Analyzing the Vulnerability of Superpeer Networks Against Attack

B. Mitra
Dept. of CSE
IIT Kharagpur, India
bivasm@cse.iitkgp.ernet.in

F. Peruani
ZIH, Technical University of Dresden, Germany
peruani@mpi-kips-dresden.mpg.de

S. Ghose, N. Ganguly
Dept. of CSE
IIT Kharagpur, India
{sujoy, niloy}@cse.iitkgp.ernet.in

ABSTRACT

In this paper, we develop an analytical framework to measure the vulnerability of superpeer networks against attack. Two different kinds of attacks namely deterministic and degree dependent attack have been introduced here. We formally model the superpeer networks with the help of bimodal structure and different attacks with the help of graph dynamics. Our analysis shows that fraction of superpeers and their connectivity have profound impact upon the stability of the network. The results obtained from the theoretical analysis are validated through simulation. The agreement between the simulation results and theoretical predictions is almost perfect.

Categories and Subject Descriptors: C.2.0 [General]: Security and protection


Keywords: Superpeer networks, attacks, complex network theory.

1. INTRODUCTION

The growing popularity of peer-to-peer networks makes them a very likely candidate for being a substrate for future internet scale information systems. In peer to peer networks, a huge number of peers are connected among themselves by some logical links forming an overlay above the physical network. Currently superpeer topologies have emerged as the most influencing topology among various overlay networks [12, 14]. Most of the commercial systems like KaZaA [1] have also adopted superpeers in their design. In this system, superpeer nodes with high bandwidth connect to each other forming the upper level in the network hierarchy. A large number of peers are connected with superpeers to get service from them.

Understanding the effect of attacks upon the large scale superpeer networks is becoming a major challenge for the p2p network community. The most prominent attack that affects the stability of the network is Denial Of Service (DoS) attack [11]. In the p2p networks, DoS drown important peers in fastidious computation so that they fail to provide any service requested by other peers. DoS attacks become far more effective when the attack is launched in distributed fashion, the feature more popularly termed as distributed denial-of-service (DDoS) attack. The perpetrator in DDoS remotely controls personal computers and directs attacks on important peers in the network through them. Worm propagation, file poisoning, sybil attack, eclipse attack are some of the important attacks that also affect the stability of the p2p networks [13].

A survey in the literature reveals that most of the commercial superpeer networks can be represented as large scale complex graphs. Attacks upon networks can also be modeled as different kinds of dynamics that take place in these complex graphs. Some analysis of graph dynamics have been done mainly by physicists. The effect of random failures and intentional attacks in various kinds of graphs are discussed by Cohen in [3, 4]. In [9], Newman et al. introduced the concept of generating function formalism. Using it, Callaway [2] found the exact analytic solutions for percolation on random graphs with arbitrary degree distribution.

In this paper, we utilize many of the aforesaid results of percolation theory to develop a generalized analytical framework to measure the stability of superpeer networks against various kinds of attacks. The attack is modeled in terms of removal of important nodes from the network. We characterize the importance of a node mainly by its connectivity and bandwidth. Two different attack models are proposed. In the naive model, nodes are removed in the sequence of their degrees while in the second (sophisticated) model, every node has a probability of removal with probability tilted towards high degree nodes. The disintegration of giant component [8] helps us to measure the stability of the attacked network. We also perform simulation to validate the theoretical results.

The rest of the paper is organized as follows. In section 2, we develop an analytical framework to measure the stability of superpeer networks. Section 3 defines and models various environmental parameters like superpeer topology and attacks. Here we also explain the simulation environment generated to mimic large superpeer networks and specify

---

1Percolation indicates the existence of a critical probability \( p_c \) such that below \( p_c \) the network is composed of isolated clusters but above \( p_c \) a giant cluster spans the entire network (i.e. the network is almost fully connected).
how to measure the stability of the network. In section 4 we utilize the developed formalism to assess the stability of superpeer networks against attacks and validate the theoretical results with the help of simulations. Finally section 5 concludes the paper.

2. DEVELOPING ANALYTICAL FRAMEWORK USING GENERATING FUNCTION FORMALISM

In this section, we derive an analytical framework for measuring the stability of overlay structures undergoing any kind of disturbances in the network. With the help of this framework, we find the critical condition for break down of the connectivity of the network. We assume that we have an infinite system, and so before any failure or attack the biggest cluster size in the system is infinite. Theoretically the question that we want to answer is how severe should be the failure or attack to make the biggest cluster size in the system finite.

We start out by giving some definitions. Let \( p_k \) be the probability of finding a randomly chosen vertex with degree \( k \). Let \( q_k \) be the probability that a node of degree \( k \) survives the failure or attack. Correspondingly \( f_k = 1 - q_k \) is the probability that a node of degree \( k \) is removed. In our framework, \( p_k \) models the ensemble of overlay structures and \( f_k \) models the disruptive events that take place in the network. We are going to establish the relationship between stability and \( p_k \) and \( q_k \) i.e. \( (1 - f_k) \) using the generating function formalism.

Generating function has been widely used to model various stochastic processes [9, 2]. A brief introduction of generating function follows.

Generating function: A generating function \( G(x) \) is formally a power series of \( x \) which encodes some probability distribution. Let us assume that \( G(x) \) generates the degree distribution of the network given by \( p_k \), then the generating function takes the form

\[
G(x) = \sum_{k=0}^{\infty} p_k x^k
\]

(1)

The connection between the generating function and the probability distribution it generates is given by

\[
p_k = \lim_{x \to 0} \frac{1}{k!} \frac{d^k G(x)}{dx^k}
\]

(2)

Another important property of generating functions is that the average of the index of the probability, i.e., for \( G(x) \) the average degree \( z \) of a vertex, can be expressed simply by

\[
z = \langle k \rangle = \sum_{k=0}^{\infty} k p_k = G'(1)
\]

(3)

Using this formalism we can formulate the generating function \( H_0(x) \) which generates the distribution of the component sizes to which a randomly selected node belongs to. Subsequently the average size of the components can be calculated from \( H_0'(1) \). When this average component size becomes infinity, it indicates the emergence of giant component and hence we can derive the critical condition for the stability of the giant component. However to formulate \( H_0(x) \), we have to use a set of generating functions that are specified below.

Some useful generating functions:

- \( H_0(x) \) generates the distribution of the component sizes that are reached by choosing a random edge and following it to one of its ends.
- \( F_1(x) \) generates the probability distribution of the outgoing edges of the first neighbor of a randomly chosen node after the process of removal of some portion of nodes is completed.
- \( F_0(x) \) is the generating function associated with the probability of a node having degree \( k \) to be present in the network after the disruptive event.

Derivation of \( F_0(x) \):

The generating function \( F_0(x) \) specifies the probability of finding a node of degree \( k \) to be present in the network after the failure or attack. Since \( p_kq_k \) is the probability of finding a node of degree \( k \) to be present after the disruptive event, applying the definition of generating function (Eq. 1), we find that \( F_0(x) \) takes the form

\[
F_0(x) = \sum_{k=0}^{\infty} p_kq_kx^k
\]

(4)

Derivation of \( F_1(x) \):

To reach the first neighbor of a randomly chosen node, we have to pick up one of its outgoing links randomly and follow it until we reach the other end. Hence the probability distribution generated by \( F_1(x) \) is same as the probability distribution of the outgoing edges of a node reached by following a random edge. Therefore we derive the generating function \( F_1(x) \), with the help of another generating function \( A(x) \) which is based upon the probability of finding a randomly chosen node connected to a node of degree \( k \).

Derivation of \( A(x) \):

If we think of an edge connecting two nodes \( i \) and \( j \) as actually two edges; one going from \( i \) to \( j \), and another from \( j \) to \( i \), then total number of such edges in the system becomes \( \sum_{k=0}^{\infty} kn_k \), where \( n_k \) is the number of nodes with degree \( k \) in the system, which can be expressed as \( n_k = Np_k \) (\( N \) being the total number of nodes in the system). The expected number of edges connected to nodes of degree \( k \) which remained present after the node removal event is \( kn_k q_k \). So, the probability of finding a randomly chosen edge connected to a node of degree \( k \) becomes

\[
p_{oo}(k) = \frac{kn_k q_k}{\sum_{k=0}^{\infty} kn_k} = \frac{k p_k q_k}{\sum_{k=0}^{\infty} k p_k} = \frac{k p_k q_k}{z}
\]

(5)

In consequence the generating function associated to the probability \( p_{oo}(k) \) is

\[
A(x) = \sum_{k=0}^{\infty} p_{oo}(k)x^k = \sum_{k=0}^{\infty} \frac{k p_k q_k}{z}x^k
\]

Since \( \sum_{k=0}^{\infty} k p_k q_k x^k \) can be expressed as \( x F_0'(x) \) therefore with the help of Eq. (3)

\[
A(x) = x F_0'(x)/G'(1)
\]

(6)

Derivation of \( F_1(x) \):

The generating function \( F_1(x) \) is based upon the probability distribution signifying the outgoing degree of a node reached
following a random edge. We know that a node having degree \( k \) arrived following a random edge has only \( k - 1 \) outgoing links that leaves from that node. Hence probability of finding an existing node (that survives after the disruptive event) of \( k - 1 \) outgoing edges reached following a random edge is \( p_{on}(k) = \frac{kp_k q_k}{z} \) as defined in Eq. (5). Therefore probability distribution of the outgoing edges of the first neighbor of a randomly chosen node can be generated by

\[
F_1(x) = \sum_{k=1}^{\infty} p_{on} x^{k-1} = \sum_{k=1}^{\infty} \frac{kp_k q_k}{z} x^{k-1} = F_1''(x)/z
\]

**Derivation of** \( H_1(x) \)

The function \( H_1(x) \) generates the distribution of cluster sizes reached by following an edge chosen uniformly at random. Without loss of generality, we assume that following an edge, we can reach either a non-existent node (node removed during deletion) or an existent node. The probability of following the randomly chosen edge and finding an existent/present node of degree zero is zero, the probability of finding an existing node of degree one is \( p_1 q_1/z \), the probability of finding an existing node of degree two is \( 2p_2 q_2/z \), and so on. So the probability of finding a node following a random edge is \( \sum_{k=0}^{\infty} kp_k q_k/z = F_1(1) \). In consequence, the probability of finding an edge that leads to a node which has been removed is \( 1 - F_1(1) \). Clearly this is also the probability of following a randomly chosen edge that leads to a zero size component. Therefore if \( s_0 \) is the coefficient that accompanies \( x^0 \) in \( H_1(x) \) then \( s_0 = 1 - F_1(1) \).

To find the full expression of \( H_1(x) \) we have still to look for the probabilities that accompany non-zero size components. We find those probabilities next with the help of induction method.

**Calculation of** \( s_1, s_2 \) etc: We calculate the probability \( s_1 \) of finding by following a randomly chosen edge a component of size 1. This is nothing else than the sum of the probabilities of following an edge and finding a node of degree \( k \) which has its other \( k - 1 \) edges connected to zero size components (all the nodes in these components are removed) (Fig. 1(a)). This can be expressed as:

\[
s_1 = \sum_{k=1}^{\infty} \frac{kp_k q_k}{z} (1 - F_1(1))^{k-1}
\]

\[
= F_1(H_1(0)) = \lim_{x \to 0} \frac{d^2(s_0 + xF_1(H_1(x)))}{dx^2}
\]

where \( p_{on}(k) = kp_k q_k/z \) and \((1-F_1(1))^{k-1}\) is the probability of taking randomly \( k - 1 \) edges and finding that all of them are attached to zero size components.

Knowing this we can easily calculate \( s_2 \), the probability of finding a component of size 2 by following a randomly chosen edge. \( s_2 \) is the sum of the probability of following a randomly chosen edge that leads to a node of degree \( k \) which is connected to \( k - 2 \) zero size components, and has also an edge that leads to a component of size 1 (Fig. 1(b)). This can be expressed as:

\[
s_2 = \sum_{k=2}^{\infty} \frac{(k-1)p_k q_k}{z} (1 - F_1(1))^{k-2} s_1
\]

\[
= F_1'(H_1(0))H_1'(0) = \lim_{x \to 0} \frac{d^3(s_0 + xF_1(H_1(x)))}{dx^3}
\]

where \((1-F_1(1))^{k-2}s_1\) is the probability of taking randomly \( k - 1 \) edges and find that \( k - 2 \) edges are attached to zero size components, and one to a size 1 component. The term \( k - 1 \) in \( s_2 \) indicates that there are \( k - 1 \) possible configurations for these edges.

Similarly we can find the probability of finding a component of size 3 by following a randomly chosen edge

\[
s_3 = \lim_{x \to 0} \frac{d^4(s_0 + xF_1(H_1(x)))}{dx^4}
\]

and so on. This suggests a self-consistency equation for \( H_1(x) \) that generates the distribution of component sizes of nodes that are reached by randomly chosen edge after the disruptive event

\[
H_1(x) = s_0 + xF_1(H_1(x)) = 1 - F_1(1) + xF_1(H_1(x))
\]

It can be easily verified that Eq. (8) leads to the correct expressions of \( s_0, s_1, ..., s_n \) by applying Eq. (2).
The stability of superpeer networks are primarily measured in terms of a certain fraction of nodes $(f_c)$ called percolation threshold [8], removal of which disintegrates the network into large number of small, disconnected components. Below that threshold, there exists a connected component which spans the entire network. This connected component is also termed as the giant component. The value of the percolation threshold $f_c$ theoretically signifies the stability of the network, higher values indicate greater stability against attack.

We take cue from condensation theory used by physicists to develop the metric to measure the percolation threshold experimentally [5, 10]. During the experiment, we remove a fraction of nodes $f_t$ from the network in step $t$ and check whether we reach the percolation point. If not then in the next step $t+1$ we remove $f_{t+1} = f_t + \epsilon$ fraction of nodes.
and indicates the percolation point. We find out the status of the network in terms of the number and size of the components formed. We collect the statistics of $s$ and $n_s$ where $s$ denotes size of the components and $n_s$, number of components of size $s$ and define the normalized component size distribution $CS_t(s) = sn_s/\sum sn_s$ at step $t$. We compute $CS_t(s)$ for all the steps starting from $t = 1$ and observe the behavior of $CS_t(s)$ after each step (Fig. 2). Initially the $CS_t(s)$ shows unimodal character confirming a single connected component (Fig. 2(a)) or bimodal character (Fig. 2(b)) confirming a large component along with a set of small components. As the fraction of nodes removed from the network increases gradually, the network disintegrates into several components. This leads to the change in the behavior of $CS_t(s)$ whereby at a particular step $t_n$, $CS_t_n(s)$ becomes monotonically decreasing function indicating $t_n$ as the percolation point (Fig. 2(c)). Therefore $t_n$ is considered as the time step where percolation occurs and the total fraction of nodes removed at that step $f_{t_n}$ specifies the percolation threshold.

3.4 Network Generation

In our simulation, the superpeer network is represented by a simple undirected graph. In order to generate the topology, every node is assigned a degree according to the bimodal degree distribution. Thereafter the edges are generated using the “matching method” [7]. Some of the edges are then rewired using “switching method” to generate sufficient randomness in the graph [6]. In our experiment, we simulate the superpeer networks by generating graphs with 5000 nodes.

4. STABILITY OF SUPERPEER NETWORKS AGAINST ATTACK

In this section, we formally analyze the effect of attacks on the superpeer networks. The theoretically derived results are verified with the help of simulation. The two kinds of attacks, namely deterministic attack and degree dependent attack are discussed separately.

4.1 Stability analysis against deterministic attack

Two cases may arise in the deterministic attack

Case 1: The removal of a fraction of superpeers is sufficient to disintegrate the network.

Case 2: The removal of all the superpeers is not sufficient to disintegrate the network. Therefore we need to remove some of the peer nodes along with the superpeers.

We analyze these two cases separately with the help of our analytical framework. From Eq. (11) the critical condition for the stability of the superpeer networks can be rewritten as

$$\sum_{k=k_{j1},k_{m}} k(k-1)p_kq_k = \langle k \rangle$$

The equation can be further expanded as below to differentiate between peers and superpeers

$$\sum_{k=k_{j}} k(k-1)p_kq_k + \sum_{k=k_{m}} k(k-1)p_k(1-f_{sp}) = \langle k \rangle$$

**Case 1:** In this case, removal of a fraction of superpeers is sufficient to disintegrate the network. If $f_{sp}$ is the critical fraction of superpeer nodes, removal of which disintegrates the giant component then $q_k = 1$ for $k = k_j$ and $q_k = 1 - f_{sp}$ for $k = k_m$. Hence according to Eq. (12),

$$\sum_{k=k_{j}} k(k-1)p_k + \sum_{k=k_{m}} k(k-1)p_k(1-f_{sp}) = \langle k \rangle$$

$$\Rightarrow f_{sp} = 1 - \frac{\langle k \rangle - k_j(k_j-1)p_{k_j}}{k_m(k_m-1)p_{k_m}}$$

As the fraction of superpeer nodes in the network is $(1-r)$, then percolation threshold for case 1 becomes $f_{tar} = (1-r) \times f_{sp}$

$$\Rightarrow f_{tar} = (1-r) \left(1 - \frac{\langle k \rangle - k_j(k_j-1)r}{k_m(k_m-1)(1-r)}\right)$$

**Case 2:** Here we have to remove $f_p$ fraction of peer nodes along with all the superpeers to breakdown the network.
Therefore $q_k = 1 - f_p$ for $k = k_l$ and $q_k = 0$ for $k = k_m$. Hence according to Eq. (12),
\[ \sum_{k=k_l} k(k-1)p_k(1-f_p) = \langle k \rangle \]
\[ \Rightarrow f_p = 1 - \frac{\langle k \rangle}{k_l(k_l-1)p_{k_l}} \]
Therefore the total fraction of nodes required to be removed to disintegrate the network for case 2 becomes $f_{\text{tar}} = rf_p + (1 - r)$.
\[ \Rightarrow f_{\text{tar}} = r \left( 1 - \frac{\langle k \rangle}{k_l(k_l-1)r} \right) + (1 - r) \quad (14) \]

Transition point: The transition from case 1 to case 2 can be easily marked by observing the value of the percolation threshold $f_{\text{tar}}$. While calculating using Eq. (13) (case 1), if the value of $f_{\text{tar}}$ exceeds the fraction of superpeers in the network $(1 - r)$, it indicates that removal of all the superpeers is not sufficient to disrupt the network. Hence subsequently we enter into case 2 and start using Eq. (14) to find the percolation threshold. During our simulation, initially only high degree superpeer nodes in the network are removed gradually until the percolation point is reached. If the percolation point is not reached even after removing of all the superpeers, we remove a fraction of peers along with the superpeers to breakdown the network. We perform each experiment for 500 times and take the average of the percolation threshold obtained in each of them. Superpeer networks with average degree $\langle k \rangle = 10$ and superpeer degree $k_m = 50$ are considered for case study. We increase the peer degree $k_l$ gradually (the peer fraction changes accordingly) and observe the change in the percolation threshold $f_{\text{tar}}$ (Fig. 3).

Observations:

a. In networks with peer degree $k_l = 1, 2$ and $3$, the removal of only a fraction of superpeers causes breakdown of the network. Moreover, the increase of peer degree from 1 to 2 and 3 further reduces the fraction of superpeers in the network which results networks with $k_l = 2, 3$ more vulnerable. Normal wisdom would expect the attack vulnerability of the network to decrease with the decrease of fraction of superpeers. But the opposite happens here. The reason is in this zone (at $k_l = 2, 3$), although peers have a larger share in the network, yet it is not large enough to form effective connections within themselves. Therefore the stability of the network is still entirely dependent on the high degree superpeers, hence now attacking even a smaller fraction breaks down the network.

b. However as peer degree increases beyond 4, a fraction of peers is required to be removed even after removal of all the superpeers to dissolve the network. This is because the high degree peers connect among themselves and they are not entirely dependent on superpeers for connectivity. This results in the steep increase of stability of the network with peer degree $k_l \geq 5$.

Peer contribution: In addition to peer degree, we study the stability of the network with respect to a new metric namely 'peer contribution'. The peer contribution controls the total bandwidth contributed by the peers which determines the amount of influence superpeer nodes exerts on the network. The peer contribution $Pr_C$ is defined by two parameters - peer degree and fraction of peers in the network. Hence $Pr_C = \frac{r_k}{C}$ where $\langle k \rangle = r_k_l + (1 - r)k_m$. We generate three set of networks having peer degree $k_l = 1, 3$ and $5$ respectively for individual peer contribution $Pr_C (0.1 \leq Pr_C \leq 0.9)$. In order to do that, we choose fraction of peers $r$ uniformly at random and adjust superpeer degree $k_m$ accordingly to keep the peer contribution $Pr_C$ and peer degree $k_l$ constant. This procedure is followed to generate one hundred networks for each set. We restrict superpeer degree $k_m \geq 20$ in order to generate realistic superpeer networks. We theoretically compute the fraction of peers required to be removed $f_p$ and percolation threshold $f_{\text{tar}}$ for individual network and calculate their average for individual $k_l$. This expected fraction of peers required to be removed $f_p$ and percolation threshold $f_{\text{tar}}$ is plotted with respect to the peer contribution $Pr_C$ (Fig. 4).

Observations:

a. It can be observed from Fig. 4 that superpeer networks having peer degree $k_l = 1$ can be disintegrated without attacking peers at all for any peer contribution $Pr_C$. This kind of attack belongs to case 1 of the deterministic attack.

b. The peers of the superpeer networks having peer con-
The above plots illustrate the impact of peer degree upon the stability of superpeer networks in face of degree dependent attack. The fraction of peers is required to be removed to disintegrate the networks having peer degree $k_i = 5$. The impact of high degree peers upon the stability of the network becomes more eminent as peer contribution $Pr_C \geq 0.5$. In this region, a significant fraction of peers is required to be removed for all the networks having peer degree $k_i = 3.5$. This kind of attack belongs to case 2 of the deterministic attack.

d. Increase in peer contribution $Pr_C \geq 0.4$ brings the percolation threshold $f_{sa}$ and fraction of peers needed to be attacked $f_c$ close to each other which implies that stability of these networks is primarily dependent upon the stability of the peers.

e. It is interesting to observe that peer contribution $Pr_C$ has two opposite effects upon stability of the networks depending on the peer degree $k_i$. The percolation threshold $f_{sa}$ increases with peer contribution $Pr_C$ for $k_i = 3.5$, but gradually reduces for $k_i = 1$. The reason behind this is, stability of the networks with peer degree $k_i = 1$ is entirely dependent upon superpeers. Since increase in peer contribution decreases superpeer contribution, it decreases stability of these networks also. On the other hand, peers having degree $k_i \geq 3$ are strongly connected among themselves, hence stability of these networks is more dependent upon peer contribution. Hence percolation threshold $f_{sa}$ increases with peer contribution $Pr_C$.

### 4.2 Stability analysis against degree dependent attack

In this kind of attack, probability of removal of a node of degree $k$ is directly proportional to $k^\gamma$ where $\gamma \geq 0$ is a real number. Hence with proper normalization, $f_s = \frac{k^\gamma}{C}$ where $C$ is the normalizing constant. Therefore probability of survival of a node having degree $k$ after a degree dependent attack is

$$q_k = 1 - \frac{k^\gamma}{C}$$

As mentioned in bimodal degree distribution, let $r$ be the fraction of peers with degree $k_i$ and rest be superpeers of degree $k_m$. If $\langle k \rangle$ is the average degree of the network then

$$p_{ki} = r = \frac{k_m - \langle k \rangle}{k_m - k_i} \quad p_{km} = (1 - r) = \frac{\langle k \rangle - k_i}{k_m - k_i}$$

From Eq. (11) the critical condition for the stability of the giant component can be rewritten as

$$\sum_{k=k_1,k_m} k(k-1)p_k q_k = \langle k \rangle$$

$$\Rightarrow \langle k^{\gamma+2} \rangle = \langle k^{\gamma+1} \rangle - C(\langle k^\gamma \rangle - 2\langle k \rangle)$$

$$\Rightarrow rk_i^{\gamma+1}(k_i - 1) + (1 - r)k_m^{\gamma+1}(k_m - 1) = C(\langle k \rangle(k_m + k_i) - k_m - 2\langle k \rangle)$$

(15)

where $\theta^k$ moment of the bimodal degree distribution can be written as $\langle k^\theta \rangle = k_m^\theta p_{km} + k_i^\theta p_{ki}$. The solution of the Eq. (15) yields a particular value of $\gamma_c$ (termed as critical exponent) and the percolation threshold becomes

$$f_c = \frac{r k_i^{\gamma_c}}{C} + (1 - r) \frac{k_m^{\gamma_c}}{C}$$

(16)

In order to evaluate the disintegration point, proper assignment of the value of normalizing constant $C$ is necessary. Since $f_s$ should be $\leq 1 \forall k$, hence the minimum value of $C = k_m^{\gamma_c}$. Assuming this condition, the Eq. (15) becomes

$$rk_i^{\gamma+1}(k_i - 1) + (1 - r)k_m^{\gamma+1}(k_m - 1) \geq k_m^{\gamma_c}(k_m + k_i) - k_m - 2\langle k \rangle$$

(17)

The solution set of the above inequality (say $S_{\gamma_c}$) can be bounded (where $0 \leq \gamma_c \leq \gamma^{\theta k}$) or unbounded (where $0 \leq \gamma_c \leq +\infty$). Each critical exponent $\gamma_c \in S_{\gamma_c}$ specifies the fraction of peers and superpeers required to be removed to
breakdown the network. Assuming equality of Eq. (17) and hence obtaining minimum value of $C$, each $\gamma_c$ results in the corresponding normalizing constant

$$C_{\gamma_c} = \frac{r k_c^{\gamma_c+1}(k_l - 1) + (1 - r) k_m^{\gamma_c+1}(k_l - 1)}{(k)(k_m + k_l) - k_m - 2(k)}$$

(18)

Hence the fraction of peers and superpeers needed to be attacked

$$f_p^{\gamma_c} = \frac{k_l^{\gamma_c}}{C_{\gamma_c}} \quad f_{sp}^{\gamma_c} = \frac{k_m^{\gamma_c}}{C_{\gamma_c}}$$

(19)

respectively and the total fraction of nodes removed $f_c^{\gamma_c}$ is obtained from Eq. (16). The $f_c^{\gamma_c}$ depends upon the critical exponent $\gamma_c \in S_{\gamma_c}$ and normalizing constant $C_{\gamma_c}$. The nature of the solution set $S_{\gamma_c}$ has profound impact upon the behavior of $f_p^{\gamma_c}, f_{sp}^{\gamma_c}$ and as well as $f_c^{\gamma_c}$. The breakdown of the network can be due to one of the three situations noted below

1. The removal of all the superpeers along with a fraction of peers.
2. The removal of only a fraction of superpeers.
3. The removal of some fraction of both superpeers and peers.

Each situation arises due to the following reasons

1. Networks having a bounded solution set $S_{\gamma_c}$ where $0 \leq \gamma_c \leq \gamma_{c}^{bd}$ exhibit this kind of behavior at the maximum value of the solution $\gamma_c = \gamma_{c}^{bd}$. Here the fraction of superpeers removed becomes $f_{sp}^{\gamma_{c}^{bd}} = 1$ and fraction of peers removed

$$f_p^{\gamma_{c}^{bd}} = \frac{k_l^{\gamma_{c}^{bd}}}{C_{\gamma_{c}^{bd}}}$$

2. Some networks have an open solution set $S_{\gamma_c}$ where $0 \leq \gamma_c \leq +\infty$. At $\gamma_c \to \infty$, $f_p^{\gamma_c}$ converges to 0 and $f_{sp}^{\gamma_c}$ converges to some $x$ where $0 < x < 1$.

3. Intermediate critical exponents $\gamma_c \in S_{\gamma_c}$ signifies the fractional removal of both peers and superpeers.

One of the major contributions of this paper is to provide an uniform attack framework, which is able to successfully capture all the important features of the deterministic attack\(^3\) as well as provide flexibility to attack the nodes in a more suitable way. Next we perform a theoretical case study that justifies our claim. We show the impact of peer degrees upon the stability of the networks against the degree dependent attack (Fig. 5). We consider superpeer networks with two superpeer degrees $k_m = 25, 50$ and fixed average degree $k_l = 10$. The change in peer contribution $Pr_C$ due to the change of peer degree $k_l$ is shown in Fig. 5(a). In Fig 5(b), we show the behavior of boundary critical exponent $\gamma_{c}^{bd}$ with the change of peer degree. The nonexistence of $\gamma_{c}^{bd}$ implies the unbounded solution set for that particular network. Fig. 5(c) describes the fraction of peers and superpeers required to be removed ($f_p$ and $f_{sp}$ respectively) as well as percolation threshold ($f_c$) for various peer degrees.

**Observations:**

a. The peer contribution $Pr_C$ increases with superpeer degree $k_m$ for a particular peer degree $k_l$ (Fig. 5(a)). In order to keep the average degree and peer degree constant, the network with higher superpeer degree results higher fraction of peers which increases the peer contribution.
b. Fig. 5(b) shows that the solution set of inequality (17) remains unbounded for the networks having peer degree $k_l \leq 4$ with superpeer degree $k_m = 25$ and peer degree $k_l \leq 3$ with superpeer degree $k_m = 50$. This implies that removal of only a fraction of superpeers disintegrate these networks (Fig. 5(c)). This kind of attack resembles case 1 of the deterministic attack. Apart from the mathematical conclusion, we explain the phenomena from the point of relative contribution of peers and superpeers in the network. The low peer degree results in low peer contribution (Fig. 5(a)) and high superpeer contribution. Hence removal of only a fraction of superpeers is sufficient to breakdown these networks.

\(^3\)Case 1 and case 2 of the degree dependent attack resembles qualitatively as well as quantitatively with the case 2 and case 1 of the deterministic attack respectively.
c. The gradual increase in peer degree increases the peer contribution and at \( k_i = 5 \) \((k_m = 25)\), the high peer contribution ensures the necessity to remove a fraction of them to breakdown the network (Fig. 5(c)). This kind of attack resembles case 2 of the deterministic attack. The inequality (17) gets a bounded solution set for the network (Fig. 5(b)). Same thing happens for network with superpeer degree \( k_m = 50 \) at peer degree \( k_i = 4 \). Note that the peer contributions get almost same values for these two networks (Fig. 5(a)). This ensures that peer contribution has profound impact upon the stability of the network specially with the networks having high peer degree \( k_i \).

Next we illustrate the different categories of degree dependent attack with the help of individual case study.

Case study 1: First we consider superpeer networks with peers degrees \( k_i = 3, 4 \) average degree \((k) = 5\) and theoretically study the stability of the networks due to the change in the fraction of superpeers. The results of the case study are noted in Fig. 6. It can be observed that the solution set of these networks up to a threshold superpeer fraction \( sp^{th} \) \((sp^{th} = 0.19 \) and 0.41 for \( k_i = 3 \) and \( k_i = 4 \) respectively\) remains bounded and the behavior of the boundary critical exponent \( \gamma_{bd} \) due to the change of fraction of superpeers is shown in Fig. 6(a). The fraction of superpeers and peers needed to be attacked for these networks is presented in Fig. 6(b). These networks exhibit the properties of case 1 of degree dependent attack hence the removal of all the superpeers is necessary to disintegrate the network along with a fraction of peers. Fig. 6(b) also represents some instances of case 2 where only some fraction of superpeers are needed to be removed.

Observations:

a. Impact upon the fraction of peers removed

The increase of superpeer fraction slowly increases \( \gamma_{bd} \) (Fig. 6(a)) which in turn gradually decreases the fraction of peers removed \( f_{pc}^{bd} \) (Fig. 6(b)). The amount of removal of peers also depends upon the degree peer \( k_i \). The high degree peers strongly connect among themselves to enhance their influence upon the stability, hence \( f_{pc}^{bd} \) gets higher values for these networks.

b. Impact upon percolation threshold

Let the percolation threshold for the networks having super peer fraction \( sp_1 (= 1 - r_1) \) and \( sp_2 (= 1 - r_2) \) (where \( sp_1 < sp_2 \)) be \( f_{pc}^{bd} \) and \( f_{pc}^{bd} \) respectively. Hence the percolation threshold for these two networks are

\[
\Delta f_{pc}^{bd} = r_1 f_{PC}^{bd} + r_2 f_{PC}^{bd} + (1 - r_1)(1 - r_2)
\]

Therefore the change in the percolation threshold when the superpeer fraction changes from \( sp_1 \) to \( sp_2 \) is

\[
\Delta f_{pc}^{bd} = r_1 f_{PC}^{bd} - r_2 f_{PC}^{bd} + (1 - r_1) - (1 - r_2)
\]

The Eq. (20) shows that the change of percolation threshold \( f_{pc}^{bd} \) becomes influenced by two opposite forces; in one hand the increase of superpeer fraction in the network makes \( \Delta sp = \Delta (1 - r) < 0 \). On the other hand, the fraction of peers in the network as well as fraction of them required to be removed decreases (Fig. 6(b)) which makes \( \Delta f_{PC}^{bd} > 0 \). Depending upon the weightage of influence, \( f_{pc}^{bd} \) either decreases \( (k_i = 3) \) or increases \( (k_i = 4) \) slowly when the fraction of superpeers is less than \( sp^{th} \).

Case study 2: We illustrate the degree dependent attack where removal of only a fraction of superpeers is sufficient to disintegrate the network. The case study is performed with a network having superpeer degree \( k_m = 25 \), average degree \((k) = 5\) and peer degree \( k_i = 2 \) and results are validated with the help of simulation. The unbounded solution set \( S_{nc} \) of the network signifies that it belongs to case 2 of the degree dependent attack. We plot the theoretically calculated (Eq. (18), (19)) fraction of peers and superpeers required to be removed to breakdown the network for each critical exponent \( \gamma_c \) (Fig. 7). In simulation, we initially remove a fraction of superpeers \( f_{sp}^{c} \) (calculated theoretically) and then start removing peers gradually to breakdown the network. The minimum peer fraction, removal of which causes the breakdown of the network produces simulated \( f_{pc}^{c} \). We perform the simulation on graphs of 5000 nodes and repeat each experiment for 500 times and take the average of the removed peer fraction. We compare simulated result with theoretically calculated \( f_{pc}^{c} \) (Fig. 7).

Observations:

a. The fraction of peers removed gradually decreases with the increase of critical exponent \( \gamma_c \) which in turn decreases the value of \( f_{pc}^{c} \). As \( \gamma_c \to \infty \), the \( f_{pc}^{c} \to 0 \) with \( f_{sp}^{c} \to x \) (where \( 0 < x < 1 \)) and eventually \( f_{sp}^{c} \) and \( f_{pc}^{c} \) both reach some steady value. It signifies that the removal of only a fraction of superpeers is sufficient to breakdown the network (Fig. 7).

b. The nonexistence of the boundary critical exponent \( \gamma_{bd} \) for the networks having more than the threshold superpeer fraction \( sp^{th} \), signifies that the solution set of these networks is unbounded and that the percolation process belongs to case 2 (Fig 6(a)). It can be observed that fraction of peers required to be removed for these networks becomes zero (Fig. 6(b)) and removal of only a fraction of superpeers disintegrates the network.

c. It is important to note that removal of only a fraction of superpeers is sufficient to disintegrate any network with peer degree \( k_i = 1, 2 \), irrespective of superpeer degree and its fraction. Mathematically it can be explained as follows

1. The inequality (17) becomes independent of \( \gamma \) as \( k_i = 1 \). Hence at \( k_i = 1 \), the solution set \( S_{nc} \) becomes unbounded which implies that removal of only a fraction of superpeers causes breakdown of the network.
2. For $k_i = 2, 2k_i \geq k_i^2$

$$\Rightarrow 2k_i \geq rk_i^2$$

$$\Rightarrow (1 - r)k_m + 2rk_i - rk_i^2 \geq 0$$

$$\Rightarrow (1 - r)k_m(k_m - 1) \geq (k)(k_m + k_i) - k_m - 2(k)$$

$$\Rightarrow rk_i^{k_i - 1}(k_i - 1) + (1 - r)k_m^{k_m - 1}(k_m - 1) \geq k_m((k)(k_m + k_i) - k_m - 2(k))$$

Since the above inequality holds for any values of $\gamma$, it indicates that any network with $k_i = 2$ has an unbounded solution set.

Case study 3: Degree dependent attack also allows to disintegrate the network by removing a fraction of both peers and superpeers (designated as case 3 of degree dependent attack). We investigate the amount of peers and superpeers needed to be removed to dissolve the network due to the change in $\gamma_c$. We deduce the results for a network having superpeer degree $k_m = 25$, average degree $\langle k \rangle = 5$ and peer degree $k_i = 3$ and results are validated with the help of simulation (Fig. 8). The simulation set up is same as described for case 2 of the degree dependent attack.

Observations:

a. Our analytical results show that this network has bounded solution set $S_n$ of the inequality (17) and all the critical exponents $\gamma_c$ less than the boundary critical exponent $\gamma_c^{bd}$ = 1.171 results in this kind of breakdown. It is evident from both theoretical and simulation results that the removal of any combination of $(f_{pc}, f_{sp})$ (obtained from the curves in Fig. 8) where $0 \leq \gamma_c < \gamma_c^{bd}$, results in the breakdown of the network.

b. Networks with unbounded solution set (Fig 7) has finite values of $\gamma_c$ (here $\gamma_c < 1.5$ (approx)) where the removal of both fraction of peers and superpeers are necessary to disintegrate the network.

5. CONCLUSION

In this paper, we have developed an analytical framework to measure the stability of the superpeer networks against attack. The wide range of attacks have been modeled in two different ways, deterministic as well as in a more general degree dependent manner. We have applied those models in our analytical framework to measure the stability of superpeers networks. In the deterministic attack, networks having peer degree $k_i \leq 3$ are very much vulnerable and removal of only a fraction of superpeers causes the breakdown of the network. But as the peer degree increases, the stability of the network increases as well. In the degree dependent attack we have formulated a critical condition whose solution set provides the critical exponent $\gamma_c$. The peers and superpeers required to be removed is dependent upon this critical exponent $\gamma_c$ and normalizing constant $C_{\gamma_c}$. The degree dependent attack model provides us with a more general scenario where various situations can be obtained only by changing the parameter $\gamma$. We believe that this can be further extended to understand the effect of churn in the network as well as combined function of both churn and attack.

6. REFERENCES