TELKOMNIKA, Vol. 11, No. 10, October 2013, pp. 5815 ~ 5823
ISSN: 2302-4046

A Safety Algorithm of P2P Routing based on Multiple-Encryption Detecting Technology

Chuiwei Lu*, Xianhao Miao, Zhiyuan Liu
Computer School, Hubei Polytechnic University, Huangshi 435003, Hubei, China
*Corresponding author, e-mail: 79738834@qq.com

Abstract

The nodes can freely join or leave the P2P network, which will lead to much false routing information that can cripple the performance of P2P network. Many hackers also utilize the weaken point to attack the P2P network. We propose a safety routing algorithm for P2P network to resist the routing attack. The algorithm adopts the multiple-encryption detecting technology. The node which launches the communicating connection will periodically detect every node in its routing path by sending some multiple-encryption detecting packets. By the responding message of the detected nodes, the malicious or disable nodes in its routing path will be accurately located and kicked out of the routing table. Simulation experiments demonstrate the algorithm can effectively improve the safety of the P2P routing and topology stability of the P2P network.

Keywords*: Routing Attack, P2P Network Safety, Multiple Encryption, Active Detection*

1. Introduction

Nowadays, the technologies of Internet and computer have obtaind great improvements. There are huge data resource and services resources dispersedly stored in Internet. The P2P network provides a probability to integrate these resources and build a new application service with a cheap and reliable method. P2P network is a kind of logic network that overlap in physical network, it can easily build many private resource-share networks on Internet. Though the node in P2P network is usually anonymous, and can join or exit the network freely, the P2P network is easily to be attacked by malicious nodes. The stable and accurate routing is the vital foundation for P2P network to work effectively. Hence, most of the attacks to P2P network are aiming at the P2P routing, it is because the breakdown of P2P routing may result in the collapse of the whole P2P network. The routing algorithm of P2P network has been the important part in P2P protocols. Therefore, many researchers and scholars have focused their interests on robust and reliable routing algorithm of P2P network.

There are numerous researches on the multiple aspects of P2P routing attacks and defense [1-8]. Paper [9] proposed a dynamic source routing protocol and security extensions to adapt and modify the inherent principles of the P2P routing security concept, and verified the applicability in a real world system. Paper [10] proposed a new trusted routing mechanism based on trust degree for P2P network which can make honest nodes play a more important part in the network, and make malicious nodes to the edge of network. Some famous attacks, such as bad mouthing attacks, self-fault problems and conflict behavior attacks, can be effectively depressed by the routing mechanism. Fujii T [11] proposed a security model for evaluating security level of routing protocol. The model defines the concept of Regular Path, and uses it as the indicator of security level. Then, some famous representative routing protocols are made comparisons from the aspects of security level and implement level based on proposed model. Paper [12] proposed an efficient routing strategy designed to control the routing path while reducing the normal routing latency. Combined with a peer-ID based signature scheme, the routing strategy can offer the initiator of each query to identify malicious nodes. This strategy also own tracer function, a key feature of routing scheme from other protocols is that alternate routing is constructed only detecting malicious nodes, which highly raise the security level of P2P routing. Lin Wang [13] et al analyzed the vulnerable aspects of

Received April 15, 2013; Revised July 3, 2013; Accepted July 16, 2013

current P2P networks, and pointed out that most structured P2P protocol maintains flaws, then proposed a DHT-based universal detecting and defending model.

The paper proposed a safety algorithm of P2P routing based on multiple-encryption detecting technology: SAP2PRMEDT. It periodically detects the attribute of relaying nodes in its routing path, and finds malicious or instable nodes and excludes them from its routing table. This measure can establish a safe and reliable routing path for every P2P connection, and help to optimize the performance of P2P network and improve its working efficiency

2. Mathematical Analysis of Attacks Factors

From the analysis in the previous section, there are several main factors that relate to the effect of routing attacks. For quantitative analysis, we assume that the P2P network nodes and resources are evenly distributed, the node degree of each node is equal, the threshold value of routing reset times is λ, the total number of P2P network nodes is N, in which numbers of malicious nodes are M, the average routing path length of each communication channel is L. This section will get a theoretical analysis and get the minimum average times n in destroying a routing path. Since the malicious nodes is randomly attack any possible routing paths, only when the same communication channel C_{sd} is destroyed more than λ times, the channel is completely disable.

For simplicity, we assume that there is only one communicating process between every two nodes in P2P network, i.e. a routing path, so there are $\frac{N(N-1)}{2}$ routing paths in N nodes. If these M malicious nodes are uniformly distributed in M routing path, they can destroy M routing path in one round attack. If these M malicious nodes are distributed in same routing path, they can only destroy one routing path in one round attack. The two cases above are two extreme cases, usually the number of the routing path that is destroyed by malicious nodes distribute in [1, M]. In addition, the longer the length of the routing path, the more the malicious nodes sneaked into the path, thus the possibility that the routing path is destroyed becomes greater. Overall, we assume the number of the destroyed routing path in one round attack is $\frac{1+M}{2}\ln L$, and this value is relatively compromised. It can be deduced that the possibility p that the malicious nodes damage the same routing path in round is $\frac{1+M}{N+(N-1)}\ln L$

Each attack that the malicious nodes launched is a random and independent event, and all for one purpose, which fits the feature of independent distributed central limit theorem. We can use this theorem to quantitatively describe the attacking event. Assume X is the number of time of the same routing path that was damaged by n rounds of attacks by the malicious nodes, then from the central limit theorem there is $X \sim b(n, p)$. When the same routing path is damaged more than λ times before the end of the communication, the path is completely disable.

$$P\{x \geq \lambda\} = \sum_{k=\lambda}^{\infty} C_n^k p^k (1-p)^{n-k} \tag{1}$$

We believe that when $P\{X \geq \lambda\} \geq 0.99$ then the damage can be identified success, thus to anti-derivate the average minimum times n that to destroy a routing path.

The formula 1 is fairly complex and difficult to calculate, so we use De Moivre-Laplace theorem to get its approximate value, the transformed formula is shown as follows.

$$\begin{aligned} P\{X \geq \lambda\} &= P\left\{\frac{X-np}{\sqrt{np(1-p)}} > \frac{\lambda-np}{\sqrt{np(1-p)}}\right\} \\ &\approx \int_{\frac{\lambda-np}{\sqrt{np(1-p)}}}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}t^2 dt} \\ &= 1-\Phi\left(\frac{\lambda-np}{\sqrt{np(1-p)}}\right) \end{aligned} \tag{2}$$

The Φ in formula 2 is a standard normal distribution function, in the paper, its value should be equal or lesser than 0.01.

$$\Phi\left(\frac{\lambda - np}{\sqrt{np(1-p)}}\right) \le 0.01 \tag{3}$$

We replace *p* in formula 3 with its true value, and get below formula.

$$\Phi\left(\frac{\lambda - n\dfrac{(1+M)\ln L}{N(N-1)}}{\sqrt{n\dfrac{(1+M)\ln L}{N(N-1)}\left(1-\dfrac{(1+M)\ln L}{N(N-1)}\right)}}\right) \le 0.01 \tag{4}$$

If assume Φ(Z)=0.01, we can find Z = −2.33 from standard normal distribution table, and get below formula.

$$\frac{\lambda - n\dfrac{(1+M)\ln L}{N(N-1)}}{\sqrt{n\dfrac{(1+M)\ln L}{N(N-1)}\left(1-\dfrac{(1+M)\ln L}{N(N-1)}\right)}} = -2.33 \tag{5}$$

In the case that *N, L, M, λ* are all known, formula 5 become a quadratic equation of variable *n*, through which we can calculate the value of *n*.

$$\begin{aligned} n &= \frac{2\lambda + 5.36(1-p) + \sqrt{29.52(1-p)^2 + 21.85\lambda(1-p)}}{2p} \\ p &= \frac{(1+M)\ln L}{N(N-1)} \end{aligned} \tag{6}$$

Using the formula 6, we can calculate the average minimum number of attack times *n* that to destroy a routing path, and also easily analysis the effect that the number of malicious nodes, the average length of routing path to the process of destroying a routing path.

Associate with formulas above, we can derive the minimum value of the attack times *n* in multiply cases. Supposing there are 10^5 nodes in P2P network, *λ* is 6, according to the theory and the formula above, when the proportion of malicious nodes in P2P network is 5%, 10%, 25%, 35% respectively, we approximatively calculate the relationship between the minimum attack times *n* and the length of routing path *L*.

From Figure 1, we can find that the length of routing path hasn't vital affect to the routing security, but the proportion of malicious nodes has. The discovery becomes the important foundation that we design SAP2PRMEDT algorithm.

3. The Design of Sap2prmedt Algorithm

Comprehensively analyzing the characteristics of P2P routing attacks, we consider that to exclude malicious nodes from P2P network is fundamental counter-measure to those attacks. So the emphasis of SAP2PRMEDT algorithm lies in attribution detection of P2P nodes with special encryption message.

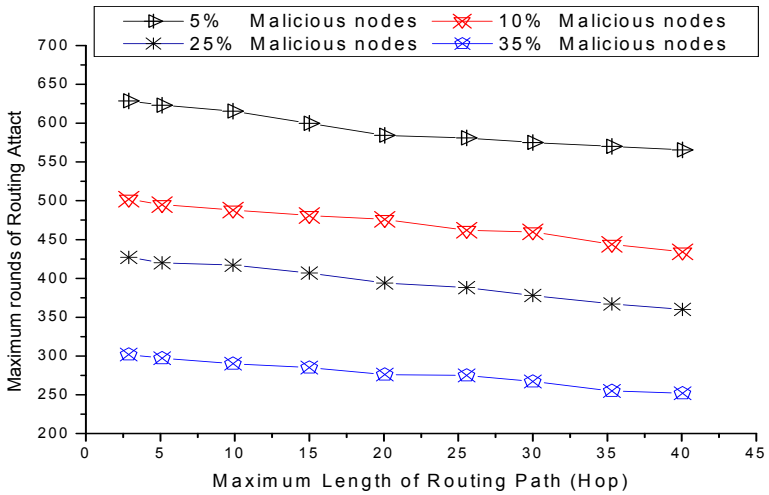

Figure 1. The length of routing path and the proportion of malicious nodes to the affect of success rate of routing attack

3.1. Process of the Initialization of Routing Path

When a node want to communicate with another node, there must be built a routing path between them first. The initial establishment of routing path does not adapt the active detecting method, because nodes need to consume certain system resources for detection, it will cost a lot to establish a routing path if uses the method of one by one detection, and it will also cost longer time, which is no good for communicating. After the establishment of routing path, the work of attribution detection of relaying nodes is merged in the routine of topological maintenance of the P2P networks. The means is efficient but at very low cost, and will not bring significant negative impact in P2P communicating. The specific way to build a routing path is described below.

The initial nodes of the communication send the request of connecting to some nodes in the routing table, when these nodes receive the request, according to the P2P handshake protocol, they forward this request to their neighbor nodes at some forwarding probability ξ. These neighbor nodes forward the request in the same way until it reaching the correct target node. The forwarding probability ξ has an inverse proportion with the load statue of the nodes. According to the method of establishing connection above, the last initiator nodes may receive several reachable routing paths to choose from. Since the overlong routing path will bring the negative impact to the communicating efficiency and the routing security, we prior choose the routing path with the least number of hops. The Small World Theory indicates that through average 6 persons, we can find the one that we want. Thus we limit the max length of the routing path be or less than 6. If there still have several routing paths are under the condition, we then choose the routing path whose total time delay is the least, that is $Min\left(\sum_{i=1,j=2}^{i=M-1,j=M} RTT_{ij}\right)$, the RTT_{ij} is the delay between two nodes.

When the routing path is chosen, the initial node of the communicating first need to record every relaying node's IP address, port number and P2P ID, and then to consult with every relaying node to get a set of asymmetric keys, and lastly to save the public keys from them for the future detection. To the malicious nodes, including the zombie nodes which were controlled by the virus, if each normal P2P network activities shows refuse or even destruction, they will be kicked out from the P2P network due to the early exposing. Thus those nodes perform their malicious activities on certain probability, like sometimes good but sometimes evil. Based on the discussion above, the pseudo-code of initially creating a routing path is as follows.

$$P_s \xrightarrow{Build \quad Connection} P_d$$

Repeat following routing search course

$P_s \xrightarrow{\xi} P_1, P_3, P_6 \ldots\ldots$

$P_1 \xrightarrow{\xi} P_4, P_5, P_2 \ldots\ldots$

$P_4 \xrightarrow{\xi} P_7, P_9, P_{10} \ldots\ldots$

……..

Until packet arriving at P_d

Obtain multiple routing path between P_s and P_d

Put those routing path into array $R\{r_1, r_2, \ldots, r_n\}$

If routing length of r_i > 6 then

  Kick r_i out of R // i =1,2,3….

End if

R = Sort by $Min\left(\sum_{i=1,j=2}^{i=M-1,j=M} RTT_{ij}\right)$ with low to high

For k = 1 to n do

  Peers in r_k send their <ID_k, IP_k, $Port_k$> to P_s

  P_s consult Asymmetric Key Pair with every peer in r_k and occupy their Public Key

  If above course is triumphantly executed then

    Adopt r_k as the ultimate routing path

    Exit cycle

  Else

    Discard r_k

    $k = k + 1$

  End if

End for

3.2. Detection Technology of Multiple Encryptions

When the routing path is established, before the end of the communicating, the initiator will use the Ping to send specific encrypted information to the receiver. Along with the Ping command, this information is forwarded to the target node upward node by node. Every time to pass a relaying node, part of the information will be modified and signed for recording the behavior of nodes. If a problem occurs to a relaying node, the relevant problem information will be send back through the Pong command to the initiator. Then the initiator will analyze the returned information and to determine the nature and location of the problem. The result of the analysis will determine whether to put the reported node into the malicious nodes' table, and to determine neither to reset the routing path or retest.

To prevent the malicious node to tamper the information along with the Ping and Pong command, we adopt the Public-key nested encryption to encrypt the transmitting information to shut down on this malicious behavior. The specific method is as follows.

Assuming a routing path is R = <P_0, P_1, P_2, P_3,…, P_m>, P_0 is the initiator. The detection begins, the P_0 randomly generates a positive integer X, and nested encrypt it with all the <*IP address, Public key*>, then to add them to Ping command to forward to the next hop P_1 node. The encrypting format of detecting information packet is as follows.

$$K_1(X, A_2, K_2, K_2(A_3, K_3, K_3(....), ..., A_{m-1}, K_{m-1}, K_{m-1}(A_m, K_m))....))$$

This is an "onion" type of nested encryption data structure, in which the K_i is the number i relaying node's public key, the A_i is the number i relay node's IP address, and 1<i<m. The follow relay nodes have to use their own private key to peel the "onion" data layers by layers if they want to use it.

In order to protect the authenticity of the random number X, the P_0 need to sign the X by its private key S_0, and forward the result $S_0(X)$ to P_1. The node P_1 received those information, it decrypts the data packet by its own private key S_1, however, it can only get and use the data X, A_2, K_2, because the last half of the information was encrypted by the P_2 node's public key K_2. Thus the P_1 is not able to decrypt this information, so that even if the P_1 happens to be malicious node, it still cannot tamper the detecting information of the follow nodes.

According to the address A_2 of the node P_2, the node P_1 encrypts $(X+1)$ by K_2 and then forward it to P_2, at the same time, P_1 has to use its own private key S_1 to sign the $(X+1)$ and send the generated data $S_1(X+1)$ together to P_2. The signature has important use on detecting

the location of malicious nodes. The node P_1 has to deliver the decrypted information $K_2(A_3, K_3, K_3(A_4, K_4, K_4(\ldots), \ldots, A_{m-1}, K_{m-1}, K_{m-1}(A_m, K_m))\ldots)$ to P_2. Then P_2 decrypts this segment of information by its private key S_2 and delivers $K_3(A_4, K_4, K_4(\ldots), \ldots, A_{m-1}, K_{m-1}, K_{m-1}(A_m, K_m))\ldots)$ to P_3. At the same time, P_2 has to encrypt $(X+2)$ by public key K_3 and sign $(X+2, S_1(X+1))$ by private key S_2, then forward those results to P_3. To repeat the operation given above until arrive at the target node P_m. If the routing path is expedite, then the X will turn into $X+m$ when the detecting information arrives. Thus the integer m actually represents the number of routing pops. In addition, the signature information will also turn into the format below.

$$S_{m-1}(X+m-1, S_{m-2}(X+m-2, \ldots S_1(X+1, S_0(X))\ldots))$$

After receiving the signature information ensuring its correctness, the node P_m will use its private key S_m to sign the information for the last time, the format is as follows

$$S_m(X+m, S_{m-1}(X+m-1, \ldots S_1(X+1, S_0(X))\ldots))$$

Last, the node P_m will attach the signature information described above to the Pong command and return it to the initial node P_0 through the same routing path. The P_0 uses the corresponding public key to decrypt the signature information, and to get its data $(X+m)$. If the value of m is the same with the value of its saved length of routing path, we then consider there is no malicious node on the routing path.

If a certain node P_{i+1} did not return the feedback information in three times detection, the node will be regard as a malicious or disabled node.

If the initiator P_0 receives wrong feedback detection information, there may be two cases, one is that the feedback information damaged during the transmission process due to some random and irresistible faults. However, the case is also a very small probability event, and has little affection to our routing detection system. Another case is that the malicious nodes in the routing path tampered the data. For example, when a malicious node produce a wrong value: $(X+i)'$, then transmits it to next hop node, which will cause the following nodes in routing path all fail in calculating the i. Since the parameter i represent the routing hop and the routing order, the P_0 need to decrypt all the i from all the $S_i(X+i)$ which were send by every relaying node. In normal situation, the sequence of the value of i is a string of continuous integer beginning at 1. If the value of i is discontinuous, the first break point is malicious node. What we need to do is just to avoid the node when resetting the routing.

The below Figure describe the process that the SAP2PRMEDT algorithm detect malicious node and reset routing path in P2P network.

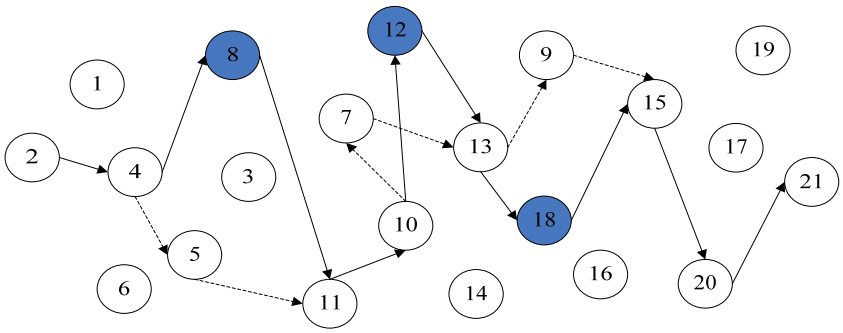

Figure 2. Reset routing path to exclude malicious nodes

In Figure 2, the node P_2 is the communicating initiator, node P_{21} is the receiver, the initial routing path is $R=<P_2, P_4, P_8, P_{11}, P_{10}, P_{12}, P_{13}, P_{18}, P_{15}, P_{20}, P_{21}>$, the length of path is 10. After using SAP2PRMEDT algorithm to detect all the relaying nodes in the routing path, the system finds that the member P_8, P_{12} and P_{18} in the routing path are bad members, then the initiator resets the routing path, and creates a new path $R'=< P_2, P_4, P_5, P_{11}, P_{10}, P_7, P_{13}, P_9, P_{15}, P_{20}, P_{21}>$. It reveals that the new routing path has avoided the bad members, which make the attacking behavior by malicious nodes become more difficult thereby enhancing the security of routing.

4. Simulation Experiments and Analysis

The simulation experiments were performed in the PC that the CPU is P4 3.2GHz and the memory is 2G, and the OS is Fedora Linux 9.0, and the simulating software is P2Psim3.5. This software is modularly simulated which is specifically designed for the P2P networks, and its function is powerful and rich and is easy to use. At the same time, the P2Psim3.5 integrates the Chord, CAN, Koorde and dozens other mainstream P2P protocols, which makes it the chief choice for stimulating in the P2P research field.

To compare the effect of developed SAP2PRMEDT algorithm, the experiments adopted Chord and Koorde algorithm as the reference. The Chord is the most common used ring network structure while the Koorde is the famous P2P network protocol based on the graph theory. These two as the reference can well reflect the effect after improving. These three algorithms all simulated 10^4 nodes and the experiment were done for three times. The purpose of the first time of the experiment is to study the effect that the SAP2PRMEDT algorithm to detect the malicious nodes in P2P networks. The experiment set up the proportion of malicious nodes as 30% and distributed evenly. The purpose of the second experiment, while the proportion of malicious nodes in P2P system gradually increases, is to study the change of the needed average attacking number of time to successfully destroy a routing path. And the purpose of the third experiment is to study how the length of routing path effects the number of average attacking time, while one of the experimental conditions are the proportion of malicious nodes in the P2P network is fixed at 20%.

It can be seen from the Figure 3 that the detecting speed of SAP2PRMEDT algorithm towards malicious nodes is relatively satisfied, the successful detecting rate after the P2P system running for 20 minutes significantly boosted. After 80 minutes past, the system has detected over 90% of malicious nodes, and after 110 minutes, almost all the malicious nodes have been detected. Thus, the effect of SAP2PRMEDT detection is significantly nice, which substantially increased the routing security in P2P networks.

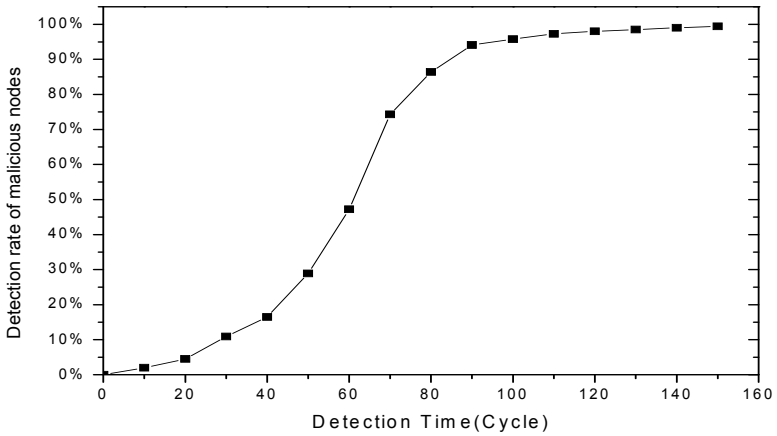

Figure 3. The malicious nodes' detection rate of SAP2PRMEDT

It can be seen from the Figure 4, with the proportion of malicious nodes in the P2P network increases, the average attacking number of time on successfully destroying a routing path declines in all the three algorithms, especially to the Koorde, coming after the Chord. The decrease rate is relatively slow to the SAP2PRMEDT, and when the proportion of malicious nodes have reached 23%, the rate of decline leveled off, which showed the excellent defense to the attack. To consider on the other side, with the proportion of malicious nodes declined, malicious nodes in the SAP2PRMEDT had to attack much more times, whose number of attacking time was much larger than the other two algorithm, which has increased the difficulty for malicious nodes to attack, so that to enhance the security of P2P system.

It can be seen from the Figure 5 that with the increase of length of routing path, the average attacking number of time on successfully destroying a routing path declines in all the

three algorithms, which might because that longer routing path is much easier to be attacked than shorter routing path. And the longer routing path takes much longer time to rebuild itself after destroyed, subsequently, within the system default time period, the number of routing path resetting will decline, and the rate of decrease is relatively smooth. This indicates that after introducing the active detecting mechanism, in the P2P system, the ability of anti routing attacking has been strengthened significantly, thereby enhancing the security of the system.

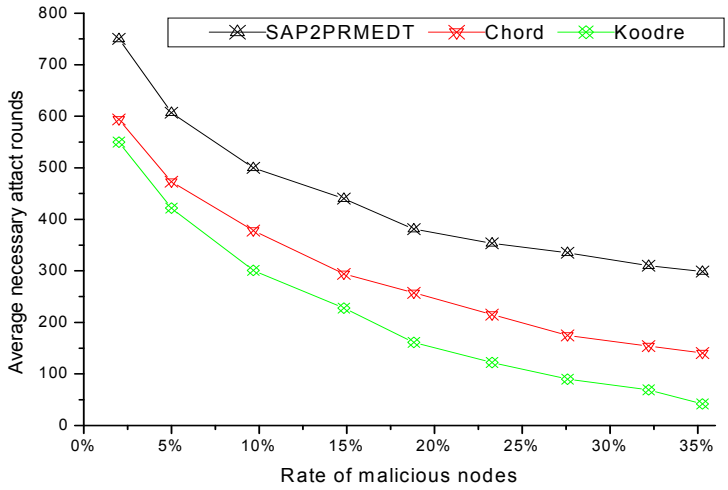

Figure 4. The affect of malicious-node rate to routing-attack rounds

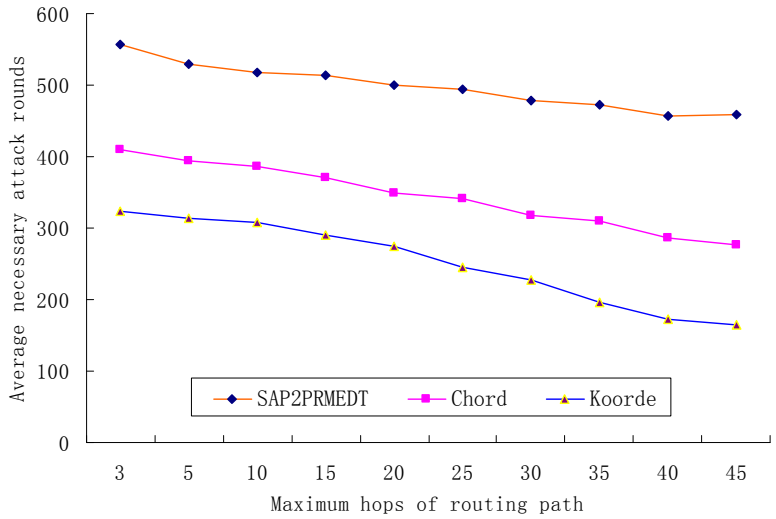

Figure 5. The affect of routing-path hops to routing-attack rounds

5. Conclusion

This paper, in connection with the P2P routing attack problem, has proposed SAP2PRMEDT algorithm which is able to detect the malicious nodes and instable nodes in the routing path within a very short period, and exclude these nodes by rebuilding a new routing path. The algorithm highly enhances the security performance of the P2P network. In addition, in the premise of maintaining the anti-attack ability, the algorithm also can optimize the maximum length of the routing path, which will reduce the communicating delay and raise the communicating performance.

Acknowledgements

Mr. Xianhao Miao and Zhiyuan Liu have done a lot of valuable contributions to the paper, and the Xianhao Miao is the corresponding author of the paper. The paper is sponsored by Key Project of Education Committee of Hubei Province (No. D20114401), and Teaching Research Project of Hubei Polytechnic University (No. 201020 and No. 2007026)

References

[1] Cholez Thibault, Chrisment Isabelle, Festor Olivier. Detection and mitigation of localized attacks in a widely deployed P2P network. *Peer-to-Peer Networking and Applications*. 2013; 6(2): 155-174.
[2] Liu Xin, Datta Anwitaman. Attack resilient P2P dissemination of RSS feed. *Peer-to-Peer Networking and Applications*. 2011; 4(3): 309-324
[3] Lu Songnian, Zhao Dandan, Zhang Aixin. An anti-attack model based on complex network theory in P2P networks. *Physic A: Statistical Mechanics and its Applications*. 2012; 391(8): 2788-2793
[4] Al-Duwairi Basheer, Mustafa Abdul-Raheem Masoud. A novel mechanism to counter P2P-based DDoS attacks. *International Journal of Internet Protocol Technology*. 2010; 5(2): 55-64
[5] Sato Fumiaki. Estimation of trustworthiness for P2P systems in a collusive attack. *International Journal of Web and Grid Services*. 2008; 4(1): 24-34
[6] Gu Jabeom, Nah Jaehoon, Kwon Hyeokchan. Defense against identity attacks in P2P networks. *Transactions on Information and Systems*. 2008; 91(4): 1058-1073
[7] Chuiwei Lu. *Research of P2P Routing Security based on Positive Detection Mechanism*. Proceeding of IEEE International Conference on Networking Architecture, and Storage. Zhangjiajie, China. 2009; 36-41
[8] Zheng X, Oleshchuk V. *Trust-based framework for security enhancement of P2PSIP communication systems*. Proceeding of International Conference on Internet Technology and Secured Transactions, London UK. 2009; 1-6
[9] Jie Xu, Xiaolin Liu, Keping Long. *A routing mechanism based on trust relationship for P2P networks*. Proceeding of IEEE International Conference on Communication Technology. Nanjing China. 2010; 1176 - 1179
[10] Kraxberger, S. *Secure Routing Approach for Unstructured P2P Systems*. Proceeding of IEEE International Conference on Emerging Security Information, Systems and Technologies. Athens, Glyfada. 2009; 210-216
[11] Fujii T Yizhi Ren, Hori Y, Sakurai K. *Security Analysis for P2P Routing Protocols*. Proceeding of IEEE International Conference on Availability, Reliability and Security. Fukuoka. 2009; 899 – 904
[12] Xu Xiang, Tan Jin. *Efficient Secure Message Routing for Structured Peer-to-Peer Systems*. Proceeding of IEEE International Conference on Networks Security, Wireless Communications and Trusted Computing, Wuhan China. 2009; 354 – 357
[13] Lin Wang. Attacks against Peer-to-peer Networks and Countermeasures. *TKK Seminar on Network Security*. 2006; 21(12): 43-51