ISSN (Online) : 2319 – 8753
ISSN (Print) : 2347 - 6710

International Journal of Innovative Research in Science, Engineering and Technology

An ISO 3297: 2007 Certified Organization, Volume 3, Special Issue 1, February 2014

International Conference on Engineering Technology and Science-(ICETS'14)
On 10th & 11th February Organized by

Department of CIVIL, CSE, ECE, EEE, MECHNICAL Engg. and S&H of Muthayammal College of Engineering, Rasipuram, Tamilnadu, India

Ensuring Host Security of Data Provenance For Trust Aware P2P Networks

R. Jayanthi[1]
II - M.E Computer & Communications, Department of ECE, EBET Group of Institutions, Kangayam, Tamilnadu, India [1]

Abstract — Malicious software typically resides furtively on a user's computer and interacts with the user's computing resources. The goal is to improve the trustworthiness of a host and its data by classifying the data to secure the file. Specifically, the new mechanism is provided that ensures the provenance of critical system information and prevents bots from utilizing host resources in peer to peer systems. Data integrity defines the security property which states that the source where a piece of data is generated cannot be spoofed or tampered. This project describes a cryptographic provenance verification approach and applying classification mechanism to find out the ratio of good and bad word count before rejection of original file. This ensures system properties and the system-data integrity, and then the application is demonstrated in the keystroke integrity verification. Specifically, it first designs and implements an efficient cryptographic protocol that provides data integrity. The protocol prevents the imitation of fake key events by malware under the reasonable assumptions. Then demonstrate the provenance verification approach by realizing the lightweight framework. If the verification fails each of the content is compared with trained data set to conclude the content must be malicious to the peer so on the attack basis is just warns the peer, otherwise the file will be received successfully and safe.

Keywords — Malware, cryptography, provenance, P2P System, classifier.

I. INTRODUCTION

A. Peer to Peer network

Peer to Peer (P2P) networks are self-configuring networks with minimum or no centralized control. P2P networks are more vulnerable to dissemination of malicious or spurious content, malicious code, worms, viruses and trojans than traditional client-server networks, due to their unregulated and unmanageable nature. For example, the notorious VBS. Gnutella worm that infected the Gnutella network, stored Trojans in the host machine. The peers in the P2P networks have to discourage from leeching on the network. It is shown that the system where peers work only for their selfish interests while breaking the rules decays to death. Policing these type of networks is extremely difficult because of the decentralized and the ad hoc nature. Besides P2P networks like the Internet are physically spread across the geographic boundaries and hence they subject to variable laws. The traditional mechanism for generating the trust and protecting client-server networks cannot be used for the pure P2P networks. This is due to the trusted central authority in traditional client-server networks is absent in the P2P networks. Introduction of central trusted authority like the Certificate Authority (CA) can reduce the difficulty of securing the P2P networks. The main disadvantage of the centralized approach says that if the central authority turns to malicious then the network will become vulnerable. If the central authority, repository, or global information is absent, then there is no silver bullet for securing P2P networks. Decentralized P2P systems are typically classified into two categories: structured P2P systems and unstructured P2P systems.

B. Architecture of P2P Systems

Peer-to-peer systems often implement an abstract overlay network and built at the Application Layer on top of native or physical network topology. Such overlay network are used for indexing and peer discovery and make the P2P system independent from the physical network topology. The content is typically exchanged directly over the underlying Internet Protocol (IP) network. Anonymous P2P systems are an exception and implement extra routing layers to obscure the identity of the source or destination of queries.

1) Structured Systems

Structured P2P networks employ a globally consistent protocol to ensure that any node can efficiently route a search to some peer that contains the desired file, even if file is extremely rare. Such guarantee necessitates a more structured pattern of the overlay links. By far the most of the common type of structured P2P network is the distributed hash table (DHT), in which a variant of consistent hashing is used to assign ownership of each file to a particular peer which is analogous to a traditional hash table's assignment of each key to a particular array slot.

2) Unstructured Systems

An unstructured P2P network is formed when the overlay links are established arbitrarily. Such networks can be easily constructed as a new peer that wants to join the network can copy existing links of another node and then form its own links over time. In an unstructured P2P network, if any peer wants to find a desired piece of data in the network then the query has to be flooded through the network to find as many peers as possible that share the data. The major disadvantage with such networks is that the queries may not always be resolved. The popular content is likely to be available at several peers and any peer searching for it is likely to find the same thing but if a peer is looking for rare data shared by only a few other peers then it is highly unlikely that search will be successful. Because of no correlation between a peer and the content managed by it and there is no guarantee that flooding will find a peer that has the desired data. The flooding also causes a high amount of signaling traffic in the network and hence such networks typically have very poor searching efficiency. Many of the popular P2P networks are unstructured.

The main requirements are:

1. A self-certification-based identity system protected by cryptographically blind identity mechanisms.
2. A light weight and simple reputation model.
3. An attack resistant cryptographic protocol for generation of authentic global reputation information of a peer.

II. DATA PROVENANCE INTEGRITY

The goal improves the trustworthiness of the network level data flow by ensuring the correct origin of affected system data, which prevents adversaries from utilizing the host resources (e.g., owned resource files). Data-provenance integrity states that the source from which a piece of data is generated can be verified. This provides the description of how data-provenance integrity can be realized for system-level data in peer to peer systems. This project focuses on a security on host-based approach for ensuring system-level data integrity and demonstrates its application for the malware detection. In comparison the network trace analysis typically characterizes malware communication behaviors for detection. Such solutions usually involve pattern-recognition and machine learning techniques, and have demonstrated effectiveness against today's malware. The cryptographic verification method is defined as a robust mechanism that ensures the true origin of the data produced by an entity such as a system stored data. Although simple, the cryptographic provenance verification method can be used to ensure and enforce correct system and network properties and appropriate workflow under a trusted computing environment.

The major contributions of this project are:

1. A self-certification based identity system is protected by cryptographically blind identity mechanisms.
2. A light weight and simple reputation model.
3. An attack resistant cryptographic protocol for generation of authentic global reputation information of a peer.

A. Architecture

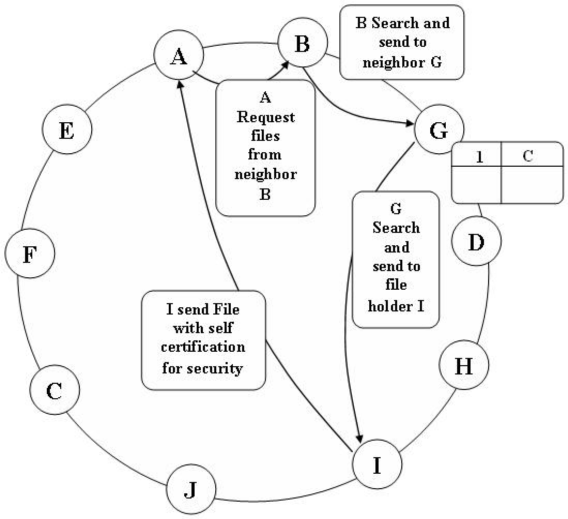
A
B
B Search and send to neighbor G
G
1
C
E
A Request files from neighbor B
D
F
G Search and send to file holder I
H
C
I send File with self certification for security
I
J

Fig.1 An Overview of Peer to peer unstructured network using self certification and reputation management.

The communication of peer to peer networks is shown in figure. Each peer request to the neighbour peer. The node A searches the file from neighbour node B and it checks whether it has the file by maintaining hash table, if not it forward the query to the its neighbour peer and it goes on continuously. If the file is found in any node it provides the response to the requested peer. The node I has the file and it maintain self certification with node A to transfer the file securely.

B. Unstructured P2P Network Formation

An unstructured P2P network is formed when the overlay links are established arbitrarily and such networks can be easily constructed as a new peer that wants to join the network can copy existing links of another node and then form its own links over time. In an unstructured P2P network, if any peer wants to find a desired piece of data in the network then the query has to be flooded through the network to find as many peers as possible that share the data. Each peer known only their neighbor peer that's represents unstructured P2P network.

C. Reputation or Hash Table Based Searching

In the unstructured P2P networks, peers willingness to share the content they have and forward the queries plays an important role during the content search process. Each and every peer in the network must maintain this table which is used to forward the peer request to the apt peer instead of its neighbor peer. The proposed system uses the distributed hash table where each and every peer has the separate hash table.

The information stored in the hash table is based on Reputation management (tracking peers past activity).It helps to perform the file searching operation efficiently. The self certificate is used for ensuring secure and timely availability of the reputation data of a peer to other peers. Since each peer stores its own reputation as local, for reputation to be reliable and elective, they have to be updated and stored securely to prevent malicious peers from the reputation system.

D. Asymmetric Cryptography Approach

Peers generate universally unique identifications locally and store them along with their public key, their current IP address. Implementation of Self Certification and Digital Signature for Secure Communication is focused in this module. RSA is used to generate public key and for data encryption and decryption. MD5 is used to generate the digital signature on the basis of the encrypted file.

E. Digital Signature Matching System

Each and every peer has the unique identity, based on this, the peer is identified and the transaction begins. The certification is attached with identity of the peer. The certification uses the concept of RSA and DS where the algorithm generates the private key and public key, these identities are attached with reputation of the given peer. The sender sends the information which is associated with its private key and signature, at receiver side receives the file and generates the signature, it will be matched with the attached signature if it matches, then concludes no more presence of malicious peer else proceeds to next module steps.

F. Data Auditing Under Training Sets

Each receiver maintains trained data sets which are used whether signature verification fails. Each content or word must be compared with trained data set and find the similarity. If the similarity is much more then it concludes the content must be malicious to the peer, so on the attack basis it just warns the peer, otherwise the file will be received successfully and safe.

III. CONCLUSION

This project presents the self-certification, an identity management mechanism, the reputation model and a cryptographic protocol that provide generation of global reputation data in a P2P network, in order to detect rogues. This describes a general approach for improving the assurance of system data and the properties of a host which has applications in preventing and identifying malware activities. The following technical contributions are made in this project: proposed the model and operations of cryptographic provenance verification in a host-based security setting and demonstrated the provenance verification approach in a lightweight framework for ensuring the integrity of outbound packets of a host.

REFERENCES

[1] Amit Chaudhary, Chander Diwaker and Sandeep Kumar, " Reputation System In Peer-To-Peer Network: Design And Classification", Journal of Global Research in Computer Science, Volume 2, No. 8, September 2011.

[2] Audun Josang, Roslan Ismail and Colin Boyd., "A Survey of Trust and Reputation Systems for Online Service Provision", 2007.

[3] Kui Xu, Huijun Xiong, Chehai Wu, Deian Stefan and Danfeng Yao, "Data-Provenance Verification For Secure Hosts", IEEE Transactions On Dependable And Secure Computing, Vol. 9, No. 2, April 2012.

[4] Patrick McDaniel, Radu Sion and Marianne Winslett, "Towards a Secure and Efficient System for End-To-End Provenance", 2010.

[5] R. Gummadi, H. Balakrishnan, P. Maniatis, and S. Ratnasamy, "Not-a-Bot: Improving Service Availability in the Face of Botnet Attacks," Proc. Sixth USENIX Symp. Networked Systems Design and Implementation (NDSI '09), 2009.

[6] Satsiou.A; Inf. & Telematics Inst., Thessaloniki, Greece; Tassiulas, L., "Reputation-Based Resource Allocation in P2P Systems of Rational Users", April 2010.

[7] V. Sharma, J.B. Grizzard, C. Nunnery, B.B. Kang, and D. Dagon, "Peer-to-Peer Botnets: Overview and Case Study," Proc. First USENIX Workshop Hot Topics in Understanding Botnets, Apr. 2007.