A Security Routing Algorithm of P2P Network Based on Asymmetric Nested Encryption

Chuiwei Lu
Computer School, Huangshi Institute of Technology
Huangshi, China
Lcwzm@tom.com

Zhengbing Hu
Department of Information Technology, Huazhong Normal University
Wuhan, China

***Abstract*—correct routing is a key factor to maintain the stability and efficiency of P2P network. There are many types of attacks aiming at P2P routing, which seriously threats the security of P2P network. We proposed an improved routing algorithm based on asymmetric nested encryption detection. The algorithm can periodically inspect every node in routing path, and eliminate bad nodes and instable nodes in routing path with little time, so as to raise the security capability of P2P routing in high. Simulation experiments demonstrate the improved algorithm can effectively enhance the routing security of P2P network.**

Keywords- Routing security; P2P network; Asymmetric nested encryption; Active detection

I. INTRODUCTION

Routing algorithm is an important part of the P2P networks, and has a critical role. A number of important functions in P2P network are required the support of the correct routing. Routing destruction would be tantamount to undermining the entire P2P network, which resembles "decapitation-war" in modern war, and its destruction effect is equivalent to attack one point but destroy all the body, and the cost is very low. Although there are many forms to attack the P2P networks, but most of the attacks is to against the P2P routing. It can be said, if there is no safe and reliable routing protocol, the P2P technology will cease to exist. Therefore, P2P routing is of great significance, and has become a research hot-point in many institutions and universities. So far, P2P routing technologies have made considerable progress and have formed three kinds of mainstream P2P routing model: centralized directory type, unstructured type with flooding-based, and structured type with DHT-based.

II. RELATED WORK

Miguel C[1] studied the attack behavior that specific destruct routing messages, and proposed an efficient control strategy. Mudhakar S [2] pointed out that the malicious nodes attack P2P routing is the biggest threat to structured P2P networks, and the first to propose a quantitative analysis method to estimate the harm extent. Ioannis [3] proposed an improved protocol to protect the routing information is accurate reach, and to publish the malicious nodes by block their communication. Peng W[4] proposed a novel DHT-based routing improved protocol: Myrmic, it can maintain more than 60% routing success rate when P2P network is violently flapped. Shruti P [5] proposed a cumulative model to estimate the reliability of routing, and uses a known as the "Pruning" technology to restrict excessive number of routing messages in the P2P networks, which can find a safe, reliable and low message-complexity routing path. In 2008, Sheila B[6] have done a more comprehensive statistics and analysis aim at the threat of P2P networks, and using DHT technology to form a new safety communications mechanisms. Krishna P[7] defined an general identity attack model, and proposed a lightweight security system with the detection, tracking malicious falsification of identity nodes, and used redirect data transmission solution to solve the confusion resulted by the identity attacks. Jayanth K[8] analyzed the negative effects of the worm in file-share P2P system, and made the quantitative analysis on P2P worm attack from the characteristics of P2P traffic, and proposed a P2P worm defense method in structured P2P networks. Naoum N[9] designed a test program to launch DDoS attacks on P2P network. It uses two kinds of "poisoning" approaches on the index table of resource information and the routing table, which launch random attacks on the structured P2P file-share network.

III. ANALYSIS OF ROUTING MECHANISM IN P2P NETWORK

Structured P2P networks can be abstracted as a directional graph $G = <P, E>$, all the nodes in P2P network are mapped to a single point of the figure and use P_i represented, Where $0 < i < n$, n represents the total number of nodes in the P2P network. $C_{ij} = <P_i, P_j>$ represents the relationship communication between members. When the node P_s needs to communicate with another node P_d, P_s will select a suitable active neighbor $P_k \in P$ in its routing table. And then P_s will send the communication needs C_{sd} to P_k. After P_k receipted, it will prefer an active neighbor node P_g according to its routing table information and forward C_{sd} out, and record simultaneously transfer path. The same operations have carried out several times then the C_{sd} will be forwarded to the P_d finally. So as to establish a routing path formed by a number of intermediate nodes and it the relay communication achieve between the initiator and the recipient. Routing paths, once established, will be maintained till the end of communication cycle automatically removed. In this transmission path the packet source address, destination address will be rewritten many times. That made a malicious node can not detect the true source, mesh nodes from the intercepted packet. Thereby protecting the node privacy, and hide the transmission path.

978-0-7695-4074-0/10 $26.00 © 2010 IEEE
DOI 10.1109/ITCS.2010.12

This approach enhances the security of routing a certain extent, has been widely used in the current P2P network resource query process, some use of the resource transfer process. Fig.1 represents a P2P network topology that contains three routing paths inside.

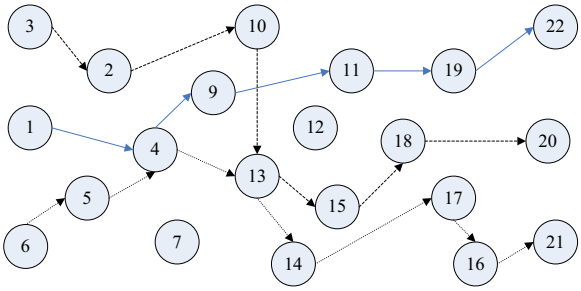

Figure 1. Routing paths in the P2P network

In fig.1, the node 1, 3, 6 are communication initiators, the node 20, 21, 22 for communication receivers. They generally need to go through several intermediate nodes relay to complete the communication task. The arrow diagram of communication line guide is called the routing path. When there is node failure or withdrawal in the routing path, the routing path has been destroyed. It is necessary to reconstruct a new path, known as the routing reset.

Based on the above characteristics of the routing topology to abstract a P2P routing path R into the form

$$R = <P_s, P_1, P_2, P_3, \cdots, P_M, P_d> \quad (1)$$

In formula 1 P_s represents the communication source nodes, P_d represents the communication purpose nodes, P_1, P_2, P_3 *and* P_M denote the relay nodes. In fact it can be seen that the subscript M represents the number of intermediate nodes. We call it as the path length or the network diameter.

IV. SECURITY ROUTING ALGORITHM BASED ON ASYMMETRIC NESTED ENCRYPTION

Comprehensive analysis and reasoning in the above sections, we find that the characteristics of P2P routing attacks and the problems in the current P2P routing mechanism. We proposed a new routing security algorithm: RAPD(Routing Algorithm based on Positive Detection). The algorithm can find possible bad nodes in the routing path, using routing-reset to avoid them, and to exchange the fail nodes information with the neighbor nodes, so as to achieve the security of routing path and information transmission.

A. First establishment of the P2P routing path

When a node wishes to communicate with another node, firstly it needs to establish a routing path between them. The initial establishment of routing paths without using positive detection mechanism, because node detection requires the consumption of certain system resources, individually detection method to establish the routing path cost too much. When the number of nodes is large, and the degree of nodes is high, it will affect the operational efficiency of P2P networks. Therefore, we mainly use the node's own routing table information, bad node list information, to establish routing paths in accordance with the original P2P communication handshake protocol. This approach can quickly build a more reliable routing path. After all, P2P networks in most of the time, most nodes are "good", which is one of the main reasons that why P2P application is rapidly development. According to Amdahl's law, such big probability of events should be given priority to its implementation and optimization. After the establishment of the routing path, reuse the periodic maintenance operations of P2P network topology to implement the relay node attributes test, the price is low but a good effect, and does not create significant negative impact to P2P communicate.

When the routing path has been chosen, the communication sponsors nodes firstly need to record each relay node IP address, port number, network identification ID. And then they need to consult with each relay node about a set of asymmetric keys, and save public key for future exploration purposes. For malicious nodes, including the virus-controlled zombie nodes, if they demonstrate the rejection and even destruction to each normal P2P network events, they would be premature to reveal themselves and being kicked out of P2P networks. So these nodes usually perform malicious activities based on certain probability. If the above record or the consultative process can not be completed, then it needs to choose another sub-optimal routing path. Based on the above discussion, the pseudo-code for the initial establishment of the routing path is shown below.

TABLE I. PSEUDO-CODE FOR FIRST ESTABLISHMENT OF ROUTING PATHS

$P_s \xrightarrow{Build\ \ Connection} P_d$
Obtain multiple routing paths between P_s and P_d
Put those routing path into array $R\{r_1, r_2, ..., r_n\}$
If routing length of $r_i > 6$ then
Kick r_i out of R // i =1,2,3….
End if
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

B. Periodical detection of P2P routing path

When a routing path established, before the end of the communication process, the communication initiator send particular encrypted information to recipient regularly using periodical maintain order Ping command of the P2P network topology. This information is transmitted to the destination

node in turn in the routing path with the Ping command. Every time passing relay node, part of the information will be updated and signed, to record the behavior of the node. If there are some problems with a relay node, the related issues information will be returned to communications initiator with Pong command. Communication initiator analysis the returned information to confirm the nature of the problem and the location of occurrence, and decide whether to put bad node into the node list according to the analysis results, whether to reset the routing path immediately or to test again. If it needs to reset the routing path, in the establishment of a new routing path, it should take the initiative to avoid the nodes in the bad node list, to obtain safe and reliable a new route path.

In order to prevent malicious nodes tampering and forwarding in random Ping and Pong command with information, we encrypted our message using asymmetric nested encryption with public key, so as to stop this malicious behavior. Specific methods are as follows.

Assume that a routing path is $R = <P_0, P_1, P_2, P_3, ..., P_m>$, P_0 as the communication sponsor, controls the important information of all relay nodes in this routing path, for example, all of the < *IP addresses, ID Identity, Public key* > triples group, as well as the sequence of relay nodes and the path length in the routing path. At the beginning of detection, P_0 randomly generated a positive integer X. It is nested with all of the < IP addresses, public key> pairs. After it is encrypted, it will be passed to the next hop node P_1 within the Ping command. The encrypted format of detection packets is as follows.

$$K_1(X, A_2, K_2, K_2(A_3, K_3, K_3(A_4, K_4, K_4(....), .., A_{m-1}, K_{m-1}, K_{m-1}(A_m, K_m))...))$$

In order to protect the authenticity of random number X, P_0 also need to sign X using its own private key S_0, and pass the results $S_0(X)$ to the P_1. In addition, because that the random number X and all the < *IP addresses, public key* > pairs are also encrypted using the public key of the node P_1, It is only P_1 have the correct private key to decrypt and use this information. Thus it has prevented the interception and tampering of other malicious nodes, so as to guarantee the reliability of detection. After receipting the above information, the node P_1 unlocks the data packets with its own private key S_1. But it can only remove and use the X, A_2, K_2 three data. Because the latter part of the information is encrypted using the public key P_2 node K_2, P_1 can not decrypt them. This ensures that if P_1 is precisely a malicious node, it still can not tamper with the detection information of the follow-up node, the same as the other relay nodes, which will further enhance the security and reliability of RAPD algorithm.

According to the address A_2 of the node P_2, the node P_1 encrypted (X +1) with K_2 and transmits it to node P_2, at the same time it must sign (X +1) with its own private key S_1, and pass the generated data S_1 (X +1) to the P2 along. This is to prevent malicious relay node to write without basis (X + i) value and to cause route damage. To preserve this signature is an important role to detect the location of a malicious node. Finally, the node P_1 also needs to pass the decrypted message K_2 (A_3, K_3, K_3 (A_4, K_4, K_4 (…), …, A_{m-1}, K_{m-1}, K_{m-1}(A_m, K_m))…) to the P_2. That is the latter part of the probe packets. P_2 decrypts this information with their own private key S_2, then passes the latter part of the packet to the node P_3, at the same time encrypted (X +2) with K_3 and also passes it to P_3, finally it needs to sign the information (X +2, S_1 (X +1)) with its own private key S_2, and passes the result S_2(X+2, S_1(X+1)) to P_3. Repeat to carry out such operations, until to the destination node P_m. If the routing path unimpeded, then when the detection information arrivals P_m, X will become $X + m$. It can be seen that in fact the positive integer m represents the number of the route hops that the probe packets has passed. In addition, the signature information also becomes into the following format, which is also a nested encrypted data.

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

Finally, the node P_m attached the above-mentioned signatures packet to Pong command, and returned it to the communication sponsor P_0 along the original routing path. P_0 decrypted the signature information using the corresponding public key K_m, and the data ($X + m$) can be removed. If the m value is consistent with the routing path length that is their own pre-saved, then considered that there are no bad nodes in the routing path. When the next maintenance cycle of a P2P network topology arrived, it proceeds to the next round of exploration process

After P_i encrypts (X + i) with its own private key S_i, attaches it to the Pong command along with the notice information that a routing path can not reach together, and returns to the communication initiator P_0. P_0 decrypt i with the corresponding public key K_i, and classifies the node P_i +1 into the bad node list based on the value of i, and instructs the node P_i to prefer other nodes in their own routing table, and continues to search for a path to reach node P_m. This is a routing path resetting process from the routing path left off, and it is also a less costly replacement.

If the communication initiator P_0 receive an error detection feedback information (usually m value is not correct), the main reason is that there are malicious nodes in the routing path tampering the data, for example, when generating data (X + i), the malicious node deliberately writes its value wrong, and then passed to the next hop node, which caused the consecutive calculation errors for i by the follow-up nodes. Because i represent the routing hops and the order of relay nodes in the routing, at this time, P_0 simply decrypts the i out which is in the signature information S_i (X + i) sent by the relay node, queue according to the routing order, and that will find the location of a malicious node.

V. SIMULATION AND ANALYSIS

Simulation experiments carried out in the PC with CPU P4 2.0GHz, 1G memory, operating system Fedora Linux 8.0, simulation software P2Psim3.5. It is a modular simulation software specifically designed for P2P network, it is powerful, rich, easy to use and integrates Chord, CAN, Koorde and a dozen other mainstream P2P protocol, P2P research field is the preferred tool software for simulation.

For comparing the advanced effect of RAPD algorithm, we chose Chord and Koorde algorithm for comparing, the former is the most commonly used loop network structure, and the latter is famous P2P network protocol based on graph theory,

so they can reflect advanced effect. The three algorithms all simulate 10^4 nodes. in the 1st experiment, we has inspected the effect of RAPD algorithm to detect bad nodes in P2P network, the ratio of bad nodes is set at 30% and be well-distributed, and set that bad nodes do not respond to detective packets with a 30% probability and tamper detection package data with 40% probability. The second experiment investigate the changes of the required average number of attacks to successfully destroy a routing path with the condition of growing proportion of malicious nodes in P2P system.

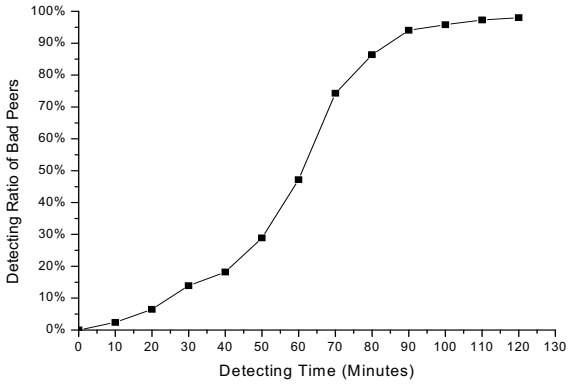

Figure 2. Detection rate of bad nodes in RAPD algorithm

As can be seen from Fig.2, RAPD algorithm for the detection rate of bad nodes is more ideal. After the P2P system has been running 20 minutes, the successful detection rate began to surge in, to 80 minutes it has detected more than 90% of the negative nodes out of the system, after 110 minutes it has detected all the bad nodes. It can be seen that the detection effect of RAPD algorithm is also very good, and substantially increased the routing security of P2P networks

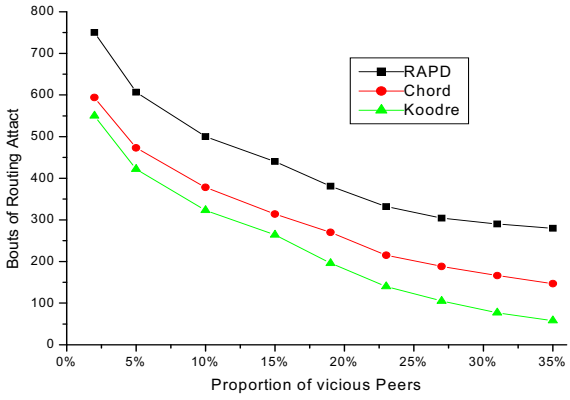

Figure 3. Relationship between the least average number of attacks and the proportion of malicious nodes

As can be seen from Fig.3, with the proportion of malicious nodes in P2P network increases, the required average number of attacks with the successful destruction of a routing path has a declining trend in the three kinds of algorithms, especially in the Koorde algorithm. Chord algorithm is followed. In RAPD algorithm, the rate of decline slowly, and after the proportion of malicious nodes reaches 23%, the rate of decline tends to smooth, which shows good anti-aggressive. On the contrary, with the reduction in the proportion of malicious nodes, the required average number of attacks in RAPD algorithm has a significant increase, much higher than the other two algorithms. It has increased the difficulty of a malicious node attack and increased the safety degree of P2P system.

VI. CONCLUSION

RAPD algorithm can effectively find the malicious nodes and the instable nodes in the routing path in a short time, and bypass these nodes through the method of reset the routing paths. Thereby it has enhanced the security performance of the system. In addition, the algorithm can limit the maximum length of routing path to reduce communication latency and to optimize communication performance in the context of maintain the system anti-attack capability. The experiments distinctly demonstrated the ratio of the malicious nodes and the unstable nodes in a P2P system have a greater impact on the effects of routing attacks. the RAPD algorithm mainly focuses on the discovery and exclusion malicious nodes and unstable nodes of system, so as to create a safe and efficient routing path.

REFERENCES

[1] Miguel Castro, Peter Druschel, Ayalvadi Ganesh. Secure routing for structured peer-to-peer overlay networks. In: Proceeding of Operating Systems Design and Implementation, OSDI'02. Boston, USA: IEEE Press. December, 2002. PP: 299-314

[2] Mudhakar Srivatsa, Ling Liu. Vulnerabilities and Security Threats in Structured Overlay Networks: A Quantitative Analysis. In: Proceeding of Computer Security Applications Conference. Tucson, Arizona: IEEE Press, December 2004. PP: 181-195

[3] I. Avramopoulos, H. Kobayashi. Highly Secure and Efficient Routing. In: Proceeding of INFOCOM. BK, USA: IEEE Press, July 2004. PP: 82-97

[4] Peng Wang, Nicholas Hopper. Myrmic: Secure and Robust DHT Routing [Technology Report]. University of Minnesota, 2006/09

[5] Shruti P. Mahambre, Umesh Bellur. Reliable Routing of Event Notifications over P2P Overlay Routing Substrate in Event Based Middleware. In: Proceeding of Parallel and Distributed Processing Symposium. Long Beach, CA: IEEE Press, 2007. PP: 1-8

[6] Sheila Becker. Security Evaluation for P2P Communication Systems: [Master Degree Thesis]: University of Luxemburg of Radu State. June 24, 2008.

[7] KPN Puttaswamy, H Zheng, BY Zhao. Securing Structured Overlays Against Identity Attacks. IEEE Transactions on Parallel and Distributed Systems. 2008, 36(9). PP: 68-80

[8] J Kannan. Implications of Peer-to-Peer Networks on Worm Attacks and Defenses. CS294-4 Project, 2003

[9] Naoum Naoumov, Keith Ross. Exploiting P2P Systems for DDoS Attacks. In Proceeding of Scalable information systems. Hong Kong, China: ACM Press, 2006. PP: 63-68