Researches on Secure Proximity Distance Defending Attack of Finger Table Based on Chord

Jie Li, Xiaofeng Qiu, Yang Ji, Chunhong Zhang
Department of Information and Communication engineering, Beijing Univ. of Posts and Telecommunications China
{zhishishizhi, qiuxiaofeng, jiyangbupt,zhangchbupt}@gmail.com

Abstract — **Structured Peer-to-peer networks overlay provide a substrate for the contribution of scalability, load balance, decentralization, and availability. However, new nodes are confronted with hijacking attacks on the Peer-to-Peer overlay with attackers. This paper proposes a model that the malicious nodes can not modify messages over the links and therefore can not control identifiers of correct nodes. However, malicious nodes could send fault messages to befool the correct nodes. Peer-to-Peer routing and location algorithm Chord is analyzed. Secure proximity distance that is security policy to defense the finger table attack is proposed. Analysis shows that the impact of attack is relative to the scale of overlay, the proportion of malicious nodes belonging to each attacker, the secure proximity distance, and the number of attackers. We conclude that the proportion of malicious nodes in the finger table will be higher than the proportion of malicious nodes in the overlay because of finger table attack.**

Keywords — **Networking security, peer-to-peer, finger table, secure proximity distance.**

1. Introduction

Structured Peer-to-Peer overlays like Chord[1], CAN[2], Tapestry[3] and Pastry[4] provide a self-organizing substrate for large-scale peer-to-peer applications. In structured peer-to-peer overlays, each node maintains links to a relatively small set of peers. All communication within the overlay, including searching resource, routing and location messages, occurs on these links. So the availability of structured peer-to-peer overlays depend on the integrity of links, however, structured peer-to-peer overlay networks can not insure the secure of links. A number of malicious nodes conspire to fool correct nodes into making links with malicious nodes, adopting the malicious nodes as their links with the goal of controlling most of the communication within the overlay.

By controlling the routing table or finger table of correct nodes, malicious nodes monitor most of links. Attackers that want to poison the finger table of correct nodes, have to gain a fraction of malicious nodes and deceive the correct by convincing means, otherwise attacking of fraudulence will be identified by corrected nodes.

In sybil attack, a single identity that presents many identities can control a substantial fraction of the system[8]. Malicious nodes can present many identities by sybil attack, by which attacker can present more identities with a small quantity of entities. Attacker can implement eclipse attack[5], since identity of routing table is not exclusive, and the routing table of each node only have a small number of networks identities. Hildrum and Kubiatowicz[12] describe a different defense against the eclipse attack based on proximity neighbor selection. Each node selects as its neighbors the nodes with minimal network delay, among all the nodes that satisfy the prefix matching for a given neighbor set member. Since a small number of malicious nodes can not be within a low network delay of all correct nodes, it is therefore difficult for them to mount eclipse attack.

Proximity distance[13] has been propose to achieve better routing efficiency. PChord is constructed on the basic of Chord which exploits proximity of underlay Internet by combining proximity routing into its routing scheme.The proximity of underlay Internet is proposed to improve reboust and efficiency of Peer-to-Peer overlay. On this paper, we propose the secure proximity distance is overlay layer, which is logical distance of nodes, so secure proximity distance is different from locality proximity distance.

This paper proposes a model that the malicious nodes can not modify messages over the links and therefore can not control identifiers of correct nodes.However, malicious nodes could send fault messgaes to befool the correct nodes. Analysis shows that the impact of finger table attack is relative to the scale of overlay, the proportion of malicious nodes belonging to each attacker, the secure proximity distance, and the number of attackers. The key idea of this paper is to propose the secure proximity distance.

The rest of this paper is organized as follows. Section 2 depicts the secure issue of Chord and the model of attack, and presents solution of secure issue, secure proximity distances purify the finger table. Section 3 deduces the conclusion and section 4 discuss the results, respectively.

2. Background, model and solution

In this section, we present some background in structured peer-to-peer overlay algorithms like chord, since this paper research the impact of finger table attack in Chord overlay. Nowadays, eclipse attack researchers omit the finger table issues of Chord, so most of analysis and technique in this paper were evaluated only in the context of Chord. Next, we describe attack models and assumptions used later in this paper including the number of attacker, secure proximity distance, and proportion of malicious nodes.

This work is supported by the National High-Tech Research and Development Plan of China under Grant No.2007AA01Z205 and 2008AA01A310.

2.1.Chord

Chord provides support for just one operation: given a key, it maps the key into a node. Data location can be easily implemented on top of chord by associating a key with each data item, and storing the key/date item pair at the node to which the key maps[1].Chord uses 160-bit circular id space, and forwards message only in clockwise direction in the circular id space. Chord nodes maintain a finger table consisting of up to 160 pointers to other live nodes. The i th entity in the finger table of node n refers to the live node with the smallest node identifier clockwise from $n+2^{i-1}$. Chord's replica function maps an object's key to the node identifier in the neighbor set of key's root, since replicas are stored in the neighbor set of the key's root for fault tolerance. Actually the expected number of nodes maintain the links is about $\log_2 N$, and the expected number of routing hops is $\frac{1}{2}\log_2 N$ [1], N denotes the number of identities on overlay, without especially statement, in this paper number of routing hops adopt above conclusion.

2.2.Attack model

We would introduce the attack model by analyzing the process of query, the query request don't distinguish lookuping predecessor message from fulfilling finger table message, since the essence of process of all messages of routing and location are same.

From Fig.1, a new node wants to find its predecessor and successor node by bootstrapped node. The node sends the request to bootstrapped node which is one of existed nodes in the overlay.

(1) New node sends the request of query to bootstrapped node to lookup its predecessor and successor node.

(2) Bootstrapped node inspects its finger table to confirm if it is predecessor or successor node of query node. If it isn't, query is routed to node C, since in bootstrapped node finger table, which is the nearest node to new node logically.

(3) Node C receives the request of query, and inspects finger table. Node C is not predecessor or successor of new node, then routes the request to node D.

(4) Node D is a malicious node, and malicious nodes that belong to attacker can conspire, then malicious node D inspect all the malicious nodes controlled by attacker, and find two malicious nodes which locate on left and right of query node, respectively. From malicious nodes list, node D detects that malicious node A' and B' close with query node logically.

(5) Node B' state that it is predecessor node of new node and node A' is the successor node of new node. Node A' state that it is predecessor node of new node and node B' is the successor node of new node. Query node would confirm that successor node of the predecessor node is successor node, and predecessor node of successor node is predecessor node, absolutely, node A' and B' suffice qualification.

(6) Query node is hijacked by attacker, and it is difficult to connect the authentic overlay. If the query searches a live node with the smallest node identifier clockwise from $n+2^{i-1}$, then the i th item of finger table points to malicious node, finger table attacker is implemented successfully.

Without malicious nodes overlay, node D should route the query to node E, and node E route the query to node B and A. Node A and B is successor and predecessor of query node respectively. However, query node can not identify the malicious message from node A' and B'. Without time after time repeat query and compare, correct node would never rectify the finger table which was poisoned.

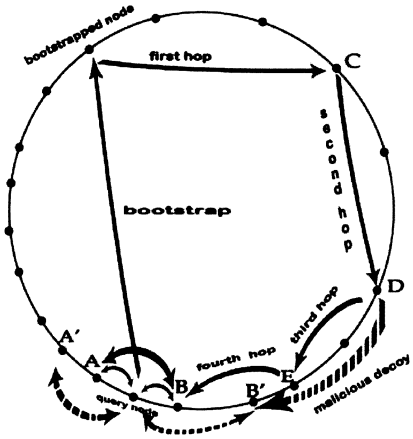

Fig.1 query the predecessor and successor node

Regardless the style of routing and location is recursive or iterative, attacker can implement the finger table attack by conspiracy of malicious nodes. On the other hand, attacker need malicious nodes which near the correct node logically, if the malicious node B' is far from new node, it will be difficult to convince the correct node that node B' is predecessor node of query node.

In this paper, although attacker can hijack new node by befooling, we mainly present the impact of attacking the finger table.

2.3.Solution and secure proximity distance

Before the description of solution, we introduce the assumption of this paper. Assumes that the number of attacker is m, and attackers can not conspire to fool correct node. However, each attacker controls many identities, which can conspire to fool correct node. Entity would present many identities as far as it can by sybil[8] attack, and assumes that the proportion of malicious node controlled by each attacker is equal. The proportion of malicious nodes on the overlay is p, and then the proportion of malicious nodes controlled by each attacker is $\frac{p}{m}$. Almost all DHT-based systems use hash function to deprived node identifier, and the identifier of nodes distribute on the chord circle even. We define the length of Chord circle is 1, and expected logical distance of two nodes adjacent in overlay where there are N nodes is $\frac{1}{N}$. We

assume that secure proximity parameter is s, and the secure proximity distance is $\frac{s}{N}$.

If $\left|\frac{ID_i-\left(n+2^{i-1}\right)}{2^{160}}\right| \le \frac{s}{N}$, the i th item of finger table identity ID_i is valid, if $\left|\frac{ID_i-\left(n+2^{i-1}\right)}{2^{160}}\right| > \frac{s}{N}$, the i th item of finger table identity ID_i would be vacant. Since a small number of malicious nodes don't close with all correct nodes, it is therefore difficult for them to mount finger table attack, so secure proximity distance is benchmark to evaluate the secure level. The smaller secure proximity distance is, the securer the overlay is, since the probability of malicious node sufficing the qualification is low.

We have to consider issues of rejection and vacancy when new node join overlay. If predecessor or successor node doesn't suffice qualification, node will be rejected because of lack of predecessor or successor node. When rejection happens, the solution of rejection issue is enlarging the secure proximity distance. When the proportion of vacancy in finger table is low, the impact of vacancy to efficiency of routing and location is negligible. Since finger table have 160 items and $\log_2 N$ identities, which insure the message routing. In order to not worsen the efficiency of routing, we give the expression of vacancy proportion in finger table.

Theorem: If there are millions of nodes in the overlay networks, the probability of vacancy in finger table is about $p_{fn} \approx e^{-s}$.

If logically distance between $n+2^{i-1}$ and smallest node identifier clockwise from $n+2^{i-1}$ is more than $\frac{s}{N}$, the i th item of finger table would be vacant, which means that other nodes locate on the $\left(1-\frac{s}{N}\right)$ area of circle.

$$p_{fn}=\left(1-\frac{s}{N}\right)^{N-1} \approx e^{-s} \quad (1)$$

If $s \ge 3$, then $p_{fn} \le 5\%$, and the probability of vacancy in finger table is low.

From Fig.1, we conclude that, if $\forall ID_m \in A, ID_n \in A$, and $\left(ID_m-K_q\right) \le \frac{n}{N} \cap \left(ID_n-K_q\right) \le \frac{n}{N}$,finger table attacking occur. Denotation A represents an attacker, identity ID_m and ID_n both are malicious nodes which belong to attacker A. Query request smallest node identifier clockwise from K_q, $\frac{s}{N}$ is secure proximity distance.

The proportion of malicious nodes controlled by each attacker is $p_m=\frac{p}{m}$. The probability of at least one malicious that suffices secure proximity distance for each attacker is p_{csc}. In equation 2, right formula subtracts p_m, because impact of finger table attack should not include the probability of malicious nodes in finger table normally, and we just evaluate the augment of proportion of malicious nodes in finger table after finger table attack.

$$p_{csc}=1-\left(1-\frac{s}{N}\right)^{N*p_m}-p_m \quad (2)$$

Finger table attack need two malicious nodes, both of which suffice secure proximity distance. One node is left adjacent, and the other node is right adjacent. So the probability of successful deceiving by each attacker is

$$p_{at}={p_{csc}}^2=\left(1-\left(1-\frac{s}{N}\right)^{N*p_m}-p_m\right)^2 \quad (3)$$

3. Theory Analysis

This section take into account scale of overlay, number of attacker, the proportion of malicious nodes in the overlay, and secure proximity distance. We conclude the impact of these parameters to finger table.

When query is routed to a malicious node, because malicious node has many malicious conspirators belonging to same attacker, if attacker befools the request successful, response would be sent to query node, query of node receives malicious response. If not, query would be transmitted to next hop normally. Assumes that he expected number of hop is l in the overlay, the probability of that query from correct node would be befooled is $p_g(l)$, and the probability of that request from malicious node would be befooled is $p_b(l)$.The probability of that next hop is correct node is $1-p$, and the probability of that next hop is malicious node is p.

$$p_g(l)=(1-p)*p_g(l-1)+p*p_b(l-1) \quad (4)$$

$$p_b(l)=p_{at}+(1-p_{at})*p_g(l) \quad (5)$$

Equation 6 is deduced from equation 4 and 5.

$$P_g(l)=1-(1-p_{at}P)^l \quad (6)$$

When the number of attacker is more than 4, proportion of malicious nodes in the overlay is less than 10%, the number of hop is not more than 10, and secure proximity parameter is not more than 5, the inequation $p_{at}*P*l=1$ is tenable. In order to make it easier to interpret this equation, we approximate the expression in equation 7 and equation 8. We have to declare that when $p_{at}*P*l=1$, the approximate is reasonable.

$$p_g(l) = p_{at} * p * l = \left(1 - \left(1 - \frac{s}{N}\right)^{N*\frac{p}{m}} - \frac{p}{m}\right)^2 * l * p \quad (7)$$

$$p_g(l) \approx \left((s-1)\frac{p}{m}\right)^2 * l * p = \frac{(s-1)^2 * l * p^3}{m^2} \quad (8)$$

The expected of number of hop is $\frac{1}{2}\log_2 N$ [1], then

$$p_g = \frac{(s-1)^2 * p^3}{2 * m^2} * \log_2 N \quad (9)$$

p_g denotes impact of finger table attack in finger table.

s denotes secure proximity parameter.

p denotes proportion of malicious node in finger table.

m denotes the number of attackers in the overlay.

N denotes the number of identities in the overlay.

4. Results

Without finger table attack, proportion of malicious nodes in finger table is equal to proportion in the overlay. From equation 9, we conclude that proportion of the malicious node in finger table is more than the proportion of malicious nodes in the overlay, which means finger table attack will enlarge p in the equation 4. When the proportion of malicious node in finger table $p' = (p + P_g) > p$, as the continuance of attack and churn of overlay, the result of attack diffusion is $P_g' > P_g$. If it is sure that messages will be routed to malicious nodes, the probability of malicious node in finger table reaches extreme.

Impact of attacks and secure proximity parameter are directly proportional to the square of relations as Figure 2. Figure 2 show the impact of finger table attack with $m = 1,5$, $p = 0.10$, $N = 32768$ and varying s. when the secure proximity distance varies from 1 to 10, the probability of finger table item befooled. When secure proximity parameter is 1, the impact of attacking finger table is about 0, however, the fraction of vacancy of finger table is $e^{-1} = 0.386$. Without secure proximity parameter, namely the secure proximity parameter is $s = N$? 10, the query would be befooled once route to malicious node.

From Figure 2, when $s = 2,3$, the impact of attack is less that 0.02, even the proportion of malicious nodes controlled by attacker is $p_m = 0.10$, so when the proportion of malicious nodes in overlay is high, rigorous secure proximity parameter is necessary.

Impact of finger table attack and the number of attackers of the square showed inverse relationship as Figure 3.

Figure 3 shows the impact of finger table attack with $s = 5$, $p = 0.10$, $N = 32768$ and varying m .When the number of attacker varies from 1 to 10, the probability of finger table item befooled. From Fig.2 and Fig.3, we conclude that when the number of attackers $m > 4$, even the secure proximity parameter $s \geq 10$, and the impact is less than 0.02, which means restriction the conspiracy of malicious node can decrease the impact of finger table attack. When proportion of malicious node controlled by attackers is $p_m \leq 0.02$, we can choose loose secure proximity parameter to insure integrity of finger table, since the impact of finger table attack is not significant.

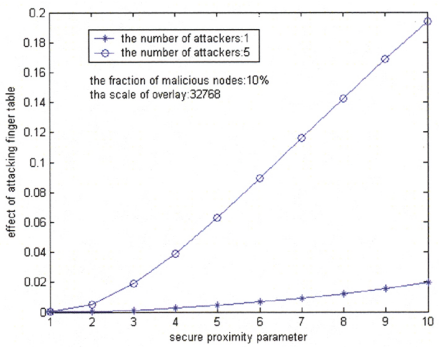

Fig. 2 secure proximity parameter impact

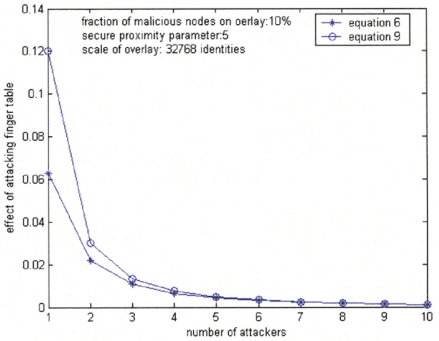

Fig. 3 number of attacker impact

5. Conclusion

This paper presents model of attack finger table for Chord under the environment of identifiers even on the Chord circle, and proposes a model that the malicious nodes can not modify messages over the links and therefore can not control identifiers of correct nodes with the assumption that the malicious nodes could send fault messages to befool the correct nodes. Research parameter includes the number of attacker, the scale of overlay, fraction of malicious nodes, and secure proximity parameter. The model is more varacious to evaluate the unsecure overlay networks environment. When the finger table attack is severe in overlay, secure proximity distance is significant to decrease the impact of finger table attack. On the other hand, secure proximity parameter produces the issue of vacancy in finger table, and the impact of vacancy is not severe.

REFERENCES

[1] I.Staica, R.Morris, D.Karger, F.Kaashoek, and H.Balakrishnan. Chord: A Peer-to-peer Lookup Service for Internet Applications.ACM SIGCOMM, 2001:149-160

[2] Sylvia Ratnasamy, Paul Francis, Mark Handley, Richard Karp, and Scott Shenker. A Scalable content-Addressable network. In Proc. ACM SIGCOMM'01, San Diego, California, August, 2001

[3] Ben Y.Zhao, John D.Kubiatowicz, and Anthony D.Joseph. Tapestry: An infrastructure for fault-relisilent wide-area location and routing. Technical Report UCB//CSD-01-1141, U.C.Berkeler, April 2001.

[4] Antony Rowstron and Peter Druschel. Pastry: Scalabel, distributed object location and routing for lager-scale peer-to-peer systems. In Proc. IFIP/ACM Middleware 2001, Heidelberg, Germany, November 2001

[5] Atul Singh, Tsuen-Wan "Johnny" Ngan, Peter Druschel, and Dan S.Wallach. Eclipse Attacks on Overlay Networks: Threats and Defenses. INFOCOM 2006

[6] E.Sit and R.Morris. Security considerations for peer-to-peer distributed hash tables. In Proceedings of 1st International Workshop on Peer-to-peer Systems (IPTPS), Cambridge, Massachusetts, Mar.2002.

[7] D S. Wallach. A Survey of Peer-to-Peer Security Issues. In International Symposium on Software Security, Tokyo, Japan, November 2002.

[8] J. Douceur. The Sybil Attack. In 1st International Workshop on Peer-to-peer Systems (IPTPS'02). Springer, 2002.

[9] efending the Sybil Attack in P2P Networks: Taxonomy, Challenges, and a Proposal for Self-Registration. IEEE, Proceedings of the 1st International Conference on Availability, Reliability and Security, 2002.

[10] Mudhakar Srivatsa and Ling Liu. Vulnerabilities and Security Threats in Structured Peer-to-Peer systems: A Quantitative Analysis. Proceedings of the 20th Annual Computer Security Applications Conference, 2004.

[11] Jian Liang, Naoum Naoumov, Keith W.Ross. The Index Poisoning Attack in P2P File Sharing Systems. Proceedings of IEEE Infocom 2006, Barcelona, Spain, April 2006.

[12] K.Hidrum and J.kubiatowicz. A asymptotically efficient approaches to fault-tolerance in peer-to-peer networks. In proceedings of 17th International Symposium on distributed Computing, Sorrento, Italy, Oct.23.

[13] Feng Hong, Minglu Li, Jiadi Yu; and Yi Wang. PChord: improvement on Chord to achieve better routing efficiency by exploiting proximity. 25th IEEE International Conference on Distributed Computing System Workshop