Efficient Immunization Algorithm for Peer-to-Peer Networks*

Hao Chen, Hai Jin, Jianhua Sun, and Zongfen Han

Cluster and Grid Computing Lab,
Huazhong University of Science and Technology, Wuhan, 430074, China
{haochen, hjin, jhsun, zfhan}@hust.edu.cn

Abstract. In this paper, we present a detail study about the immunization of viruses in Peer-to-Peer networks exhibiting power-law degree distributions. By comparing two different immunization strategies (randomized and degree-based), we conclude that it is efficient to immunize the highly connected nodes in order to eradicate viruses from the network. Furthermore, we propose an efficient updating algorithm for global virus database according to the degree-based immunization strategy.

1 Introduction

In the past several years, Peer-to-Peer (P2P) networks have emerged as effective ways for communication and cooperation among geographically distributed computers. P2P systems depend on voluntary participation of peers without any centralized control and hierarchial organization, from which the underlying infrastructure is constructed. In P2P networks (e.g. SETI@Home, Freenet, Gnutella, Napster), through cooperation of all peers, tremendous computation and storage resources unoccupied on individual computers can be utilized to accomplish some kinds of tasks jointly. Individual computers communicate with each other directly without a central point of coordination. P2P systems are often built at the application level and use their own communication protocols to form a virtual network over the underlying physical network. The topology of the virtual network shares some common properties of complex networks in other disciplines of science, and has a significant impact on performance, scalability and robustness of P2P systems.

Recently, a large proportion of research effort has been devoted to the study and modeling of a wide range of natural systems that can be regarded as networks, focusing on large scale statistical properties of networks other than single small networks. Some reviews on complex networks can be found in [11]. From biology to social science to computer science, systems such as the Internet [8], the World-Wide-Web [5], social communities, food web and biological networks can be represented as graphs, where nodes represent individuals and links represent interactions among them. Despite this simple definition, these networks often exhibit high degree of complexity due to the wiring

* This paper is supported by National Science Foundation of China under grant 60125208 and 60273076.

L. Bougé and V.K. Prasanna (Eds.): HiPC 2004, LNCS 3296, pp. 232–241, 2004.

entanglement during their growth. Researches on these networks have revealed some commonalities. Specially, many of these networks have complex topological properties and dynamical features that can not be explained by the classical graph model of random networks, the Erdos-Renyi model [4].

These diverse networks can be characterized more accurately by small world phenomenon and power-law degree distributions. The first demonstration of small world effect was introduced by the classic experiment by Stanley Milgram [10], which showed that people could find a short sequence of acquaintances in order to deliver a message to each other, and are often referred to as "six degrees of separation". In networks with power-law degree distributions, the probability distribution of the degree of the node is approximately proportional to $k^{-\gamma}$, where k is the node degree and γ is a constant. Such networks are often called scale-free networks. However, the degree distribution in random graph networks follows a Poisson distribution.

One important characteristic of P2P networks, like some other complex networks, is that they often show high degree of tolerance against random failures, while they are vulnerable under intentional attacks [6]. Such property has motivated us to carry out a study about the virus spreading phenomena and hacker behaviors in P2P networks from a topological point of view. In our study, we choose Gnutella as our testbed, due to its large user community and open architecture. Some previous works have been done on the measurement and analysis of Gnutella network [12, 14, 9], such as bottleneck bandwidth [14] and search algorithms [1]. But few work has been devoted to investigate the behaviors of virus spreading and intrusion from a topological view. The main contributions of this paper are two-fold: firstly, an optimal immunization strategy is given; secondly, we propose an efficient information updating algorithm for P2P networks based on the immunization strategy.

The rest of this paper is organized as follows. Section 2 describes the immunization model of P2P networks. In section 3, we propose an information updating algorithm for P2P networks. In section 4, we give our conclusions and point out some directions for future work.

2 Immunization Model of P2P Networks

Some previous works [1, 6] indicate that P2P networks often display small-world phenomenon and power-law degree distributions. Many topological properties of such power-law networks are much different from those of networks modelled by random graphs [4]. One of the most important property is the network resilience, which measures the network robustness and weakness by random removal or targeted deletion of vertices in the network. In this section, we first review some materials about the resilience of P2P network, from which implications for designing immunization strategy of P2P network will be introduced.

2.1 Network Resilience and Its Implication for Immunization of P2P Networks

There are a variety of different strategies of removing nodes form a network, and different networks may show varying degrees of resilience to these strategies. For example, one

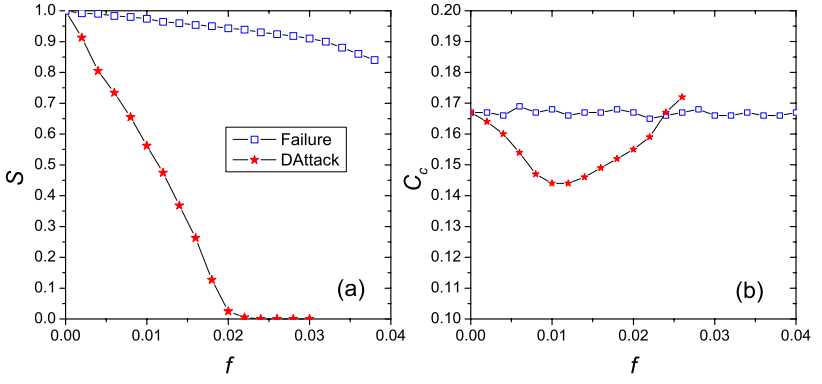

Fig. 1. Results for random failures (open square) and degree-based (star) attacks of nodes measured by the relative size of largest cluster S, the average closeness centrality C_c as functions of the fraction of removed nodes f in Gnutella network

could remove some nodes randomly in a network, or other nodes with high degrees. Removal of important nodes may affect the network significantly. With the removing of nodes from a network, some paths between pairs of nodes is broken. The average length of these paths increases. Eventually nodes are isolated in different clusters, and communications between them become impossible. Some real networks display high degree of robustness against random failures of nodes, but they are also very vulnerable under attacks of the high degree nodes.

In the following, we illustrate the network resilience of Gnutella network based on our previous work [6]. To explain the damages caused by attacks and random failures, we measure two parameters: the relative size of the largest cluster S (defined as the ratio between the size of the largest cluster and the size of original network) and the average closeness centrality[1] C_c (defined as the average of the closeness centralities of all nodes in the largest cluster).

As shown in Fig.1, Gnutella network shows high degree of tolerance against random failures. However, the fault tolerance comes at the expense of attack vulnerability: rapid decreasing of the relative size of the largest cluster and the average closeness centrality in early stage. In Fig.1 (b), after the critical point, the largest cluster becomes much smaller than the initial size of the network, which causes the fallback of average path length in such clusters and the increasing of C_c correspondingly. A more detailed description can be referred to [6].

It is intuitive that the attacks on high degree nodes are analogous to the malicious behavior of hackers in reality. They often make hosts malfunctioning by brute attacks. In addition, there are other dangers caused by computer viruses or backdoor programs (programs left by hackers that reside in hosts and can be used to intrude into other hosts

[1] Closeness centrality is the measurement of the shortest path length of one node to all others in the network. See [6] for a more detailed description.

without breaking down systems). If they are not controlled properly, they will spread to the whole system. Hence, the problems remained to us are: how efficiently can we stop the spreading of viruses (in the rest of this paper, viruses stand for both computer viruses and backdoor-like programs unless explicitly stated)? How can one node inform other nodes when it detects a virus? Imagine that if one successfully intrude into one host and spread in a P2P network, DDOS(Distributed Denial of Service) attacks could be easily performed. These are exactly the topics to be discussed in the following sections.

2.2 Modeling Immunization of P2P Network

One model of the spread of a virus over a network is the SIR (*susceptible-infective-recovered*) model [7]. This model assumes that the nodes in the network can be in three states: *susceptible* (one node is healthy but could be infected by others), *infective* (one node has the virus, and can spread it to others), or *recovered* (one node has recovered from the virus and has permanent immunity, so that it can never be infected again or spread it).

Another widely used model of virus spreading is called SIS (*susceptible-infective-susceptible*) model [7]. The main difference with SIR model is that one node can be infected again without permanent immunity, even though it once recovered from the virus. Comparing the two models, we know that the SIS model is more suitable for modeling the spread of computer virus or intrusion in P2P networks, since viruses or intrusions in the network can be cured by antivirus software or be blocked by intrusion detection system. But without a permanent virus-checking or intrusion-detecting program, they have no way to defend the subsequent attacks by the same virus or intrusion. Hence, we use SIS model to investigate the effect of virus spreading in P2P networks.

In SIS model, regarding P2P networks as graphs, we represent individuals by nodes, which can be either "healthy" or "infected", and represent connections between individuals by links, along which the infection can spread. Each node (susceptible) is infected with rate ν if it is connected to one or more infected nodes. At the same time, an infected node is cured with rate δ, defining an effective spreading rate $\lambda = \nu/\delta$ for the virus. Without lose of generality, we set $\delta = 1$. Viruses whose spreading rate exceeds a critical threshold λ_c will persist, while those under the threshold will die out shortly. This model can be used to investigate epidemic states of viruses of P2P networks, in which a stationary proportion of nodes is infected.

P2P networks often exhibit power-law degree distributions [1, 6], similar to other complex networks. A widely used theoretical model for such power-law networks is the Barabasi and Albert (BA) model [3], which describes the growth of complex networks by two basic features: the growing nature of the networks and a preferential attachment rule. The algorithm of BA model is as following: Staring with a small number (m_0) of nodes, at every step we add a new node with *m* edges that link the new node to *m* different nodes already in the system. The probability that a new node will be connected to node *i* depends on the degree k_i of node *i*, such that $\prod(k_i) = k_i / \sum_j k_j$. After *n* steps, we obtain a network with degree distribution $p_k(k) = 2m^2k^{-3}$. In the following, we use the BA model to deduce a theoretical framework of the prevalence of virus, and then compare with the real data obtained from Gnutella network [6].

In order to take into account the different connectivity of all the nodes, we denote the density of infected nodes with degree k by $\rho_k(t)$, where the parameter t indicates the time evolution, and denote the average density of all infected nodes in the network by $\rho = \Sigma_k p(k)\rho_k$. According to the mean-field theory as in [13], we have the following equation:

$$\frac{d\rho_k(t)}{dt} = -\rho_k(t) + \lambda k[1 - \rho_k(t)]\Theta(\lambda). \tag{1}$$

The first term in the right-hand side describes the probability that an infected node is cured. The second term is the probability that a healthy node with degree k is infected, proportional to the infection rate λ, the probability $1 - \rho_k(t)$ that a node with degree k is healthy and the probability $\Theta(\lambda)$ that a given link point to an infected node. The probability $\Theta(\lambda)$ is proportional to the average degree $\langle k\rangle = \sum_k kp(k)$ of all the nodes, and it can be written as:

$$\Theta(\lambda) = \frac{\sum_k kp(k)\rho_k(t)}{\langle k\rangle}. \tag{2}$$

Imposing the stationary condition $\frac{d\rho_k(t)}{dt} = 0$ when the system is at large times such that the number of infected nodes are balanced with the number of healthy nodes, we find the stationary density as:

$$\rho_k = \frac{k\lambda\Theta(\lambda)}{1 + k\lambda\Theta(\lambda)}. \tag{3}$$

Using a continuous k approximation, we calculate $\Theta(\lambda)$ for BA model, whose average degree is $\langle k\rangle = \sum_k kp(k) = \int_m^\infty k2m^2k^{-3}dk = 2m$, as:

$$\Theta(\lambda) \simeq m\int_m^\infty \frac{\lambda\Theta(\lambda)}{k(1 + k\lambda\Theta(\lambda))}dk = \frac{e^{-1/m\lambda}}{(1 - e^{-1/m\lambda})m\lambda}. \tag{4}$$

By combining equations (3) and (4) we have:

$$\rho \simeq 2m^2\int_m^\infty \frac{k^{-2}\lambda\Theta(\lambda)}{1 + k\lambda\Theta(\lambda)}dk = \frac{2e^{-1/m\lambda}}{1 - e^{-1/m\lambda}}. \tag{5}$$

The ρ is the stationary density of all infected nodes after time evolution of the stochastic cycle of SIS model. Equation (5) shows an explicit relationship between the infection density ρ and the effective infection rate λ, which can be used to evaluate different immunization strategies in the following section. The detailed calculations of $\Theta(\lambda)$ and ρ are shown in the appendix.

2.3 Immunization Strategies of P2P Networks

As discussed in section 2.1, the power-law networks exhibit different behaviors under random failures and intentional attacks, from which two intuitive immunization strategies may be regarded as randomized and degree-based immunizations. In the randomized immunization strategy, a proportion of nodes randomly chosen in the network are immunized, and these immune nodes will not be infected and do not spread the virus to their

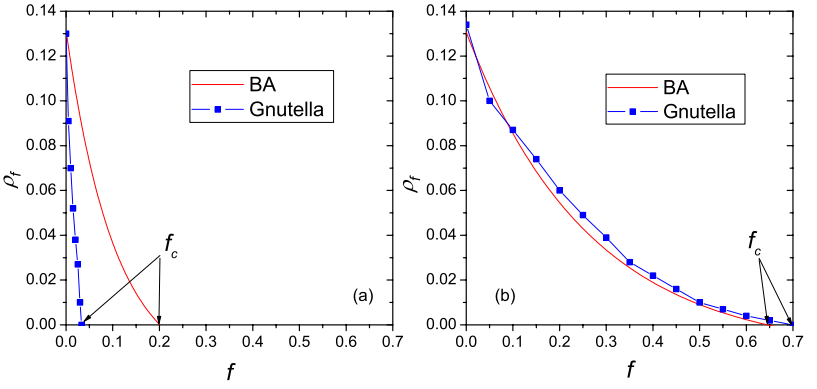

Fig. 2. Results for randomized and degree-based immunization measured by the density of infected nodes ρ_f as a function of the fraction of immune nodes f

neighbors. Accordingly, in the degree-based strategy, nodes are chosen for immunization if their degrees are greater than a predefined value.

Let us illustrate an example of installing a distributed intrusion detection system or a distributed firewall in a P2P system according to the two strategies. In both of the immunization strategies, the spreading dynamical properties can be considered as follows: suppose that a proportion of nodes are infected in the initial state, the immune nodes (intentionally protected by the IDS or firewall) in the system do not transmit viruses as if the links to their neighbors were eliminated, and the non-immune nodes spread viruses to their neighbors, but at the same time these nodes are cured with probability δ such as some nodes may install their personal antivirus program (not the same as the IDS or firewall mentioned above) or update the operating system. After a long time evolution, according to the mean-field theory, the system comes into balance between the infected and healthy nodes. In such a process, the two strategies have remarkably different impact on the density of infected nodes at the critical point of balance.

In the randomized case, for a fixed spreading rate λ, defining the fraction of immunized nodes in the network as f, we get the effective spreading rate $\lambda(1-f)$, and substitute it into equation (5), we obtain

$$\rho_f = \frac{2e^{-1/m\lambda(1-f)}}{1-e^{-1/m\lambda(1-f)}}. \tag{6}$$

Clearly, in the case of degree-based immunization, we can not use equation (5) to deduce an explicit formula as in the randomized case, but we will use simulations to compare the difference between the theoretical BA model and the real data of Gnutella network.

Our simulations are implemented with a fixed spreading rate $\lambda = 0.15$, the smallest node degree $m = 3$ and the number of nodes $N = 34206$ the same as the real data of the topology collected from Gnutella network [6]. Initially a proportion of healthy nodes are infected in the network. In Fig.2 (a), we plot the simulation results of degree-based immunization for BA network (line) and Gnutella network (square-line). As the

increasing of f, ρ_f decays much faster in Gnutella network than in BA model, and the linear regression from the largest values of f yields the estimated thresholds $f_c \simeq 0.03$ in Gnutella network, $f_c \simeq 0.2$ in BA network. The value of f_c in Gnutella network indicates that the Gnutella network is very sensitive to the degree-based immunization, and the immunization of just a very small fraction (3%) of nodes will eradicate the spreading of virus. On the other hand, in Fig.2 (b), the simulation results of randomized immunization are plotted for Gnutella Network (square-line), which is in good agreement with the theoretical prediction (line) by equation (5), except for a larger value of $f_c \simeq 0.7$ compared with the value $f_c \simeq 0.64$ of BA network. Based on the analysis above, it is evident that the degree-based immunization is really better than randomized immunization, which also inspires us designing an efficient immunization algorithm for P2P networks in the next section.

3 Efficient Immunization Algorithm for P2P Networks

Since the degree-based immunization is more effective than randomized immunization, one problem arises naturally that how efficiently we can inform other nodes when one node finds a virus in real networks. Returning to the example discussed in section 2.3, we consider that in a distributed IDS or firewall system for P2P network, if one node detects an intrusion or a virus, how can it transfer the information to others to update their local database in an efficient way? An intuitive solution is that it transfers the update information by visiting the neighbor with the highest degree, followed by a node with the next highest degree, since the immunized nodes are always with high degrees. In such a way, one can walk down a degree sequence all having high degrees. First, we formulate the highest degree in the network as a function of the network size.

Generally, the highest degree k_{max} of a node in a network depends on the size of the network. In [2], Aiello *et al.* assumed that the highest degree was approximately the value above which there was less than one node of that degree in the network on average, i.e, $np_k = 1$. This implies that for the power-law degree distribution $p_k \sim k^{-\gamma}$, $k_{max} \sim n^{1/\gamma}$. However, this assumption is not accurate in many real networks, where there are nodes with significantly higher degrees than this in the network.

Given a specific degree distribution p_k, the probability that there are m nodes with degree k and no nodes with higher degree is

$$p_{k_{max}} = \binom{n}{m} p_k^m (1 - P_k)^{n-m}, \tag{7}$$

where $P_k = \sum_{k'=k}^{\infty} p_{k'}$ is the cumulative probability distribution and n is the number of nodes in the network. Hence, the probability Π_k that the highest degree in the network is k is

$$\Pi_k = \sum_{m=1}^{n} \binom{n}{m} p_k^m (1 - P_k)^{n-m} = (p_k + 1 - P_k)^n - (1 - P_k)^n, \tag{8}$$

and the expected value of the highest degree is $k_{max} = \sum_k k\Pi_k$.

The probability Π_k tends to zero for both small and large values of k. Thus, in most case, a good approximation to the mean value of the highest degree is given by the modal value. Based on equation (8), we find that the maximum of Π_k occurs when

$$\frac{d\Pi_k}{dk} = (\frac{dp_k}{dk} - 1)(p_k + 1 - P_k)^{n-1} + p_k(1 - P_k)^{n-1} = 0, \tag{9}$$

where $\frac{dP_k}{dk} = p_k$. Assuming that p_k is sufficiently small for $k \geq k_{max}$ that $np_k \ll 1$ and $P_k \ll 1$, we can write equation (9) as

$$\begin{aligned} \frac{dp_k}{dk} &= -p_k[(\frac{1 - P_k}{p_k + 1 - P_k})^{n-1} - 1] \\ &= -p_k[(\frac{1}{p_k + 1})^{n-1} - 1] \simeq -np_k^2, \end{aligned} \tag{10}$$

where $0 < (p_k + 1 - P_k)^{n-1} < (1 - P_k)^{n-1} < 1$.

For BA model, the probability distribution of degree is $p_k = 2m^2k^{-3}$. Substituting it into equation (10), we have

$$k_{max} \simeq \sqrt{\frac{2m^2n}{3}}. \tag{11}$$

As described above, the node that detected an intrusion or a virus transfers the information to other nodes through a degree sequence in which all the nodes have the highest degree. For simplicity, suppose that the degrees of the nodes in the sequence all approximate to k_{max}, then the number of steps needed to transfer the information in the network of size n is

$$s = \frac{n}{k_{max}} \simeq \sqrt{\frac{3n}{2m^2}}. \tag{12}$$

We performed simulations of the real data of Gnutella network with a power-law exponent $\gamma = 2.0$ [6], and compared the simulation results with the theoretical prediction of BA network in equation (12). The number of nodes range from $N = 10^3$ to $N = 10^4$.

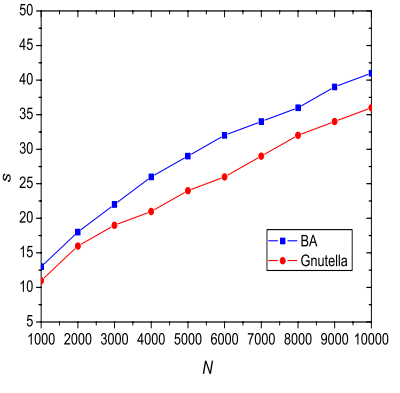

Fig. 3. The number s of steps needed to transfer information through high degree nodes as a function of the network size N

Fig.3 shows that the algorithm of transferring update information based on high degrees in Gnutella network is as efficient as the prediction of the theoretical BA model. We need only $s = 11$ steps to update all high degree nodes in Gnutella network with $N = 1000$ nodes, and $s = 36$ steps in Gnutella network with $N = 10000$ nodes. The discrepancy between the BA network and the Gnutella network is mainly due to the difference of power-law exponent of $\gamma = 3.0$ in BA network and $\gamma = 2.0$ in Gnutella network. Hence, in implementing a real distributed IDS or firewall system, we can update the global information effectively utilizing the highly connected nodes.

4 Conclusions

In this paper, based on the simple SIS model, we analyze the influence of virus spreading on P2P networks with two different immunization strategies namely randomized and degree-based immunization, and performe theoretical modeling and real data simulations. The results show that the degree-based strategy is more efficient than the randomized strategy, which also motivate us to design an effective information transferring algorithm for updating global virus databases. These methods are highly valuable in the implementation of real systems such as distributed IDSs or firewalls. As mentioned in the text, the immunization model is not flexible enough to analyze both randomized and degree-based strategies, hence, our future work is to improve the model to make it suitable for the analysis of both strategies.

References

1. L. Adamic, R. Lukose, A. Puniyani and B. Huberman, "Search in Power-Law Networks", *Phys. Rev. E*, Vol.64, 2001.
2. W. Aiello, F. Chung, and L. Lu, "A random graph model for massive graphs", *Proceedings of the thirty-second annual acm symposium on Theory of computing*, pp.171-180, 2000.
3. A. L. Barabasi and R. Albert, "Emergence of scaling in random networks", *Science*, Vol.286, pp.509, 1999.
4. B. Bollobas, *Random Graphs*, Academic Press, New York, 2nd ed, 2001.
5. A. Broder, R. Kumar, F. Maghoul, P. Raghavan, and R. Stata, "Graph structure in the web", *Computer Networks*, Vol.33, pp.309-320, 2000.
6. H. Chen, H. Jin, and J. H. Sun, "Analysis of Large-Scale Topological Properties for Peer-to-Peer Networks", *Proceedings of International Symposium on Cluster Computing and the Grid*, 2004.
7. O. Diekmann and J. A. P Heesterbeek, *Mathematical epidemiology of infectious diseases: model building, analysis and interpretation*, JohnWiley & Sons, New York, 2000.
8. M. Faloutsos, P. Faloutos, and C. Faloutsos, "On Power-law Relationships of the Internet Topology", *Computer Communications Review*, Vol.29, pp.251-262, 1999.
9. M. A. Jovanovic, "Modeling Large-scale Peer-to-Peer Networks and a Case Study of Gnutella", Master thesis, Department of Electrical and Computer Engineering, University of Cincinnati, 2000.
10. S. Milgram, "The small-world problem", *Psychology Today*, Vol.1, pp.62-67, 1967.

11. M. E. J. Newman, "The structure and function of complex networks", *SIAM Review* , Vol.45, pp.167-256, 2003.
12. M. Ripeanu, I. Foster and A. Lamnitchi, "Mapping the Gnutella Network: Properties of Large-Scale Peer-to-Peer Systems and Implications for System", *J. Internet Computing*, 2002.
13. R. P. Satorras and A. Vespignani, "Epidemic Spreading in Scale-Free Networks", *Phys. Rev. Lett*, Vol.86, pp.3200-3203, 2001.
14. S. Saroiu, K. P. Gummadi and S. D. Gribble, "Measuring and analyzing the characteristics of Napster and Gnutella hosts", *Multimedia Systems*, Vol.9, pp.170-184, 2003.

A Appendix. Calculations of $\Theta(\lambda)$ and ρ

$$\begin{aligned}
\Theta(\lambda) &= \frac{\sum_k kp(k)\rho_k(t)}{\langle k\rangle} \simeq \int_m^\infty \frac{2m^2k^{-3}k^2\lambda\Theta(\lambda)}{2m(1+k\lambda\Theta(\lambda))} \\
&= m\int_m^\infty \frac{\lambda\Theta(\lambda)}{k(1+k\lambda\Theta(\lambda))}dk \\
&= m\lambda\Theta(\lambda)\int_m^\infty (\frac{1}{k} - \frac{\lambda\Theta(\lambda)}{1+k\lambda\Theta(\lambda)})dk \\
&= m\lambda\Theta(\lambda)([\ln k]_m^\infty - [\ln(1+k\lambda\Theta(\lambda))]_m^\infty) \\
&= m\lambda\Theta(\lambda)(\lim_{M\to\infty} \ln\frac{M}{1+M\lambda\Theta(\lambda)} - \ln\frac{m}{1+m\lambda\Theta(\lambda)}) \\
&= -m\lambda\Theta(\lambda)\ln\frac{m\lambda\Theta(\lambda)}{1+m\lambda\Theta(\lambda)} \Rightarrow \\
-\frac{1}{m\lambda} &= \ln\frac{m\lambda\Theta(\lambda)}{1+m\lambda\Theta(\lambda)} \Rightarrow \Theta(\lambda) \simeq \frac{e^{-1/m\lambda}}{(1-e^{-1/m\lambda})m\lambda}.
\end{aligned}$$

$$\begin{aligned}
\rho &= \Sigma_k p(k)\rho_k \simeq \int_m^\infty \frac{2m^2k^{-3}k\lambda\Theta(\lambda)}{1+k\lambda\Theta(\lambda)}dk \\
&= 2m^2\lambda\Theta(\lambda)\int_m^\infty (\frac{1}{k^2} - \frac{\lambda\Theta(\lambda)}{k} + \frac{(\lambda\Theta(\lambda))^2}{1+k\lambda\Theta(\lambda)})dk \\
&= 2m^2\lambda\Theta(\lambda)([-\frac{1}{k}]_m^\infty - \lambda\Theta(\lambda)[\ln k]_m^\infty + \lambda\Theta(\lambda)[\ln(1+k\lambda\Theta(\lambda))]_m^\infty) \\
&= 2m^2\lambda\Theta(\lambda)(\frac{1}{m} + \lambda\Theta(\lambda)\lim_{M\to\infty}\ln\frac{1+M\lambda\Theta(\lambda)}{M} + \lambda\Theta(\lambda)\ln\frac{m}{1+m\lambda\Theta(\lambda)}) \\
&= 2m^2\lambda\Theta(\lambda)(\frac{1}{m} + \lambda\Theta(\lambda)\ln\frac{m\lambda\Theta(\lambda)}{1+m\lambda\Theta(\lambda)}) \\
&= 2m^2\lambda\Theta(\lambda)(\frac{1}{m} - \lambda\Theta(\lambda)\frac{1}{m\lambda}) \qquad (\textit{by substituting } \Theta(\lambda) \textit{ into the above equation}) \\
&\simeq 2m\lambda\Theta(\lambda) \qquad (\textit{the lowest order of } \lambda \textit{ is remained}) \\
&= \frac{2e^{-1/m\lambda}}{(1-e^{-1/m\lambda}).}
\end{aligned}$$