Middle-East Journal of Scientific Research 24 (Techniques and Algorithms in Emerging Technologies): 37-40, 2016
ISSN 1990-9233;

DOI: 10.5829/idosi.mejsr.2016.24.TAET23317

Sensitive Data Leak-Detection Using Mobility Key Service

T.R. Srinivasan and S. Saranya

Computer Science and Engineering, Vivekanandha Institute, Tiruchengode, India

Abstract: Statistics from tackle security problems, research institutions and government organizations show that the number of data-leak instances has grown-up rapidly in recent years. There exist solutions detecting unintentional sensitive Data leaks caused by human mistakes and to provide alerts for organizations conventional technologies for data leakage avoidance rely on the terminal or boundary control which is difficult for data leakage in spread environment. However, this secrecy requirement is hard to satisfy in practice, as detection servers may be compromised or outsourced. In this paper, we present a privacy- preserving data-leak detection (DLD) solution to solve the issue where a special set of sensitive data digests be used in detection. The advantage of our method is that it enables the data owner to hand securely over the detection operation to a semi-honest provider without revealing the sensitive data to the provider. For this, the user profile is created using the local knowledge base so that only required data can be given intended for avoiding data leakage and misuse.

Key words: Data Leak · Network Security · Privacy · Collection Intersection · User profile

INTRODUCTION

Detecting and preventing data leaks requires a set of corresponding solutions, which may include data-leak detection, data confinement, stealthy malware detection and policy enforcement. Data leakages are detected and avoided and the misuse is a great difficult issue for organizations. Network dataleak detection (DLD) typically searches for any incidence of sensitive data patterns and performs deep packet inspection (DPI). DPI is a technique to analyze payloads of TCP/IP packet for inspecting application layer data, e.g., HTTP header/content. Alerts are triggered and traffic passes a threshold when the amount of sensitive data found. This challenge becomes more difficult when trying to detect and prevent data leakage and misuse performed by an insider having legal permissions to access the organization's systems and its sensitive data.

The dataleak detection solution which can be deployed and outsourced in a semi-honest detection environment. The fuzzy fingerprint technique be used data privacy during data-leak detection operations.

By compressing data security prevention boundary to data itself, they make all kinds of security control mechanisms associated with data usage more closely and change the static and passive data protection conception. "Kang" [1] also designed a hardware architecture which integrates data and signature management software especially for data leakage protection of mobile storage device. "Kuhn" [2] applied trusted computing to disk encryption and secure latent control, preventing data from leaking in work-in- progress. "Yin Fan" [3] proposed a reliability-based distributed data leakage protection model to extend files from leaking. Berger's [4] trusted virtual datacenter (Trusted Virtual Datacenter, TVDc) structure that is grouping the virtual machine and the underlying data resources further based on TVDs according to security needs of centralized data services, etc.

DLD provider can be prevented from collecting exact knowledge on sensitive data and the collection of potential leaks is composed of noises and real leaks. The data owner who post-processes the potential leaks sent back by the DLD provider and then determines whether there is any real data leak.

Model and Overview: The privacy-preserving data-leak detection problems are of with a threat model, a security goal and a privacy goal.

Corresponding Author: T.R. Srinivasan, Computer Science and Engineering, Vivekanandha Institute, Tiruchengode, India.

Security Goal and Threat Model: Three causes for sensitive data to appear on the outbound traffic of an organization, including the legitimate data use by the employees.

Case I Inadvertent Data Leak: The sensitive data is accidentally leaked into the outbound traffic by a legitimate user. This paper focuses on detecting this type of accidental data leaks over supervised network channels. Inadvertent data leak may be due to human errors such as forgetting to use encryption, inaccurately forwarding an internal email and attachments to outsiders, or due to application flaws [5]. A supervised network channel is either an unencrypted or an encrypted channel where the content in it can be extracted and checked by an authority

Case II Malicious Data Leak: A rogue insider or a piece of stealthy software may take sensitive personal or organizational data from a host. Because the malicious opponent can use strong private encryption, [6] steganography or secret channels to disable content-based traffic inspection, this type of leaks is out of the scope of our network-based solution.

Case III Legitimate and Intended Data Transfer: The sensitive data is send through a legitimate user intended for legitimate purposes. In this paper, we assume that the data owner is aware of legitimate data transfers and permits such transfers. So the data owner can tell whether a piece of sensitive data in the network traffic is a leak using legitimate data transfer policies.

Privacy Goal and Threat Model: To prevent the DLD provider from ahead knowledge of sensitive data during the detection process, we need to set up a privacy goal [7] that is corresponding to the security goal above. We model the DLD provider as a semi-honest challenger, who follows our protocol to carry out the operations, but may attempt to gain knowledge about the sensitive data of the data owner. Our privacy goal is defined as follows. The DLD provider be given digest of sensitive data from the data owner and content of network traffic to be examined. We present a privacy-preserving DLD model with a new fuzzy fingerprint mechanism to improve the data defense against semi-honest DLD provider. We generate digests of sensitive data through a one-way function and then hide the sensitive values among other non-sensitive values via fuzzification.

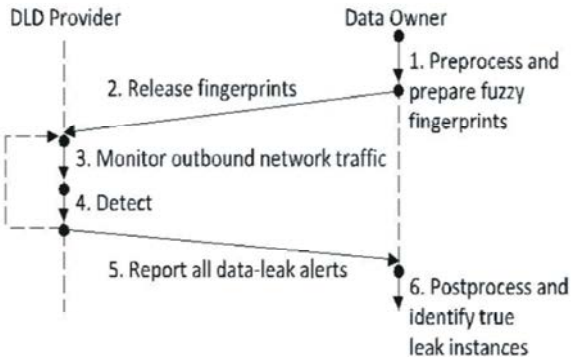

Fig. 1: Proposed Model of Privacy-preserving Data-Leak Detection

DLD provider is used to accessible the traffic count in plaintext. Therefore, in the event of a data leak [8], the DLD provider may learn sensitive information from the traffic, which is near be expected for all deep packet in1spection approaches. Our solution limits the amount of maximal information learned during the detection and provides quantitative certification for data privacy.

Our goal is to offer DLD source solutions to scan massive content for sensitive data exposure and minimize the possibility that the DLD provider learns about the sensitive information.

Scalability: This means that the processing content at variety of scales. For examples, megabytes to terabytes enables the DLD provider which offers an on-demand content inspection.

Privacy: It keeps the sensitive data confidential, not disclosed to the DLD provider or any attacker breaking into the detection system [9].

Accuracy: The ability to identify all leaks and only real leaks in the content, which implies low false negative/positive rates for the detection.

Privacy-Preserving Data-Leak Detection: Network-based Data-Leak Detection (DLD) technique is the main feature in detection of revealing contents of sensitive data. Instead, only a small amount of specialized digests are needed. Proposed technique is the fuzzy fingerprint detection which detects an accidental data leaks due to human errors or application flaws. The privacy-preserving [10] feature of our algorithms minimizes the exposure of sensitive data and enables the data owner to delegate safely the detection to others.

The two most important players in proposed abstract model are the organization (i.e., data owner) and data-leak detection (Server).

DLD provider inspects the network traffic for potential data leaks. The inspection can be performed offline without causing any real-time delay in routing the packets. The DLD provider may attempt to gain knowledge about the sensitive data.

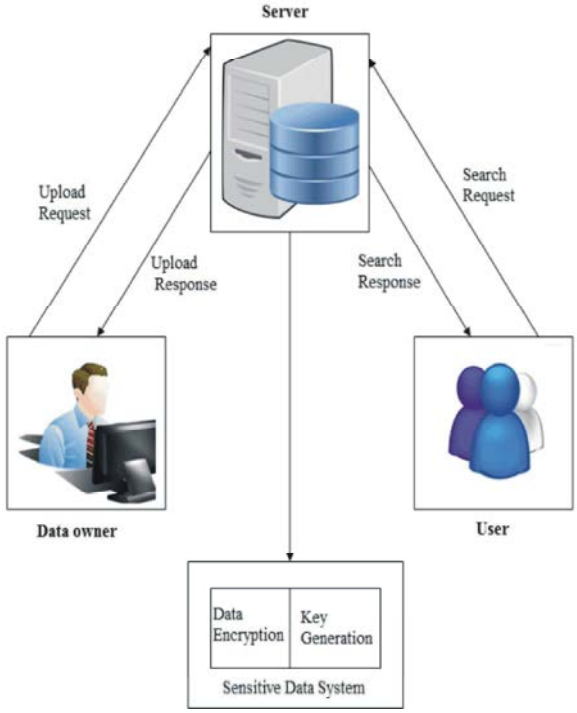

Fig. 2: Privacy-Preserving Data-Leak Detection

High Secure Encryption: The process of making data unreadable by other humans or computers for the purpose of preventing others from gaining access to its contents. Encrypted data is generated using an encryption program such as PGP, encryption machine, or a simple encryption key and appears as garbage until it is decrypted. As first publicly accessible, from the NSA for the classification "top secret" approved cipher, the Advanced Encryption Standard (AES) is one of the most frequently used and most secure encryption algorithms available today [11].

Collection Intersection: To protect dynamically changing data like source code or documents with constant development or keystroke data, the digests continuously updated the detection. The question is raised on issues of how the dynamic data detection efficiently detect the dynamic data to investigate the community with network based approach.

Data Leakage: Proposed privacy-preserving data-leak detection method supports a practical data-leak detection as a service and knowledge is minimized such that a DLD provider gain during process. Six different operations are executed by data owner and DLD provider. PREPROCESS run is included by the data owner which prepare the digests of sensitive data, RELEASE of data owner sends digests to the DLD provider, MONITOR and DETECT collects the outgoing traffic of organization and compute the digests of traffic content and identifies potential leaks, REPORT for the DLD provider used to return data-leak alerts to the data owner where there may be false positives (i.e., false alarms) and POSTPROCESS pinpoints the true data-leak instances. When trying to detect and prevent data leakage and misuse performed by an insider having legitimate permissions to access the organization's systems and its sensitive data.

Dual Key Description: Encryption key has to be split it out into two key if encryption key has eight digit first four digit first off send throw system then second 4digits second off key will be provided via mobile phone(SMS service). If receiver enters the companied key, only he can decrypt the received file.

An Active Data Leakage Prevention Model

Model Idea: The main idea of active data leakage prevention model is to add Secure Data Container (abbreviated as SDC) to achieve active security, as shown in Figure 2, SDC is equivalent to adding a protection shell for documents. Data and security attributes are encrypted and packaged, which are transparent to the upper applications. SDC is a dynamic virtual isolation environment for processes, controlling file access, network access and inter-process communication for processes accessing sensitive content.

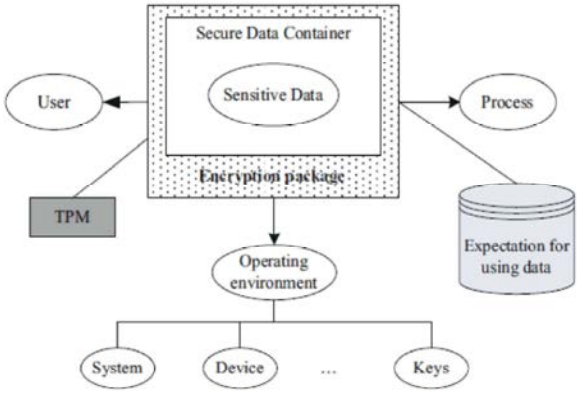

Fig. 3: Active Data Leakage Prevention Model

The Process can only use the decrypted data in SDC. All operations to write data to a non-trusted storage or sent data to non-trusted process will be prohibited. Neither authorized normal users nor illegal processes can leak protected sensitive data out. The integrity itself and data encryption or decryption keys of the SDC are guaranteed by the underlying TPM module. When processes access sensitive data, SDC will actively detect the integrity and security of the related usage environment, involve platforms, hardware platforms and decryption keys, etc. It ensures that data is used by authorized users in trusted environment and complies with data protection usage expectation by authenticating users and processes [12].

CONCLUSIONS

From this, we conclude that the privacy-preserving detection method is used to secure sensitive data from the exposure. Using some special digests, the disclosure of the sensitive data is kept to a minimum during detection. The conducted extensive experiments to validate the accuracy, privacy and efficiency of our solutions. We propose an active data leakage prevention model. By adding a secure data container to execute security prevention mechanism, the model can ensure that data be used in a trusted and controllable environment. Based on the model an implementation framework of active data leakage protection is given.

ACKNOWLEDGMENT

This research was supported by my Head of the Department, Prof. T.R. Srinivasan. I thank Mr. Chandra Mohan for assistance with a technique and Mr. Sathish for comments that greatly improved manuscript.

We thank famous persons Dr. Karunanithi and Dr. K.C.K. Vijaykumar, for sharing their pearl of wisdom with us during this research and we thank Mrs. E. Menaka and Mr. R. Rajagopal reviewers for their so-called insights. Mrs. Leema Mathayi helped with some technical stuff. We are also immensely grateful to my family for their comments and support, although any errors are our own and should not tarnish the reputations of these esteemed persons.

REFERENCES

1. Aho, A.V. and M.J. Corasick, 1998. 'Efficient String Matching: An Aid to Bibliographic Search', Commun. ACM, 18(6).
2. Borders, K. and A. Prakash, 2009. 'Quantifying Information Leaks In OutboundWeb Traffic', in Proc. 30th IEEE Symp. Secur, pp: 129-140.
3. Borders, K. and E.V. Weele, 2009. 'Protecting Confidential Data On Personal Computers With Storage Capsules', in Proc.18th USENIX Secur. Symp, pp: 367-382.
4. Burkhart, M. and M. Strasser, 2010. 'SEPIA: Privacy-Preserving Aggregation Of Multi-Domain Network Events And Statistics', in Proc.19th USENIX Conf. Secur. Symp., pp: 15.
5. Jung, J., A. Sheth and B. Greenstein, 2008. 'Privacy oracle: A System For Finding Application Leaks With Black Box Differential Testing', in Proc. 15th ACM Conf. Comput.Commun. Secur., pp: 279-288.
6. Kleinberg, J. and C.H. Papadimitriou, 2001. 'On The Value Of Private Information', in Proc. 8th Conf. Theoretical Aspects Rationality Knowl., pp: 249-257.
7. Geravand, S. and M. Ahmadi, 2013. 'Bloom Filter Applications In Network Security: A State-Of-The-Art Survey', Comput. Netw., 57(18): 4047-4064.
8. Croft, J. and M. Caesar, 2011. 'Towards Practical Avoidance Of Information Leakage In Enterprise Networks', in Proc. 6th USENIX Conf. Hot Topics Secur.(HotSec), pp: 7.
9. Karjoth, G. and M. Schunter, 2002. 'A Privacy Policy Model For Enterprises', in Proc. 15th IEEE Comput. Secur. Found. Workshop, Jun., pp: 271-281.
10. Jagannathan, G. and R..N Wright, 2005. 'Privacy-Preserving Distributed K-means Clustering Over Arbitrarily Partitioned Data', in Proc. 11th Int. Conf. Knowl. Discovery Data Mining, pp.
11. Li, K. and Z. Zhong, 2009. 'Privacy-Aware Collaborative Spam Filtering, " IEEE Trans. Parallel Distrib. Syst., 20(5): 725-739.
12. Madhusudhanan, B., S. Chitra and C. Rajan, 2015. Mobility Based Key Management Technique for Multicast Security in Mobile Ad Hoc Networks, The Scientific World Journal, Hindawi Publishing Corporation, 2015.