International Journal of Computer Applications Technology and Research
Volume 3– Issue 11, 679 - 681, 2014, ISSN: 2319–8656

Enhanced Detection System for Trust Aware P2P Communication Networks

K. Kalaivani
Computer Science and Engineering
V.S.B Engineering college
Karur, India

C.Suguna
Computer Science and Engineering
V.S.B Engineering College
Karur, India

Abstract: Botnet is a number of computers that have been set up to forward transmissions to other computers unknowingly to the user of the system and it is most significant to detect the botnets. However, peer-to-peer (P2P) structured botnets are very difficult to detect because, it doesn't have any centralized server. In this paper, we deliver an infrastructure of P2P that will improve the trust of the peers and its data. In order to identify the botnets we provide a technique called data provenance integrity. It will ensure the correct origin or source of information and prevents opponents from using host resources. A reputation based trust model is used for selecting the trusted peer. In this model, each peer has a reputation value which is calculated based on its past activity. Here a hash table is used for efficient file searching and data stored in it is based on the reputation value.

Keywords: provenance, p2p system, trust, reputation, hash table

1. INTRODUCTION

A botnet is a collection of compromised hosts (bots) that are remotely controlled by the botmaster. In the centralized architecture the botmaster can send commands to the bots through the command and control (C&C) channel. The disadvantage of this is, the bots can be easily identified and removed. In order to overcome this problem botmasters have implement botnets in peer to peer (p2p) networks. Botnets are dynamic in nature. It can range from larger network to smaller network [5].

Peer to Peer (P2P) network can be configured by itself and it is decentralized. Comparing to the traditional client server network, the p2p network can dispense more malicious or forged content, malicious code, worms, viruses and Trojans because of its decentralized nature. It is shown that the system where peers work only for their selfish interests. Policing these types of networks is extremely difficult because of the decentralized and the ad hoc nature. In the centralized approach, the disadvantage is that the central authority can be turned into malicious. If there is no central authority, repository, or global information means, then there is no mechanism for protecting the P2P networks. Structured P2P systems and unstructured P2P systems are the two types of decentralized network.

The structured peer-to-peer network is structured on a specific topology, and the protocol ensures that any node can efficiently search the network for a file or resource, even if the resource is extremely rare. The unstructured p2p network is formed by establishing arbitrary connection between the peers. The multicast overlays such as peer-to-peer overlays (e.g. Gnutella) are also incorporated in the overlay networks. When the overlay links are acknowledged randomly an unstructured P2P network is formed. The p2p network can be easily build, if a new peer that wants to join the network can replica the existing links of another node and then forms its own links over time.

The decentralized P2P networks are independent of a single server and they have to invest in server farms to guarantee the scalability of their systems [6]. There are some obvious advantages to decentralizing the C7C mechanism "self-healing" nature of the network, along with "servers" that actually share files. Each peer is frequently advertising its presence, as well as requesting updates from other peers [3]. There are many approaches that have been for detecting the botnets such as botminer and botgrep. The botminer can detect the botnets that have similar malicious activities [4]. It can differentiate between the authenticated and malicious user. But botminer has some limitations [1]. It can't able to identify the user with mixed characteristics. The botgrep can identify the botnets based on the network traffic. The disadvantage of this method is it is hard to maintain the traffic information. So, in order to detect the botnets we are using a security property called data provenance integrity. It is used to verify the origin of data.

The p2p systems have the advantage of having a hash table. It can provide the efficient and quick way of retrieving the data. It can identify the rare files more easily [2]. In this paper we use this hash table for file searching. In addition to this, for verifying the identity of the peers a self certificate based approach is used. These certificates can be exchanged between the peers during communication. These certificates are generated by the peer itself. The Self Certification mechanism helps to identify the malicious content and makes the search more effective. In self certification each peer is having their own certificate authority. The certificate authority issues the identity certificate. The self certification is used for assuring secure and appropriate availability of the reputation information of a peer.

The main requirements for these are:

1. A self certificate based identity approach
2. A trivial and straightforward reputation model.

3. An attack resistant cryptographic protocol for creation of reliable global reputation information of a peer.

2. DATA PROVENANCE INTEGRITY

In P2P file-sharing systems, all peers are both sender and receiver of resources and can contact each other directly without any middle agents. Thus the security becomes a problem in unstructured p2p network. For this, data provenance integrity is used as the security property. It is used to verify the origin of the data. It is a security control and it is mainly related to the data integrity. By using this property we can identify the unwanted changes in data. It can enhance the trustworthiness of data. Data provenance integrity measures the level of trustworthiness of both data and data providers by assigning values to them. Based on these values, peers can make their more informed decisions whether to use the data or not. The way in which the data was collected is also an important aspect when determining the trustworthiness of the data. For example, if a number of self-governing sources provide the same data, such data is most expected to be true. Reputation model is serving as a data provenance property.

In p2p networks peers cooperate to perform a function. Among the independent peers some may be authentic and provide high quality service and some may be malicious which provide harmful services such as the nodes may violate the rules and behave maliciously so as to obtain some personal gain and it is possible that a third party may inject some misbehaving nodes so as to disrupt the network service to gain a competitive edge in a commercial market. In order to overcome these problems a reputation based model is used. For each peer in the network, the trust management system maintains a universal trust label. When the system is reputation-based, the label is a combination of the local opinions of all peers in the network which is based on their previous experiences with other peers. To compute it, each opinion of the peer is weighted by the reputation of the opiner, thus peers with good reputations are more influential than those with poor reputations or no reputation.

An objective of the trust management system is to permit only trusted peers to get high reputations with high prospect. This allows trusted agents to easily recognize the malicious agents and potentially cut them from transactions. For example, a file served from a un-trusted peer might be given a low integrity value by the receiving host. Likewise, the routing protocol might keep away from forwarding messages through dishonest peers. These overall reputations permit each peer to get advantage from the experiences of all other peers in the network.

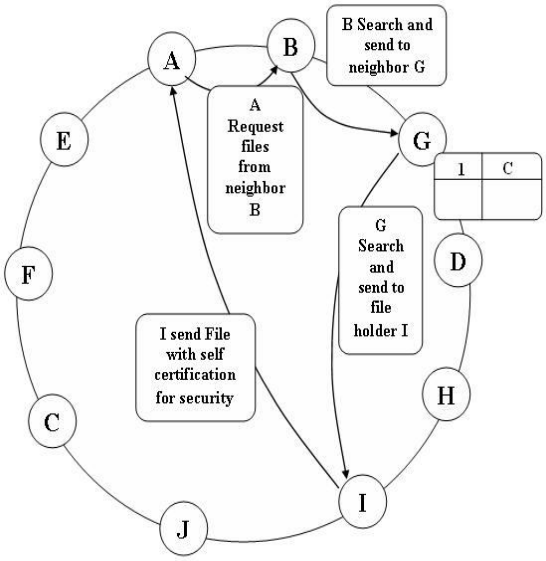

Fig 1: An overview of peer to peer unstructured network using self certification

The conversation of peer to peer networks is shown in fig 1. The node A can send the file searching request to node B. Node B can search the file in its hash table but the file is not present so it can forward the request to the node G. G can found the file in its hash table and it can forward the request to the node I because I contains that appropriate file. The node I has the file and it maintains self certification with node A to transfer the file securely.

Based on the reputation value the file searching is done in p2p network. The peer with highest reputation value is considered to be authenticated peer and the peer with lowest reputation value is considered as malicious peer. In order to receive a file from a peer, the reputation value is checked. If the value is high means then the file can be received or else it will be rejected. Thus we can prevent the peer from getting malicious data by using this reputation value. The data in the hash table is also stored based on these values.

In the decentralized P2P network, if a peer wants to find out a required file in the network then the request is passed through the network. This results in finding many peers that share their data. The major disadvantage of passing the request in the network is that the request may not always be determined. If the peer is searching for some popular data or file then the probability of finding the data is high. In other case where a peer is searching for rare data which is shared by only some peers, then it is highly doubtful that the search will be successful. As the peer and the content management store the data separately, there is no guarantee that the request will find a peer that has the required data. This passing of request causes a high amount of traffic in the network. These networks typically have poor searching efficiency. Popular P2P networks are generally unstructured. Thus a hash table is used for file searching.

A. Neighbor peer extraction

Self-governing nodes are available in many network services and some rules are required for these nodes to work together and to attain a given network functionality. To understand such network service, the neighbor node must communicate with other subset of nodes in the network e.g., send packets to neighbors, and receive packets from neighbors and so on. To assure the correct functionality of the network service such that every node can get service with desired performance, nodes must follow the predefined protocols when they participate in the communication with their neighbors.

Peer's construct its neighbors by sending connection request with its own certification likewise all peers shares its identity and its self certificate with its neighbor peer's. This certificate is compare by neighbor peer when the peer send file search request to the one of neighbor peer thus avoids unwanted flow of packets which reduce the traffic. Source peer select the neighbor peer as per highest reputation metric/value. The reputation value is calculated under each peer's total entry in the hash table or its load status.

B. Self Certification based approach

All the peers share its identity and its self certificate with its neighbor peers. The self certification is attached with identity of the peer. This certificate is compared by the neighbor peer when the peer sends file searching request to one of its neighbor peer. If the certificate is matched then the peer searches its hash table with the file name. It uses the concept of RSA and DSS.

C. Hash table based searching

In the unstructured P2P networks, peers distributing the file they have and forwarding the file searching request will plays an important role. The hash table is maintained in all the peers. It consists of its own file and the files which are received from other peers. If a file searching request is arriving means then the peer can search it in hash table. The file is present means then the request is forwarded to the peer which is the owner of that file. The anticipated system uses the hash table where each and every peer has the detached hash table.

Based on the reputation value, which is calculated based on the past activities of peer, the files are stored in the hash table. This stored information helps to achieve the file searching operation skillfully. On receiving the file searching request the peer first checks the reputation value. If the value is high then the peer can accept the request. A self certificate is exchanged between the peers to ensure the secure access of data. The hash table is used for forward the file searching request to the appropriate peer instead of the neighbor peer.

D. Trustworthy peer communication

After the successful match of self certification, each neighbor peer must recognize the source file request . Thus, the malicious peers which attempt to compromise with other peer address can be avoided. Then the source a request from the starting neighbor peer is advanced to the next peer with the hidden identity of source peer thus reduces the possibilities of unwanted rebroadcast packets and also malicious behaviors. So the initiate neighbor peer acts as temporary source, then forward the source request to its neighbor peer and this will continues until the file is found. The destination peer sends the file to the initiate neighbor and the initiate neighbor peer sends the file to the source thus avoid the length of the packet flow and control the misbehave peers.

3. CONCLUSION

This project presents the reputation model and a cryptographic protocol that provide generation of global reputation data in a P2P network, in order to detect rogues. This can improve the trustworthiness of peer and its data. It can also prevent the peers from getting malicious data from a malicious peer. The main technological contributions that are proposed in this project are the model and operations of cryptographic provenance verification in a host-based security setting and shows the provenance verification approach in a lightweight framework for ensuring the integrity of outbound packets of a host.

4. REFERENCES

[1] Junjie Zhang, Roberto Perdisci, Wenke Lee, Xiapu Luo, and Unum Sarfraz "Building a Scalable System for Stealthy P2P-Botnet Detection" IEEE transactions on information forensics and security, vol. 9, no. 1, january 2014.

[2] S. Rhea, D. Geels, T. Roscoe, and J. Kubiatowicz, "Handling churn in a DHT," in *Proc. Annu. Conf. USENIX Annu. Tech. Conf.*, 2004, pp. 127–140.

[3] S. Stover, D. Dittrich, J. Hernandez, and S. Dietrich, "Analysis of the storm and nugache trojans: P2P is here," in *Proc. USENIX*, vol. 32. 2007, pp. 18–27.

[4] G. Gu, R. Perdisci, J. Zhang, and W. Lee, "Botminer: Clustering analysis of network traffic for protocol- and structure-independent botnet detection," in *Proc. USENIX Security*, 2008, pp. 139–154. Forman, G. 2003.

[5] D. Dagon, G. Gu, C. Lee, and W. Lee, "A taxonomy of botnet structures," in Proc. 33rd Annu. Comput. Security Appl. Conf., 2007, pp. 325-33.

[6] A. Binzenhofer, D. Staehle, and R. Henjes, "On the stability of chordbased P2P systems," i*n Proc. IEEE Global Telecommun. Conf.*, vol. 2. Nov./Dec. 2005, pp. 884–888.