International Journal On Engineering Technology and Sciences – IJETS™
ISSN(P): 2349-3968, ISSN (O): 2349-3976
Volume IV, Issue VIII, August - 2017

Security Ensured Source Discovery on Internet Data Communication using Cloud Resources

Mrs. S. Logeshwari, MCA., MPhil., Research Scholar,
Mrs. J.S. Subhashini, MCA., M.Phil., Assistant Professor,
Department of Computer Science,
SSM College of Arts & Science, Komarapalayam, Tamilnadu, India

Abstract

The data communication over the web or any local network is carried out with the support of the Internet Protocol (IP). Static and dynamic IP addresses are logically assigned to the hosts that are connected to the network environment. The online services can be attacked from various sources. The source discovery operations are performed using the IP traceback schemes. The user request traces are managed under the log files in the Internet Service Providers (ISP). The traceback logs are analyzed to discover the source. The intruders can identify the topology of the Internet Service Providers during the IP traceback operations.

The IP traceback operations can be performed with the support of the cloud resources. The IP traces are maintained and processed under the cloud resources with authentication and security features. The access control mechanisms are employed in the cloud based traceback services. The security ensured traceback services are performed using the Framework for Authentication in Cloud-based IP Traceback (FACT). The temporal packet based authentication process is applied in the FACT scheme. The temporal access tokens are combined with the traffic flows and delivered to the end host. Complex incentive management model is used in the FACT model. The header size is the key factor in the IP header marking process.

The security ensured source discovery scheme is build with the cloud resources based on the enhanced FACT model. The optimal marking scheme is combined with the Framework for Authentication in Cloud-based IP Traceback (FACT) schemes. The incentive assignment scheme is enhanced with resource usage support levels. The temporal token integrity verification is carried out with limited computation complexity. The cloud clod resources and ISP data are managed with security using the Incentive based Framework for Authenticating Cloud IP Traces (IFACT) scheme. The service request duration and frequency measures are analyzed to control the DDoS attacks.

Index Terms: IP traceback, Cloud based traceback services, Temporal tokens and Distributed Denial of Service (DDoS) attacks

1. Introduction

The most important Cloud entity and the principal quality driver and constraining influence are, of course, the user. The value of a solution depends very much on the view it has of its end-user requirements and user categories. Four broad sets of nonexclusive user categories: System or cyber infrastructure (CI) developers; developers of different component services and underlying applications; technology and domain personnel who integrate basic services into composite services and their orchestrations and delivers those to end-users; and, finally, users of simple and composite services. User categories also include domain specific groups and indirect users such as stakeholders, policy makers and so on. Functional and usability requirements derive, in most part, directly from the user profiles.

Cyber infrastructure developers are responsible for development and maintenance of the Cloud framework. They develop and integrate system hardware, storage, networks,

interfaces, administration and management software, communications and scheduling algorithms, services authoring tools, workflow generation and resource access algorithms and software and so on. They must be experts in specialized areas such as networks, computational hardware, storage and low level middleware, operating systems imaging and similar. In addition to innovation and development of new "cloud" functionalities, they also are responsible for keeping the complexity of the framework away from the higher level users through judicious abstraction, layering and middleware. One of the lessons learned from, for example, "grid" computing efforts is that the complexity of the underlying infrastructure and middleware can be daunting and, if exposed, can impact wider adoption of a solution.

Service authors are developers of individual base-line "images" and services that may be used directly, or may be integrated into more complex service aggregates and workflows by service provisioning and integration experts. In the context of the Virtual Computing Laboratory (VCL) technology, an "image" is a tangible abstraction of the software stack.

Services integration and provisioning experts should be able to focus on creation of composite and orchestrated solutions needed for an end-user. They sample and combine existing services and images, customize them, update existing services and images and develop new composites. They may also be the front for delivery of these new services; they may oversee the usage of the services and may collect and manage service usage information, statistics, etc. This may require some expertise in the construction of images and services, but, for the most part, their work will focus on interfacing with end-users and on provisioning of what end-users need in their workflows.

2. Related Work

Although many IP traceback methods have been proposed, the majority of research efforts over the past decade in this area can be broadly classified into three categories: marking-based, logging-based and hybrid approaches. We briefly survey the related works accordingly below.

2.1. Marking-based Approaches

In marking-based traceback, routers embed identity information in the IP headers of passing packets to convey network path information to an end-host. Existing MBT methodes can further be divided into Deterministic Packet Marking (DPM) [9] and Probabilistic Packet Marking (PPM). Typically, DPM embeds the first ingress border router's identity information on packets in a deterministic manner, while PPM probabilistically augments packets with partial path information as they traverse in the network [1]. The goal of DPM is to locate the attack source, and the main purpose of PPM is to identify the attack path.

As a representative work in deterministic marking, Belenky et al. proposed to store the source address in the marking fields of passing packets. Although deterministic marking incurs less computational overhead to trace back to the attack source at the end-host side, it lacks incremental deployment property since it assumes that ingress routers are always traceback enabled. Moreover, it may overload the ingress routers by marking each passing packet compared with the probability based measure. To reduce the number of marked packets, authors in [2] presented a flow-level deterministic marking method for traceback. More recently, Yu et al. [6] proposed a marking on demand (MOD) scheme based on the DPM mechanism to dynamically distribute router IDs in both temporal and space dimensions.

One of the pioneering probabilistic marking solutions was proposed by Savage et al., which probabilistically marks packets with router's identity information as they traverse routers through the Internet. Later on, different variants of PPM [4] have been proposed to

improve the scalability and efficiency of probabilistic marking. Adler revealed that an inherent tradeoff exists in PPM between the number of header bits used and the number of packets required to reconstruct the attack path. PPM based approaches are able to reconstruct the attack path only after receiving sufficient marked packets at the end-host, and may generate false positives. Dong et al. presented a comparative summary of different PPM schemes. For the PPM approach proposed by Savage et al. , more than 2500 packets are required to convey network path information to the destination. Other methods require 103×105 collected packets depending on the number of bits used for marking and awareness of network topology.

An important assumption in PPM is that packets in the flow of interest are much more frequent than other normal packets [9]. Otherwise, PPM will incur a long completion delay or even fail for the path reconstruction under low frequency traffic scenarios. Another shortcoming of PPM is that it is difficult to identify the origin of a single packet. While for applications such as attack mitigation, fault diagnosis or path validation [5], it is preferable to achieve fast traceback as well as single packet traceback. The above challenges in marking-based approaches motivate us to design a novel robust traceback acceleration mechanism with the ability to trace a single packet as presented in this work.

2.2. Logging-based Approaches

Logging-based traceback involves the storing of packet digests at intermediate routers on the path toward end-hosts, thus achieving single packet traceback. Zhang et al. presented a topology-aware single packet IP traceback system. The main disadvantage of logging-based traceback lies in that large storage space is required for packet logs. To reduce the storage requirement for logging, Lee et al. [8] proposed flow digesting on routers instead of logging individual packets. Sample Trace is another flow-level logging method using existing xFlow (sFlow, NetFlow and IPFIX) function and BGP information to implement traceback.

2.3.. Hybrid Approaches

Hybrid approaches [3] take advantages of both packet marking and logging, to reduce the number of marked packets when conducting the traceback process and alleviate the high storage overhead at routers. Duwairi et al. proposed two hybrid traceback schemes, distributed link-list traceback (DLLT) and probabilistic pipelined packet marking (PPPM), to reduce the number of packets needed for constructing attack paths in PPM through utilizing packet logging. In DLLT, if a router decides to mark a packet, it first stores the marking information which was written by the previous marking router, and then marks the packet by overwriting the marking field with its IP address [7]. A link list is therefore established to guide the marking information collection from the end-host. PPPM is a logging-assisted marking scheme, which loads traceback messages into packets going to the same destination of these traceback messages. Gong et al. presented a hybrid solution, called HIT, which reduces the storage overhead at routers to one half and could track a single IP packet. The basic idea of HIT is to recursively mark the accumulated information of multiple routers on packets, and log these accumulated path information at some of the routers on the path. Nevertheless, HIT requires relatively large marking field per packet and high storage on logged routers, since the logging is performed on a per-packet basis. RIHT [3] is a hybrid IP traceback scheme for efficient packet logging aiming to have a fixed storage requirement.

The proposed traceback message delivery scheme in PPPM is a related work to our design. Several fundamental differences exist between PPPM and our design. First, PPPM assumes that IP header has enough space to hold the traceback message of a router. While OPM/AOPM is designed based on a general message model, and message fragmentation is explicitly supported. Second,

we consider the trigger- based traceback, and distinguish the internal-flow and external-flow, which is different from PPPM. Same as most existing PPM methods, PPPM traces all traffic flows. Thus, it marks every passing packet in a probabilistic manner. In OPM/AOPM, once routers are triggered to generate traceback messages, they mark passing packets in a deterministic manner until all message fragments are delivered to the end-host. Importantly, their message delivery schemes are different. In PPPM, a router always swaps the marking information of a received marked packet with one of its buffered traceback messages. The frequent swapping operation incurs very high router processing overheads and complicates the implementation.

3. IP Traceback Services in Clouds

IP traceback is an effective solution to identify the sources of packets as well as the paths taken by the packets. It is mainly motivated by the need to trace back network intruders or attackers with spoofed IP addresses, for attribution as well as attack defense and mitigation. For example, traceback is useful in defending against Internet DDoS attacks. It also assists in mitigating attack effects [2]; DoS attacks, for instance, can be mitigated if they are first detected, then traced back to their origins, and finally blocked at entry points. In addition, IP traceback can be used for a wide range of

While many different IP traceback approaches have been proposed, none of them has achieved universal acceptance or practical deployment. The risk of leaking network topology information ranks as the major challenge in hindering the acceptance of traceback techniques. ISPs (Internet Service Providers) are normally reluctant to allow any external party to gain visibility into their internal structure, since such exposure not only leaks sensitive information to their competitors [5], but also makes their networks vulnerable to attacks. For example, an adversary may misuse traceback services to reconstruct an ISP's network topology [6]. As a result, ISPs will not wish to participate if the deployment of traceback could leak any sensitive information. Incremental deployability is another important factor for a viable IP traceback solution; it is unrealistic to expect all ISPs to deploy IP traceback services in their networks at the same time [7]. Unfortunately, existing IP traceback mechanisms are inadequate in providing guarantees on privacy and support for incremental deployment. Besides technical shortcomings, economic inefficiency, such as lack of financial incentive for ISPs, also hinders the practical deployment of existing traceback solutions.

The advent of cloud services, offer a new appealing option to support IP traceback service over the Internet. It provides an opportunity to design a traceback system that is incrementally deployable. Cloud storage also increases the feasibility of logging traffic digests for forensic traceback. With a proper access control mechanism, cloud-based traceback can alleviate ISP's privacy concerns of disclosing its internal network topology. In addition, the pay-per-use nature of cloud service provides incentives to encourage ISPs to deploy traceback service in their networks. Consequently, migrating traditional traceback solutions to cloud becomes more of a natural choice.

In this work, we first present a novel cloud-based traceback architecture, which exploits increasingly available cloud infrastructures for logging traffic digests, in order to implement forensic traceback. Such cloud-based traceback simplifies the traceback processing and makes traceback service more accessible. It not only possesses privacy-preserving and incremental deployment properties, but also increases robustness against attacks and presents high financial motivation. Yet, regulating access to cloud-based traceback service becomes an important problem. In this paper, we also address the access control problem in the cloud based traceback architecture. To this end, we

propose a framework for authentication in cloud based IP traceback, named FACT, which enhances traditional authentication protocols such as the password-based scheme in cloud based traceback. Our key idea is to embed temporal access tokens in traffic flows and then deliver them to end-hosts in an efficient manner. The proposed method not only ensures that the user requesting for traceback service is an actual recipient of the packets to be traced, but also adapts well to the limited marking space in IP header. Evaluation studies using real-world Internet traffic datasets demonstrate the feasibility and effectiveness of our proposed FACT traceback authentication scheme.

4. Problem Statement

Cloud based traceback systems are build with access control policies. Framework for Authentication in Cloud-based IP Traceback (FACT) is adapted for authenticating traceback service queries**.** FACT is a temporal token based authentication framework used in cloud environment. FACT embeds temporal access tokens in traffic flows, and then delivers them to end-hosts in an efficient manner. The following drawbacks are identified from the existing system Complex token delivery process. The marking scheme is not optimized. Incentive management process is not supported. Service request based attacks are not efficiently handled.

5. Security Ensured Source Discovery on Internet Data Communication

IP traceback methods are applied to discover the source and traversed paths of packets. Cloud based traceback architecture is build with traceback services deployed in Internet Service Providers (ISPs). Framework for Authentication in Cloud-based IP Traceback(FACT) is used to handle IP traceback queries. The FACT scheme is enhanced with optimal marking scheme, attack control and incentive management mechanisms. The Framework for Authentication in Cloud-based IP Traceback (FACT) is enhanced with optimal marking schemes. Incentive estimation for ISPs is integrated with the system. The ISP protection is improved to handle service request based attacks. Token digest management overhead is controlled in the system.

The internet attack discovery operations are carried out using the IP traceback method. The cloud resources are used to support IP traceback operations. Cloud storage and computational resources are provided for the IP traceback operations. Authentication is provided with temporal tokens. Resource usage time limit is verified with the tokens. The cloud based IP traceback system is divided into five major modules. They are Traceback Coordinator, Traceback Servers, Authentication Process, Attack Discovery and Traceback Process. The traceback coordinator manages the resources and traceback requests. Traceback servers are placed to interact with the ISPs. User verification process is carried out under the authentication process. Service request based attacks are handled in attack discovery process. Traceback process delivers the traceback results to the users.

6. Performance Analysis

The IP traceback systems are designed to discover the source of the network requests. The Internet Service Provider (ISP) maintains the request details for the users. The requests logs are used for the source discovery process. The cloud based traceback schemes are used to find out the IP traceback operations. The storage and computational resources are adapted for the IP traceback analysis. The system is tested with Framework for Authenticating Cloud IP Traces (FACT) and Incentive based Framework for Authenticating Cloud IP Traces (IFACT) techniques.

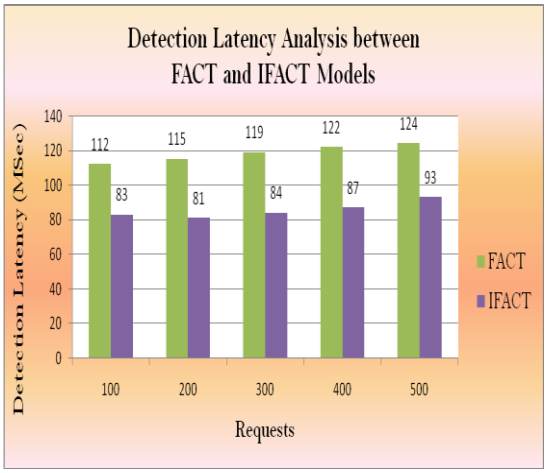

Figure No: 6.1. Detection Latency Analysis between FACT and IFACT Models

The cloud based IP traceback process tested with three performance measures. They are detection latency, false positive rate and false negative rate. The detection latency analysis compares the class identification period for the IP traceback requests. Figure 6.1. Shows the detection latency analysis between the Framework for Authenticating Cloud IP Traces (FACT) and Incentive based Framework for Authenticating Cloud IP Traces (IFACT) methods. The analysis result shows that the Incentive based Framework for Authenticating Cloud IP Traces (IFACT) scheme reduces the detection latency 20% than the Framework for Authenticating Cloud IP Traces (FACT) scheme.

The false positive rate and false negative rate measures are employed to estimate the decision making accuracy level of the system. The false positive rate analysis is estimated with the positive discriminatory results and the falsely assigned positive results. Figure 6.2. Shows the False Positive Rate analysis between the Framework for Authenticating Cloud IP Traces scheme and Incentive based Framework for Authenticating Cloud IP Traces (IFACT) schemes. The analysis result shows that the Incentive based Framework for Authenticating Cloud IP Traces (IFACT) scheme reduces the False Positive Rate 30% than the Framework for Authenticating Cloud IP Traces (FACT) scheme.

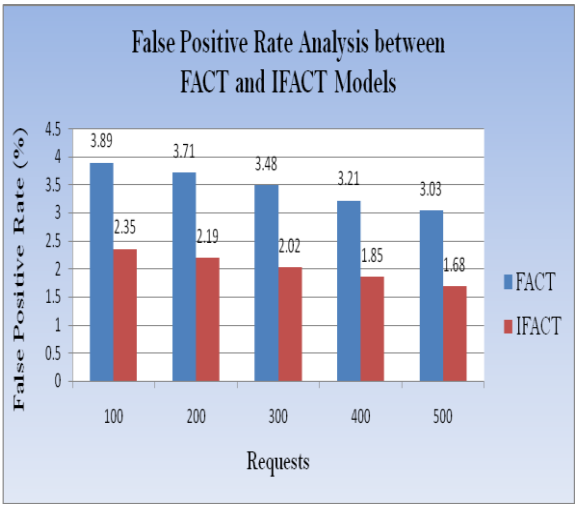

Figure No: 6.2. False Positive Rate Analysis between FACT and IFACT Models

The false negative rate analysis is estimated with the negative discriminatory results and the falsely assigned negative results. Figure 6.3. Shows the False Negative Rate analysis between the Framework for Authenticating Cloud IP Traces (FACT) scheme and Incentive based Framework for Authenticating Cloud IP Traces (IFACT) schemes. The analysis result shows that the Incentive based Framework for Authenticating Cloud IP Traces (IFACT) scheme reduces the False Negative Rate 25% than the Framework for Authenticating Cloud IP Traces (FACT) scheme.

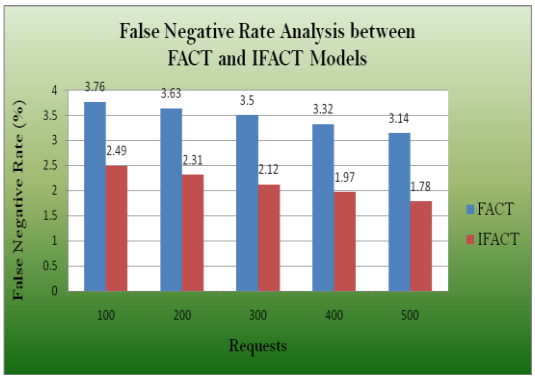

Figure No: 6.3. False Negative Rate Analysis between FACT and IFACT Models

Conclusion and Future Work

The cloud authentication based IP traceback discovery process is developed to utilize the cloud resources for the IP traceback operations. The temporal authentication scheme is used to verify the user requests. The topology protection is provided with the temporal authentication process. Disclosing ISP's internal network topologies, poor incremental deployment and lack of incentives for ISPs parameters are used. The FACT scheme is enhanced to support incentive management model. Attack control schemes are integrated with the system. Optimal marking models are used to manage IP headers. The (FACT) is improved with incentive scheme and attack resistant models. The IP traceback operations can be performed with privacy preserved query submission models. The traceback data values can be maintained and processed in encrypted cloud storage environment.

References:

[1] Long Cheng, Dinil Mon Divakaran, Vrizlynn L. and L. Thing, "Opportunistic Piggyback Marking for IP Traceback", IEEE

[2] V. Aghaei-Foroushani and Zincir-Heywood, "IP traceback through (authenticated) deterministic flow marking: an empirical evaluation," EURASIP Journal on Information Security, 2013.

[3] M.-H. Yang and M.-C. Yang, "RIHT: A novel hybrid IP traceback scheme," IEEE Trans. on Information Forensics and Security, vol. 7, no. 2, pp. 789–797, 2012.

[4] S. Yu, W. Zhou, R. Doss and W. Jia, "Traceback of DDoS attacks using entropy variations," IEEE Trans. Parallel Distrib. Syst., vol. 22, no. 3, pp. 412–425, 2011.

[5] T. H.-J. Kim, C. Basescu, L. Jia, S. B. Lee, Y.-C. Hu and A. Perrig, "Lightweight source authentication and path validation," in SIGCOMM '14, 2014, pp. 271–282.

.

Transactions on Information Forensics and Security, 2016.

[6] S. Yu, W. Zhou, S. Guo and M. Guo, "A dynamical deterministic packet marking scheme for DDoS traceback," in GLOBECOM '13, 2013.

[7] H. Tian and J. Bi, "An incrementally deployable flow-based scheme for IP traceback," IEEE Communications Letters, vol. 16, no. 7, pp. 1140–1143, 2012.

[8] T.-H. Lee, W.-K. Wu and T.-Y. Huang, "Scalable packet digesting schemes for IP traceback," in ICC '04, 2004, pp. 1008–1013.

[9] Y. Xiang, W. Zhou and M. Guo, "Flexible deterministic packet marking: An IP traceback system to find the real source of attacks," IEEE Trans. Parallel Distrib. Syst., vol. 20, no. 4, 2009