Measuring Robustness of Superpeer Topologies[*]

Bivas Mitra, Sujoy Ghose, Niloy Ganguly
Department of Computer Science and Engineering
Indian Institute of Technology, Kharagpur, India
E-Mail: bivasm@cse.iitkgp.ernet.in

Abstract

In this paper, we propose *an analytical framework* based on percolation theory to assess the robustness of superpeer topologies in face of user churns and/or attacks targeted towards important nodes. It is observed in practice that in spite of churn of peers, superpeer networks show exceptional robustness and do not disintegrate into disconnected components. With the help of the analytical framework developed, we formally measure its stability against user churn and validate the general observation. The effect of intentional attacks upon the superpeer networks is also investigated. Our analysis shows that fraction of superpeers in the network and their connectivity have profound impact upon the stability of the network. The results obtained from the theoretical analysis are validated through simulation. The simulation results and theoretical predictions match with high degree of precision.

1 Introduction

Peer to peer (p2p) networks have recently become a popular media through which huge amount of data can be shared. P2p file sharing systems, where files are searched and downloaded among peers without the help of central servers, have emerged as a major component of Internet traffic [1, 2]. Peers are connected among themselves by some logical links forming an overlay above the physical network. Superpeer topologies have emerged as the most influencing structure among the overlay networks. Most of the commercial systems like KaZaA have also adopted superpeers in their design [3]. In this system, superpeer nodes with higher bandwidth and connectivity connect to each other forming the upper level in the network hierarchy. Each superpeer works as a server on behalf of the set of client peers who form the lower level of network hierarchy [4, 5].

Peers in the superpeer system join and leave the network randomly without any central coordination. This churn of nodes might partition the network into smaller fragments and breakdown communication among peers. But in practice, superpeer overlay networks exhibit stable behavior against churn. Consequently the possible breakdown of the network is a rare event [6]. However the stability of the overlay network can get

[*] An initial version of this paper has been presented as poster in ACM SIGCOMM 2006.

severely affected through intended attacks targeted towards the important peers [7]. A comprehensive study of stability of the superpeer networks against all these dynamics that take place in the network, is the primary focus of this paper. The main contribution of the paper lies in developing a quantitative measure to analyze the stability of networks against both churn and attack.

A survey of the literature reveals that most of the commercial superpeer networks can be modeled as complex graphs [4, 8, 9]. Some analysis of dynamics of complex graphs have been done mainly by the physicists. These approaches can be utilized to understand the various properties of superpeer overlay networks. Effect of random failures and intentional attacks in various kinds of graphs are discussed by Cohen *et al.* in [10, 11]. It has been observed that Internet, which can be modeled by power law networks is resilient to random failure, but is highly sensitive to intentional attack [12, 13]. In [14], Newman *et al.* introduced the concept of generating function formalism. Using it, Callaway [15] found the exact analytic solutions for percolation[1] on random graphs with arbitrary degree distribution. In this paper, we utilize many of the aforesaid results of percolation theory and propose a generalized equation to measure stability[2] of any given p2p overlay structures in face of churn of peers as well as attacks mounted on them.

We characterize the topology of the network by a probability distribution P and dynamics of the nodes by another probability distribution Q. Using these, we develop an analytical framework to examine the stability of generalized graphs where the vertices undergo some dynamics. The stability of superpeer networks is measured using the concept of *giant component*[3]. We also perform simulations to verify the goodness of our theoretical results.

The rest of the paper is organized as follows. Section 2 proposes an analytical framework to find the amount of disturbances required to disrupt the giant component of the network. Section 3 models the superpeer topologies as generalized random graph and also models the churns and attacks mounted on the network. In section 4 we mathematically analyze the effect of churn in the superpeer overlay networks and validate the results with the help of simulation. In section 5 the effect of targeted attack upon superpeer networks is discussed. Finally section 6 concludes the paper.

2 Developing analytical framework using generating function formalism

In this section, we use generating function to derive the general formula for measuring the stability of overlay structures undergoing any kind of disturbances in the network. We explain the basic concept behind development of the framework without going into mathematical details. Generating function is a formal power series whose coefficients encode

[1] Percolation 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).

[2] In this paper, we do not differentiate between the terms stability and robustness. They are therefore used interchangeably.

[3] Giant component is a technical term used in percolation theory which signifies the largest connected component in the network whose size is of the order of size of the network [16]

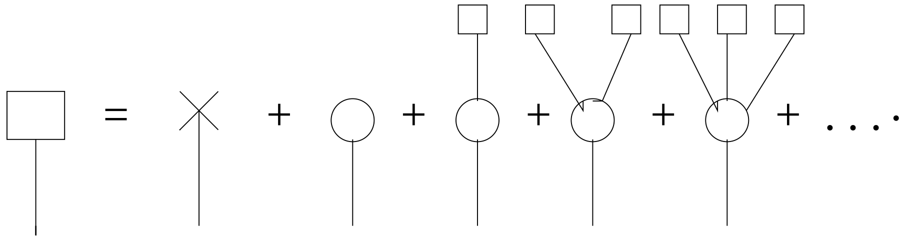


Figure 1: Schematic representation of the sum rule for the connected component of vertices reached by following a randomly chosen edge. The probability of each such component (left-hand side) can be represented as the sum of the probabilities (right hand side) of having no vertex (which has been removed), only a single vertex, having a single vertex connected to one other component, or two other components, and so forth. The entire sum can be expressed in closed form as Eq. (1) and similarly (2).

information about a sequence that is indexed by the natural numbers [14]. This generating function can be used to understand different properties of graphs. For example, let the generating function $G_0(x)$ generate the probability distribution of the vertex degrees k. Therefore $G_0(x) = \sum_{k=0}^{\infty} p_k x^k$ where p_k is the probability that a randomly chosen vertex in the graph has degree k. Importance of the generating function lies in the convenient way it can be used to understand various properties of the graph - for instance, the average degree z of a vertex in the case of $G_0(x)$ is given by $z = \langle k \rangle = \sum_k kp_k = G_0'(1)$. Higher moments can be calculated from higher derivatives also. Here we are using the generating function to explain a slightly more complicated concept.

Let q_k be the probability that a vertex of degree k be present in the network after the removal of a fraction of nodes. In our formalism f_k (=1 $- q_k$) and p_k specifies the churn/attack model and network topology respectively whose stability is subjected to examination. The formalism helps us to locate the transition point where the giant component breaks down into smaller components. $p_k.q_k$ specifies the probability of a node having degree k to be present in the network after the process of removal of some portion of nodes is completed. Hence

$$F_0(x) = \sum_{k=0}^{\infty} p_k.q_k x^k$$

becomes the generating function for this distribution. Distribution of the outgoing edges of the first neighbor of a randomly chosen node can be generated by

$$F_1(x) = \frac{\sum_k kp_k q_k x^{k-1}}{\sum_k kp_k} = F_0'(x)/z$$

where z is the average degree [15].

Let $H_1(x)$ be the generating function for the distribution of the component sizes that are reached by choosing a random edge and following it to one of its ends. Except when we are precisely at the phase transition where giant component appears, typical component size is finite. Moreover as chance of a component containing a closed loop of edges goes down exponentially with size of the graph, it becomes negligible for large graph [14]. Therefore the component may be conceptualized as a treelike structure that contain zero node if the node at the other end of the randomly selected edge is removed, which happens with probability $1 - F_1(1)$. The edge may otherwise lead to a node with k other edges leading out of it other than the edge we came in along, distributed according to $F_1(x)$ (Fig. 1). That means $H_1(x)$ satisfies a self-consistency condition of the form [15]

$$H_1(x) = 1 - F_1(1) + xF_1(H_1(x)). \tag{1}$$

The distribution for the component size to which a randomly selected node belongs to is similarly generated by (Fig. 1) $H_0(x)$ where

$$H_0(x) = 1 - F_0(1) + xF_0(H_1(x)). \tag{2}$$

Therefore the average size of the components becomes

$$H_0'(1) = \langle s \rangle = F_0(1) + \frac{F_0'(1)F_1(1)}{1 - F_1'(1)}$$

which diverges when $1 - F_1'(1) = 0$, that is the size of the component becomes infinite. We present an intuitive explanation for this critical condition of giant component formation. $F_1'(1)$ represents the average outgoing links of the first neighbor of a randomly chosen node. After the node removal process, if this average number of outgoing links is more than one, then the network should percolate, i.e. it is possible to find an infinite cluster of connected nodes. But if it is less than one, then it is very likely that by following a random edge, we land in a node that has no outgoing link and thus no chance of reaching another existing node. Therefore

$$F_1'(1) = 1 \Rightarrow \sum_{k=0}^{\infty} kp_k(kq_k - q_k - 1) = 0 \tag{3}$$

Significance of the Eq. (3) lies in the fact that it states the critical condition for the stability of giant component with respect to any type of graphs (characterized by p_k) undergoing any type of failure and attack (characterized by q_k). Using this formalism, we investigate the stability of superpeer networks in face of attack.

3 Environmental definition

In this section, we formally model the superpeer networks and churn/attack to utilize the analytical framework. Also we define the stability metric and explain the simulations undertaken to verify the theoretical results.

3.1 Topology of the superpeer overlay networks

The different types of overlay networks can be modeled using the uniform framework of probability distribution p_k, where p_k is the probability that a randomly chosen node has degree k. So the degree distribution p_k signifies the topology of the overlay network. In this paper, we model the superpeer overlay networks as mixed poisson network. In mixed poisson network, interconnection between superpeers are selected to approximate a E-R graph [17, 18] which follows Poisson distribution. Similarly the degree distribution of peers follow Poisson distribution. The average degree of the superpeers are much higher than peers. Mathematically, if r be the fraction of peers in the network[4] and rest are superpeers then degree distribution of the network

$$p_k = rp_{k_{pr}} + (1-r)p_{k_{spr}}$$

where degree distribution of peers $p_{k_{pr}} = \frac{\langle k_p \rangle^{k_{pr}} e^{-\langle k_p \rangle}}{k_{pr}!}$ and superpeers $p_{k_{spr}} = \frac{\langle k_{sp} \rangle^{k_{spr}} e^{-\langle k_{sp} \rangle}}{k_{spr}!}$ follow Poisson distribution with average degree $\langle k_p \rangle$ and $\langle k_{sp} \rangle$ respectively and $\langle k_p \rangle << \langle k_{sp} \rangle$. The average degree of the mixed poisson network becomes

$$\langle k \rangle = r\langle k_p \rangle + (1-r)\langle k_{sp} \rangle$$

3.2 Different kinds of churn and attack models

As defined in the previous section, let q_k be the probability that a vertex of degree k be present in the network after the removal of a fraction of nodes. In our framework q_k is used to specify the churn and attack models.

- In churn, the probability of removal of any randomly chosen node is degree independent and equal (constant) for all other nodes in the graph. Therefore the presence of any randomly chosen node having degree k after this kind of failure is $q_k = q$ (independent of k).

- In targeted attack, the nodes having high degrees are progressively removed. Formally $q_k = 1$ when $k < k_m$ but $0 \leq q_k < 1$ otherwise. This removes a fraction of nodes from the network with degree $\geq k_m$. Formally

  $q_k = 0$ when $k > k_m$

  $0 \leq q_k < 1$ when $k = k_m$

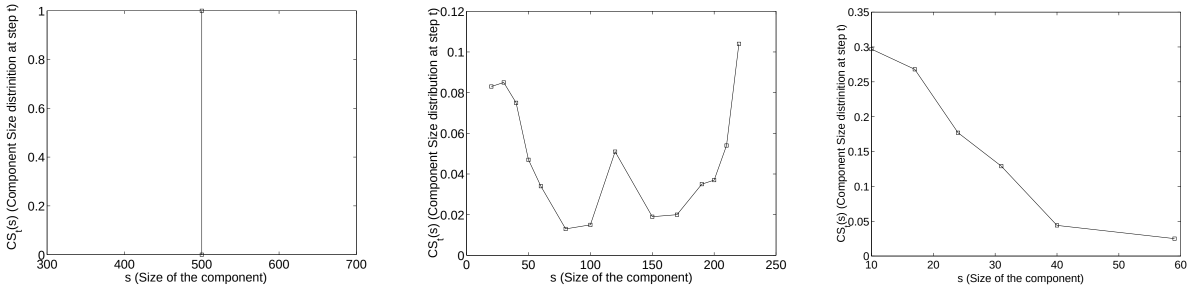
  $q_k = 1$ when $k < k_m$.

  This removes all the nodes from the network with degree greater than k_m and a fraction of nodes having degree k_m.

3.3 Stability metric

The stability and robustness of overlay networks are primarily measured in terms of certain fraction of nodes (f_c) called percolation threshold [15, 16], 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 percolation

[4] If total number of nodes in the network is N and out of them n_p is the number of peers then $r = \frac{n_p}{N}$.



(a) Initial component size distribution (only single giant component of size 500).

(b) Intermediate component size distribution.

(c) Component size distribution at percolation point.

Figure 2: The above plots represent the change in the component size distribution during percolation process and indicates the percolation point.

threshold f_c theoretically signifies the stability of the network, higher value indicates greater stability against churn and attack.

We take cue from condensation theory used by physicists to develop the metric to measure the percolation threshold experimentally [19, 20]. 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 from the network and check again. This process is continued until we reach the percolation point. After each step, 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 $CS_t(s) = sn_s / \sum_s sn_s$ 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 alongwith 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 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 Simulation environment

The superpeer overlay structure is represented by a simple undirected graph stored as an adjacency list. In order to generate the topology, every node is assigned a degree according to the mixed poisson degree distribution. Thereafter the edges are generated using the “matching method” [21]. Some of the edges are then rewired using “switching method” to generate sufficient randomness in the graph [22]. In our experiment, we simulate the overlay network by generating graphs with 5000 nodes.

Churn or attack on a peer effectively means deletion of the node and its corresponding edges. We implement this phenomena by removing a fraction of nodes in each step depending on the disrupting event in the network. In the case of churn, nodes are randomly selected using a time-seeded pseudo-random number generator and its edges are removed from the adjacency list. For targeted attack, high degree nodes in the network are removed sequentially in each step until the percolation point is reached. We perform each experiment for 500 times and take the average of the percolation threshold.

4 Stability of superpeer networks against churn

The superpeer networks mostly suffer from the churn of peers which can be modeled by the random failure of nodes in complex graph. In this section, we use our equation to show that stability of the superpeer networks is quite unaffected due to churn of peers. We validate the theoretical results with the help of simulation. At first, we present the result for generalized random graph and then customize it for superpeer networks.

Generalized random graph

In this section, we discuss the effect of random failure in a generalized random graph. If $q = q_r$ is the critical fraction of nodes whose presence in the graph is essential for the stability of the giant component after this kind of failure then according to Eq. (3)

$$\sum_{k=0}^{\infty} kp_k(kq_r - q_r - 1) = 0$$

$$\Rightarrow q_r = \frac{1}{\frac{\sum_{k=0}^{\infty} k^2 p_k}{\sum_{k=0}^{\infty} kp_k} - 1} \Rightarrow q_r = \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1}$$

where $\langle k^2 \rangle = \sum_{k=0}^{\infty} k^2 p_k$ and $\langle k \rangle = \sum_{k=0}^{\infty} kp_k$ are the second and the first moment of the degree distribution respectively. Now if f_r is the critical fraction of nodes whose random removal disintegrates the giant component then $f_r = 1 - q_r$. Therefore percolation threshold

$$f_r = 1 - \frac{1}{\frac{\langle k^2 \rangle}{\langle k \rangle} - 1} \tag{4}$$

This is the well known condition [10] (derived differently) for the disappearance of the giant component due to random failure. Note that, we have reproduced it to show that it can also be derived from the proposed general formula (Eq. (3)).

Superpeer networks

In mixed poisson network, let r be the fraction of peers in the network and rest be superpeers. Superpeer nodes are connected to each other to form an E-R network [17, 18] with average degree $\langle k_{sp} \rangle$. Similarly peers connected with superpeers forms another E-R graph with an average degree $\langle k_p \rangle$ where $\langle k_p \rangle << \langle k_{sp} \rangle$. Now we examine the stability of this kind of superpeer network undergoing churn. In mixed poisson network, first and second moment of the degree distribution becomes $\langle k \rangle = r\langle k_p \rangle + (1 - r)\langle k_{sp} \rangle$ and $\langle k^2 \rangle = r\langle k_p^2 \rangle + (1 - r)\langle k_{sp}^2 \rangle$ respectively. If k is a random variable following Poisson distribution then it can be shown that $\langle k^2 \rangle \approx \langle k \rangle^2 + \langle k \rangle$. Hence according to Eq. (4),

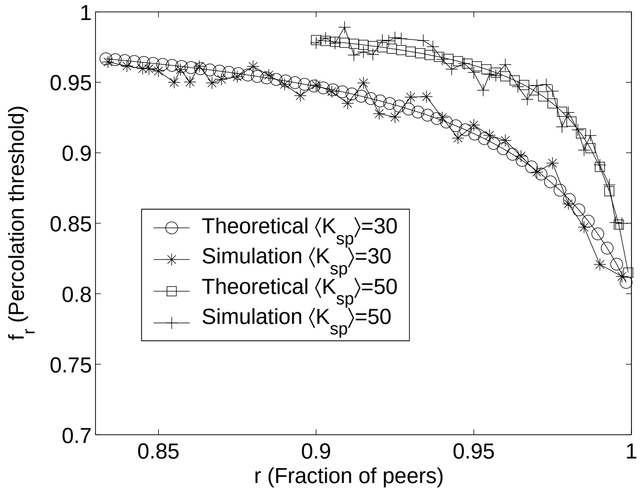


Figure 3: The above plots represent a comparative study of theoretical and simulation results of stability for two mixed poisson networks undergoing churn. Here X-axis represents the fraction of peer nodes (r) exists in the network and Y-axis represents the corresponding percolation threshold (f_r). We keep the average degree $\langle k \rangle = 5$ fixed and vary the mean superpeer degree $\langle k_{sp} \rangle = 30, 50$ for two plots.

percolation threshold becomes

$$f_r = 1 - \frac{r\langle k_p \rangle + (1-r)\langle k_{sp} \rangle}{r\langle k_p \rangle^2 + (1-r)\langle k_{sp} \rangle^2}$$

Substituting for $\langle k_p \rangle$, we get

$$f_r = 1 - \frac{\langle k \rangle r}{\langle k \rangle^2 - 2\langle k \rangle(1-r)\langle k_{sp} \rangle + (1-r)^2\langle k_{sp} \rangle^2 + r(1-r)\langle k_{sp} \rangle^2} \tag{5}$$

Feasible fraction of peers : Since the mean peer degree $\langle k_p \rangle$ needs to be > 0 to be connected in the network therefore

$$\frac{\langle k \rangle - (1-r_r)\langle k_{sp} \rangle}{r_r} > 0$$

$$\Rightarrow r_r > 1 - \frac{\langle k \rangle}{\langle k_{sp} \rangle}$$

That means we can form a connected superpeer network with prescribed peer and superpeer degrees only if the fraction of peers in the network is greater than the feasible peer fraction (r_r). For $\langle k_{sp} \rangle = 30, 50$ this feasible fraction r_r becomes $0.833, 0.90$ respectively. Below that fraction, there does not exist any network, therefore our theoretical analysis as well as simulations are performed with peer fraction r above the feasible fraction r_r. Using Eq. (5), we study the variation of percolation threshold (f_r) due to the change in the fraction of peers (r). We validate the analytically derived result with the help of simulation. We perform the simulation on two mixed poisson networks with average superpeer degree $\langle k_{sp} \rangle = 30$ and 50, keeping the average degree $\langle k \rangle = 5$. Comparative study reveals that networks having higher superpeer degree exhibit more robustness than with lower superpeer degree for any peer-superpeer ratio. *It can be observed from Fig. 3 that simulation results match closely with theoretical predictions which shows the success of our theoretical framework.*

Observations:

1. It is important to observe that for the entire range of peer fractions, the percolation threshold f_r is greater than 0.7 which implies that superpeer networks are quite robust against churn.

> Since churn affects peers and superpeers depending upon their individual fraction in the network, peers are affected much more than superpeers. The removal of a significant number of low degree peers alongwith a few high degree superpeers have less impact upon the stability of the networks. Practical experience also ensures that superpeer networks exhibit high robustness in face of churn.

2. Another significant observation is, lower fraction of superpeers in the network (specifically when it is below 5%) results in a sharp fall of f_r, that is the vulnerability of the network drastically increases when the fraction of superpeers is below 5%.

> Higher fraction of superpeers results in low mean peer connectivity. Therefore most of the peers are only connected to superpeers (and not within themselves), hence stability of the network depends entirely upon superpeers. As fraction of superpeer reduces below 5%, mean peer degree becomes quite high (4 to 5). This gives rise to situations where some peers are not connected to the superpeers at all, but only connected to fellow peers. Hence removal of individual peers also result in the removal of fellow peers. This produces an avalanche effect which results in a drastic reduction of stability of the network in this region.

5 Stability of superpeer networks against occasional attack

Stability of the superpeer networks is challenged by various kinds of attacks on prominent peers or superpeers. The attack model has been formally defined in section 3. In this section, we analyze the effect of this kind of targeted attack upon superpeer networks where r be the fraction of peers and rest are superpeers. In the case of targeted attack two cases may arise

Case 1 Removal of a fraction of superpeers is sufficient to disintegrate the network. This happens when the percentage of superpeers is relatively higher than peers.

Case 2 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. (3) the critical condition for the stability of the giant component can be rewritten as

$$\sum_{k=0}^{\infty} k(k-1) p_k q_k = \langle k \rangle$$

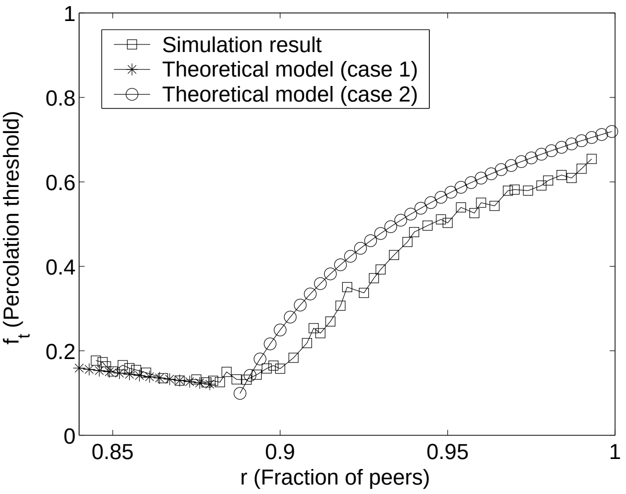


Figure 4: The above plot represents the behavior of the mixed poisson network in face of targeted attack found experimentally and compares it with the proposed theoretical model. Here X-axis represents the fraction of peer nodes (r) exists in the network and Y-axis represents the corresponding percolation threshold (f_t). We keep the average degree $\langle k\rangle = 5$ and mean superpeer degree $\langle k_{sp}\rangle = 30$ fixed. Case 1 and case 2 of the theoretical model represent Eq. (7) and (8) respectively.

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

$$\sum_{k=0}^{k_{max}-1} k(k-1)p_k q_k + \sum_{k=k_{max}}^{\infty} k(k-1)p_k q_k = \langle k\rangle \tag{6}$$

where all the nodes having degree less than k_{max} are peers and rest are superpeers.
Case 1: In this case, removal of a fraction of superpeers is sufficient to disintegrate the network. If f_{sp} be the critical fraction of superpeer nodes, removal of which disintegrates the giant component then $q_k = 1$ for $k < k_{max}$ and $q_k = 1 - f_{sp}$ for $k \geq k_{max}$. Hence according to Eq. (6),

$$\sum_{k=0}^{k_{max}-1} k(k-1)p_k + \sum_{k=k_{max}}^{\infty} k(k-1)p_k(1-f_{sp}) = \langle k\rangle$$

$$\Rightarrow f_{sp} = 1 - \frac{\langle k\rangle - \sum_{k=0}^{k_{max}-1} k(k-1)p_k}{\sum_{k=k_{max}}^{\infty} k(k-1)p_k}$$

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

$$\begin{aligned}\Rightarrow f_t &= (1-r)\left(1 - \frac{\langle k\rangle - \sum_{k=0}^{k_{max}-1} k(k-1)p_k}{\sum_{k=k_{max}}^{\infty} k(k-1)p_k}\right)\\ &= (1-r)\left(1 - \frac{\langle k\rangle - r\sum_{k=0}^{\langle k_p\rangle+\delta} k(k-1)\frac{\langle k_p\rangle^k e^{-\langle k_p\rangle}}{k!}}{(1-r)\sum_{k=\langle k_p\rangle+\delta+1}^{\infty} k(k-1)\frac{\langle k_{sp}\rangle^k e^{-\langle k_{sp}\rangle}}{k!}}\right)\end{aligned} \tag{7}$$

where mean peer degree $\langle k_p \rangle = \frac{\langle k \rangle - (1-r)\langle k_{sp} \rangle}{r}$ and we choose suitable value of δ depending on the standard deviation of the Poisson distribution. δ ensures the inclusion of all peer and superpeer degrees around their respective means $\langle k_p \rangle$ and $\langle k_{sp} \rangle$ during the calculation of above equations.

Case 2: Here we have to remove f_p fraction of peer nodes alongwith all the superpeers to breakdown the network. Therefore $q_k = 1 - f_p$ for $k < k_{max}$ and $q_k = 0$ for $k \geq k_{max}$. Hence according to Eq. (6),

$$\sum_{k=0}^{k_{max}-1} k(k-1)p_k(1-f_p) = \langle k \rangle$$

$$\Rightarrow f_p = 1 - \frac{\langle k \rangle}{\sum_{k=0}^{k_{max}-1} k(k-1)p_k}$$

Therefore the total fraction of nodes required to be removed to disintegrate the network for case 2 becomes $f_t = rf_p + (1-r)$.

$$\begin{aligned} \Rightarrow f_t &= r\left(1 - \frac{\langle k \rangle}{\sum_{k=0}^{k_{max}-1} k(k-1)p_k}\right) + (1-r) \\ &= r\left(1 - \frac{\langle k \rangle}{r\sum_{k=0}^{\langle k_p \rangle + \delta} k(k-1)\frac{\langle k_p \rangle^k e^{-\langle k_p \rangle}}{k!}}\right) + (1-r) \end{aligned} \qquad (8)$$

where mean peer degree $\langle k_p \rangle = \frac{\langle k \rangle - (1-r)\langle k_{sp} \rangle}{r}$.

Transition point: The transition from case 1 to case 2 can be easily marked by observing the value of percolation threshold f_t. While calculating using Eq. (7) (case 1), if the percolation threshold f_t 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. (8) to find percolation threshold.

We validate our theoretical model of attack on mixed poisson network with the help of simulation. In simulation, we consider a mixed poisson network with average degree $\langle k \rangle = 5$ and mean superpeer degree $\langle k_{sp} \rangle = 30$. We increase the fraction of peers gradually keeping average degree $\langle k \rangle = 5$ fixed and observe the change in the percolation threshold f_t (Fig. 4). It is important to note that when the fraction of superpeers in the network is high, it is possible to breakdown the network only by removing a fraction of superpeers and modeled as case 1 (Eq. (7)). But when the fraction of superpeers is below some threshold, a fraction of peers should be attacked alongwith the superpeers to stop percolation in the network and modeled as case 2 (Eq. (8)).

Observations: In the networks with peer fraction $r < 0.89$ (where mean peer degree $0 < \langle k_p \rangle \leq 2$), the removal of only a fraction of superpeers causes breakdown hence makes these networks vulnerable. Moreover, increase of peer fraction r in this range increases mean peer degree from 1 to 2 that makes networks with $\langle k_p \rangle = 2$ more vulnerable. Normal wisdom would expect the attack vulnerability of the network to decrease with the increase of fraction of peers. But the opposite happens here. The reason is in this zone, 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 superpeers, hence now attacking even a smaller fraction

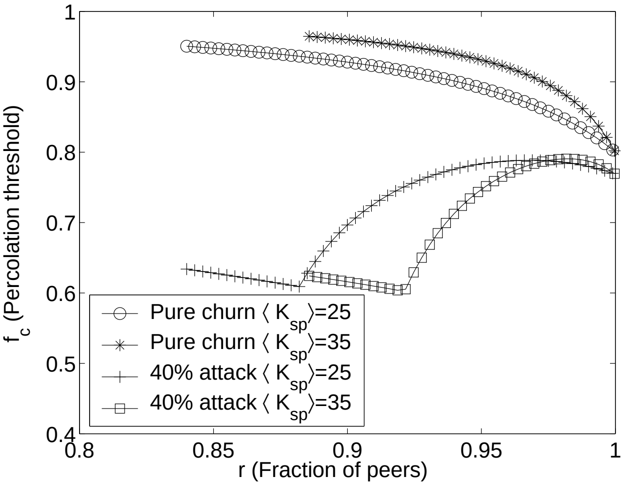


Figure 5: We plot percolation threshold f_c for various peer fraction r. Two different mixed poisson networks have been considered with average superpeer degree $\langle k_{sp} \rangle = 25, 35$ with fixed average degree $\langle k \rangle = 5$. Feasible fraction of peers are considered only. We compare theoretically the stability of these two networks against pure churn and combination of churn (60%) and attack (40%). Comparative study shows that the impact of the combination is more severe for the network having higher mean superpeer degree ($\langle k_{sp} \rangle = 35$) specially when the fraction of superpeers in the network is quite high. As the fraction of superpeers decreases, the influence of the superpeers as well as attack upon the stability of the network decreases. Hence the percolation threshold of both networks becomes close to each other.

breaks down the network.
However as peer fraction becomes ≥ 0.89, the mean peer degree increases to 3 and 4 and 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 are not entirely dependent on superpeers for connectivity. This results in the increase of stability of the network with peer degree $\langle k_p \rangle = 4$.

6 Conclusion

In this paper we have developed a common analytical framework to evaluate the robustness of superpeer networks against various disturbances in the network. We have modeled superpeer networks by mixed poisson degree distribution. We have also modeled the churn of peers as random failure of nodes. It has been observed from both theoretical and simulation results that superpeer networks remain robust for user churn. Next we have analyzed the behavior of superpeer networks in face of targeted attack. Unlike churn, in this case increase of peers improves the stability of the network and the rate of improvement is almost linear to the fraction of peers present in the network.

Our analysis has shown that presence of superpeers impart conflicting advantages for churn and attack. Hence proper mix of fraction of superpeers with peers is necessary to improve the robustness of the network in face of combination of churn and attack. It appears from Fig. 5 that when percentage of attack is 40%, the network having lower superpeer degrees ($\langle k_{sp} \rangle = 25$) performs better than network having higher superpeer degree ($\langle k_{sp} \rangle = 35$). So to obtain optimized performance, it is upto the design engineers to choose the correct superpeer to peer ratio depending on the working environment. The theoretical framework developed in this paper will help them to easily and accurately calculate the ratio.

References

[1] Q. Lv, P. Cao, E. Cohen, K. Li and S. Shenker : "Search and Replication in Unstructured Peer-to-Peer Networks", ACM International Conference on Supercomputing, New York, USA, 2002.

[2] N. Ganguly, A. Deutsch : "Developing Efficient Search Algorithms for P2P Networks Using Proliferation and Mutation", Proceedings of the International Conference on Artificial Immune Systems, Catania, Italy, 13-16 September 2004.

[3] KaZaA website. http://www.kazaa.com.

[4] Y. J. Pyun, D. S. Reeves : "Constructing a Balanced, log(N)-Diameter Super-peer Topology", Proceedings of the 4^{th} International Conference on Peer-to-Peer Computing, Zurich, Switzerland, August 2004.

[5] B. Yang, H. Garcua-Molina : "Designing a Super-Peer Networks", Proceedings of the International Conference on Data Engineering (ICDE), Los Alamitos, CA, March 2003.

[6] D. Stutzbach, R. Rejaie, S. Sen : "Characterizing Unstructured Overlay Topologies in Modern P2P File-Sharing Systems", Proceedings of ACM SIGCOMM/USENIX Internet Measurement Conference, Berkeley, CA, October 2005.

[7] S. Saroiu, P. K. Gummadi, S. D. Gribble : "Measuring and Analyzing the Characteristics of Napster and Gnutella Hosts", Multimedia Systems Journal, 8(5), Nov. 2002.

[8] A. Montresor : "A Robust Protocol for Building Superpeer Overlay Topologies", Proceedings of the 4^{th} International Conference on Peer-to-Peer Computing, Zurich, Switzerland, August 2004.

[9] Rdiger Schollmeier : "Signaling And Networking In Unstructured Peer-To-Peer Networks", Ph.D. Thesis, Technical University, Munchen, pp 75-86, 2005.

[10] R. Cohen, K. Erez, D. Avraham, S. Havlin : "Resilience of the Internet to Random Breakdown", Vol. 85, No. 21 Physical Review Letters, 2000.

[11] R. Cohen, K. Erez, D. Avraham, S. Havlin : "Resilience of the Internet under Intentional Attack", Vol. 86, No. 16 Physical Review Letters, 2001.

[12] R. Albert, H. Jhong, A. L. Barabsi : “Error and Attack Tolerance of Complex Networks”, Nature, 406, 2000.

[13] T. Tanizawa, G. Paul, R. Cohen, S. Havlin, H. E. Stanley : “Optimization of Network Robustness to Waves of Targeted and Random Attack”, Physical Review E., 71, 047101, 2005.

[14] M. E. J. Newman, S. H. Strogatz, D. J. Watts : “Random Graphs with Arbitrary Degree Distributions and Their Application”, Physical Review E, 2001.

[15] D. S. Callaway , M. E. J. Newman, S. H. Strogatz, D. J. Watts : “Network Robustness and Fragility: Percolation on Random graphs”, Vol. 85, No. 21 Physical Review Letters, 2000.

[16] M. Molloy, B. Reed : “The Size of the Giant Component of a Random Graph with a Given Degree Sequence”, Combinatorics, Probability and Computing 7, 295-298, 1998.

[17] P. Erdos, A. Renyi : “On Random Graphs I”, Publ. Mathematical, 6, Debrecen, 1959, 290-297.

[18] P. Erdos, A. Renyi : “On the Evolution of Random Graphs”, Publ. Math. Inst. Hangar Acad. Sci., 5, 1960, 17-61.

[19] S. N. Majumdar, M. R. Evans and R. K. P. Zia : “Nature of the Condensate in Mass Transport Models”, Physical Review Letters 94, 180601, 2005.

[20] F. Peruani, A. Deutsch, M. Baer : “Nonequilibrium clustering of self-propelled rods”, Physical Review E 74, 030904(R), 2006.

[21] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan., D. Chklovskii, and U. Alon : “Network Motifs : Simple Building Blocks Of Complex Networks”, Science 298, 824827, 2002.

[22] R. Milo, N. Kashtan, S. Itzkovitz, M. E. J. Newman, U. Alon : “On the Uniform Generation of Random Graphs with Prescribed Degree Sequences”, eprint arXiv:cond-mat/0312028, 2003.