that the prediction of this analytical approximation is in good agreement with the simulation results. According to these figures we can clearly see that the percolation threshold p_c of the mid degree nodes defense strategy of nodes with initial degree between $3 \leq k \leq 6$ is higher than the same p_c for low and high degree nodes defenses. We can see that by defending the mid degree nodes, the network becomes more robust to degree based attacks. For very high degree node defense we can see that the network robustness actually decreases with the added edges. This is due to adding edges to nodes that are already well connected leading to only a small increase in the network structural robustness, but also to increased vulnerability due to stronger targeting of these nodes under degree based attacks. We conclude that the mid degree nodes are very important to the robustness of the network and by inserting new links between those nodes we can better defend our network and make it much more resilient to such attacks.

In order to check the sensitivity of the general results to different attack strategies we also simulated targeted attacks with degree dependence of $\pi(k) = \alpha\,(k/K_{max})^2$ and $\pi(k) = \alpha\sqrt{k/K_{max}}$. In both cases the mid degree nodes defense proves to be the most efficient. Results are presented in Fig. 3.

We will now test these results for the special case of the Internet topology. We simulate the Internet using real data of Internet structure according to the DIMES project [20]. We use the measured map of the Internet's AS level autonomous systems as the initial network and then use the procedure presented earlier to study the effect of the different defense strategies discussed above on this network's robustness and stability to degree based attacks. Fig. 4 shows the simulation results for p_c vs the number of additional links E for different α values and for different defense strategies. The figures verify our expectations based on the random network case. We can see that the critical threshold, p_c, in the case of mid degree nodes defense strategy is higher than the same p_c for low and high degree nodes defenses. We see that by defending the mid degree nodes the Internet becomes more robust to degree based attacks. This result verifies our previous observation that defending the mid degree nodes leads to the highest improvement of network robustness against degree based attacks.

In summary, we have studied both offensive and defensive aspects of large scale-free networks. From the offensive point of view, we have studied the network robustness under degree based attacks where the vulnerability of each node depends on its degree. We showed that partial knowledge of the network structure and connectivity is sufficient to destroy the network by removing a very small fraction of the nodes. For example, we found that in a scale-free network with parameter $\gamma = 2.5$ and attack parameter $\alpha = 1$ the percolation threshold reduces drastically from $p_c = 1$ for random attack [11] to about $p_c \cong 0.2$. In intentional attack when the network structure and topology are completely known , $p_c \cong 0.07$.

The Internet, as an example for a large scale-free network [21], can be destroyed completely with very little effort even without full information about the net-

work topology. These results are applicable to other networks and also important when we consider immunization strategies where the high degree nodes through which the virus spreads are not known in advance.

From the defensive point of view, we have studied the effect of different defense strategies on the network robustness and tested different defense strategies based on adding a constant number of new links to specific groups of nodes. We showed that by defending the mid degree nodes the network becomes more robust to degree based attacks. We find that the nodes with the mid degrees $3 \leq k \leq 6$ are very important to the network stability. We have shown that in defense strategies of large scale-free networks, one should focus on the mid degree nodes in order to achieve the best results. We verified this result in the case of the Internet and showed that defending the mid degree nodes makes the Internet more stable to degree based attacks.

This work was partially supported by the BSF.

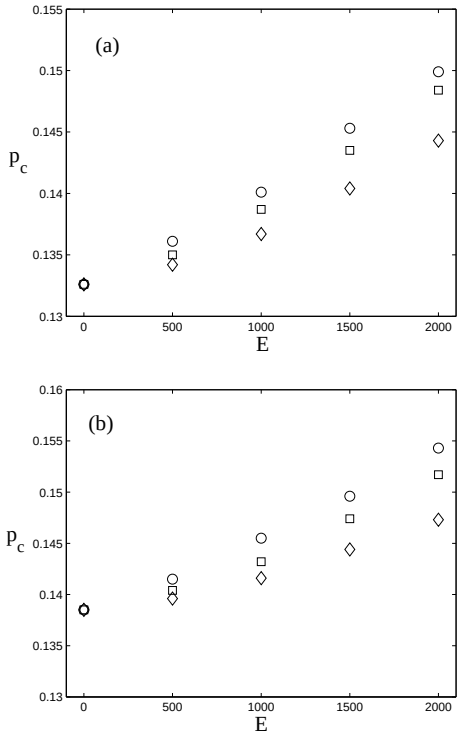

FIG. 4. Values of p_c vs E according to simulation data based on AS graphs of the Internet from the past years. The symbols represent low degree ($1 \leq k \leq 2$, $\square$), mid degree ($3 \leq k \leq 6$, $\circ$) and high degree ($7 \leq k \leq 10$, $\diamond$) defense strategies for degree based attack with $\pi(k) = \alpha\,(k/K_{max})$. (a) The attack parameter is $\alpha = 1$ (b) The attack parameter is $\alpha = 0.7$

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A Statistical Physics Approach* (Cambridge University Press, 2006).

[3] A. Barrat, M. Barthélemy and A. Vespignani, *Dynamical Processes on Complex Networks* (Cambridge University Press, 2010).

[4] M. E. J. Newman, *Networks: An Introduction* (Oxford University press, 2010).

[5] M. E. J. Newman, *SIAM Review* **45**, 167 (2003).

[6] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).

[7] R. Albert and A.-L. Barabási, Phys. Rev. Lett. **85**, 5234 (2000).

[8] P. L. Krapivsky, S. Redner, and F. Leyvraz, Phys. Rev. Lett. **85**, 4629 (2000). S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin, Phys. Rev. Lett. **85**, 4633 (2000)

[9] F. Liljeros *et al.* Nature **411**, 907 (2001).

[10] R. Cohen and S. Havlin, *Complex Networks: structure, robustness and function* (Cambridge University Press, Cambridge, UK, 2010).

[11] R. Albert, H. Jeong, and A.-L. Barabási, Nature **406**, 378 (2000).

[12] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin, Phys. Rev. Lett, **85**, 4626 (2000);**86**, 3682 (2001).

[13] A. Barrat, M. Barthélemy, R. Pastor-Satorras, A. Vespignani, Proc. Natl. Acad. Sci. U.S.A , **101**, 3747 (2004);A. Barrat, M. Barthélemy, A. Vespignani , Phys. Rev. Lett. **92**, 228701 (2004).

[14] A. E. Motter and Y.-C. Lai, Phys. Rev. E 66, 065102(R) (2002).

[15] Y. Moreno, R. Pastor-Satorras, A. Vazquez, A. Vespignani, Europhysics Letters, **62**, 292 (2003).

[16] M. Molloy and B. Reed, Random Structures and Algorithms **6**, 161 (1995).

[17] Z. Dezsö and A.-L. Barabási, Phys. Rev. E **65**, 055103(R) (2002).

[18] A. Vázquez and Y. Moreno, Phys. Rev. E **67**, 015101(R) (2003).

[19] R. Xulvi-Brunet, W. Pietsch, and I. M. Sokolov, Phys. Rev. E **68**, 036119 (2003).

[20] Y. Shavitt and E. Shir, ACM SIGCOMM Computer Communication Review, **35**, 71 (2005).

[21] L. K. Gallos, P. Argyrakis, A. Bunde, R. Cohen and S. Havlin, Phys. Rev. Lett. **94** 188701 (2005).

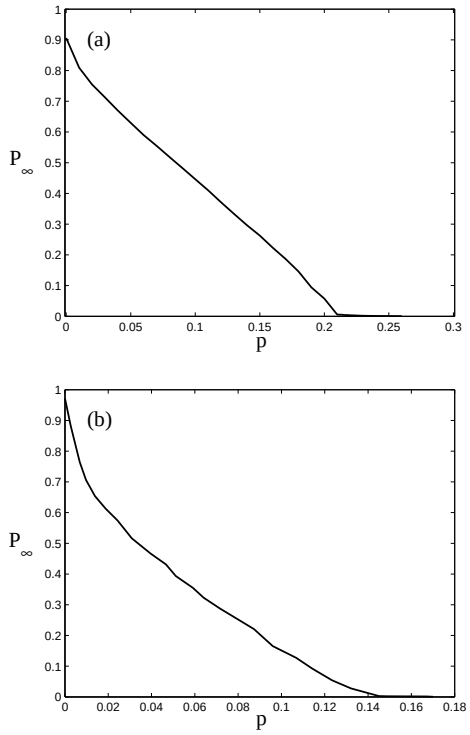

FIG. 5. Relative Size of the giant component, P_∞ as a function of the fraction of removed nodes, p for mid degree defense strategy with $E = 2000$ and attack parameter $\alpha = 1$ for degree based attack with $\pi(k) = \alpha\,(k/K_{max})$ for (a) simulated scale free network with $\gamma = 2.5$, $N = 10^5$ nodes and lower cutoff $m = 1$ and (b) the AS map network according to the DIMES project [20].