Impact Factor Value 4.046 **e-ISSN: 2456-3463**
National Conference on "Recent Advances in Engineering and Technology" SAMMANTRANA 19
Organized by Government College of Engineering, Nagpur
International Journal of Innovations in Engineering and Science, Vol 4 No.8, 2019
www.ijies.net

Enhanced Detection of Transformed Data Leaks using Hybrid Encryption

Swati P. Bisen[1], Kapesh Raghatate[2]

[1]Student CSE, RCERT Chandrapur,
[2]Professor CSE, RCERT Chandrapur

Abstract -*The data leak of sensitive information on systems has a serious threat to organization data security. Research give that the improper encryption on files and communications due to human errors is one of the leading causes of information loss. T identify the exposure of important data by monitoring the content in storage and transmission. detecting of exposure of sensitive data system is challenging due to data transformation in the content. Transformations result in highly unpredictable leak patterns. In this paper, To detecting complex data-leak patterns an utilize sequence alignment techniques used for. The system achieves good detection accuracy in recognizing transformed leaks. It gives similar system of our algorithms in graphics processing unit to achieve high analysis data. The processes of collective privacy preservation, organizations have to cope with some interesting conflicts. Fo instance when personal information undergoes analysis processes that produce new facts about users' shopping patterns, hobbies, or preferences, patterns. In general, this scenario is beneficial to both users and organizations. However, organizations data sharing in a sharing system, the goal is not only to protect personally identifiable information but also sensitive knowledge represented by some strategic patterns. To demonstrate the high multithreading scalability of the data leak detection method required by a requirement of organization.*

Keyword: *Security, leak, data-leak patterns, privacy preservation.*

INTRODUCTION

The leak of sensitive data on computer systems poses a serious threat to organizational security. Research show that the improper encryption on files and communications due to human errors is one of the leading causes of data loss. Organizations requires a tools to identify the exposure of sensitive data by screening the content in storage and transmission, i.e., to detect sensitive information being stored or transmitted in the clear. When the detectin of sensitiv information is challenging due to data transformation in the content. Transformations result in highly unpredictable leak patterns. In this paper, we utilize sequence [1]. alignment techniques for detecting complex data-leak asymmetric cryptography, facilitate the creation of a verifiable association between a public key and the identity of the holder of the corresponding private key, for uses such as authenticating the identity of a specific entity, ensuring the integrity of information, providing support for non repudiation, and establishing an encrypted communications section.[3] We extensively evaluate the accuracy of our solution with several types of datasets under a multitude of real-world data leak scenarios. Identity Authorization This module allows the user to register their identity into the system with proper input parameters. The key generation centers play a vital role in it, which generates public/ secret parameters . [4] The key authorities consist of a central authority and multiple local authorities. Assume that there are secure and reliable communication channels between a central authority and each local authority during the initial key setup and generation phase. Each local authority manages different attributes and issues corresponding attribute keys to users. They grant differential access rights to individual users based on the users' attributes.[5] The key authorities are assumed to be honest but curious. That is, they will honestly execute the assigned tasks in the system however they would like to learn information of encrypted contents as much as possible.[6]

LITERATURE SURVEY

Unknown cyber-attacks are increasing because existing security systems are not able to detect them. Big data Research techniques that can extract information from a

variety of sources to detect future attacks. The event of new and previously unknown attacks, detection rate becomes very low and false negative increases. To defend against these unknown attacks. Does not detect future Advanced Persistent Threat (APT) detection.[1]
Big data security analytics is used for the growing practice of organization to gather and analyze security data to detect vulnerabilities and intrusions. Security and Information Event Monitoring (SIEM) system. The malicious and targeted attacks have become main subject for government, organization or in dust. Big data analytics is the process of analyzing big data to find hidden patterns, unknown correlations and other useful information that can be extracted to make better decisions. It is used effectively and at the same time, hackers can leave their targets forever.[2] Zero Day Attack Signatures Detection Using Honeypot: Musca, Mirica, E. ; Deaconescu, R. IEEE 29-31 May 2013. Unexpected behavior. Fault distribution studies show that there is a correlation between the number of lines of code and the number of faults. The LCS algorithm on the data content of a number of user going to the same services. Threat is a data content threat that tries to exploit computer application that are unknown to others or undisclosed to the software developer. Vulnerability window which is the time between the first exploitation of vulnerability and when software developers start to develop a countermeasure to that threat. [3]
Cloud Model based Outlier Detection Algorithm for Categorical Data: Dajiang Lei Liping Zhang And Lisheng Zhang, Vol. 6, No. 4, August, 2013. Statically data but there will be a large number of categorical data in real life. Some outlier detection algorithm shave been designed. for categorical data. There are two main problems of outlier detection for categorical data, which are the similarity measure between categorical data objects and the detection efficiency. Outlier detection algorithm for categorical data. Efficient outlier detection can help us make good decisions on erroneous data or prevent the negative influence of malicious and faulty behavior. New data mining techniques try to reduce the influence of outliers or eliminate them entirely. The information manner may result in the loss of important hidden information.[4][5] Cloud Computing-Based Forensic Analysis for Collaborative Network Security Management System: Zhen Chen*, Fuye Han, Junwei Cao, Xin Jiang, and Shuo Chen.1, February 2013. Internet security problems remain a major challenge with many security concerns such as Internet worms, spam, and phishing attacks. Botnets, well-organized distributed network attacks, consist of a large number of bots that generate huge volumes of spam or launch Distributed Denial of Service (DDoS) attacks on victim hosts. A distributed security overlay network with a centralized security center leverages a peer-to-peer communication protocol used in the UTMs collaborative module. These new security rules are enforced by collaborative UTM and the feedback events of such rules are returned to the security center. Collaborative network security management system can not identify the intrustion.[6].

IMPLEMENTATION

Understanding privacy in data mining requires understanding how privacy can be violated and the possible means for preventing privacy violation. In general, one major factor contributes to privacy violation in data mining: the misuse of data. Users' privacy can be violated in different ways and with different intentions. Although data mining can be extremely valuable in many applications (e.g., business, medical analysis, etc), it can also, in the absence of adequate safeguards, violate informational privacy. Privacy can be violated if personal data The proposed method has several advantages.

1. To avoid the attacker.
2. Secrecy of the data should be maintained.
3. The scheme is robust to withstand brute force attacks.

- Securely transforming the data from one place to other by using key attribute.
- It contains the transmission of the data to the long distances.
- Sign to Sign Key parameter in according to the level of authority.
- Hybrid data encryption

Identity Authorization

This module allows the user to register their identity into the system with proper input parameters. The key generation centers play a vital role in it, which generates public/ secret parameters The key authorities consist of a central authority and multiple local authorities. The secure and reliable communication channels between a central authority and each local authority during the initial key setup and generation phase. The authority manages different attributes and issues corresponding attribute keys to users. They important access rights to individual users based on the users' attributes. The key authorities are used to be secured. That is, they will honestly execute the assigned tasks in the system however they would like to learn information of

encrypted contents as much as possible. Confidential Data Interchange This is an entity who owns confidential messages or data and wishes to store them into the external data storage node for ease of sharing or for reliable delivery to users in the extreme networking environments. If a user possesses a set of attributes satisfying the access policy of the encrypted data defined by the sender, and is not revoked in any of the attributes, then he will be able to decrypt the ciphertext and obtain the data.Part of security psychology involves developing a high level of mistrust for anything new to see how this affects people's / device's attitude. For example, let's take Encryption as an example. The objective of encryption is to make the encrypted data look like perfectly random noise. Good algorithm will make it totally random. So no amount of analysis on the output stream would reveal any pattern.

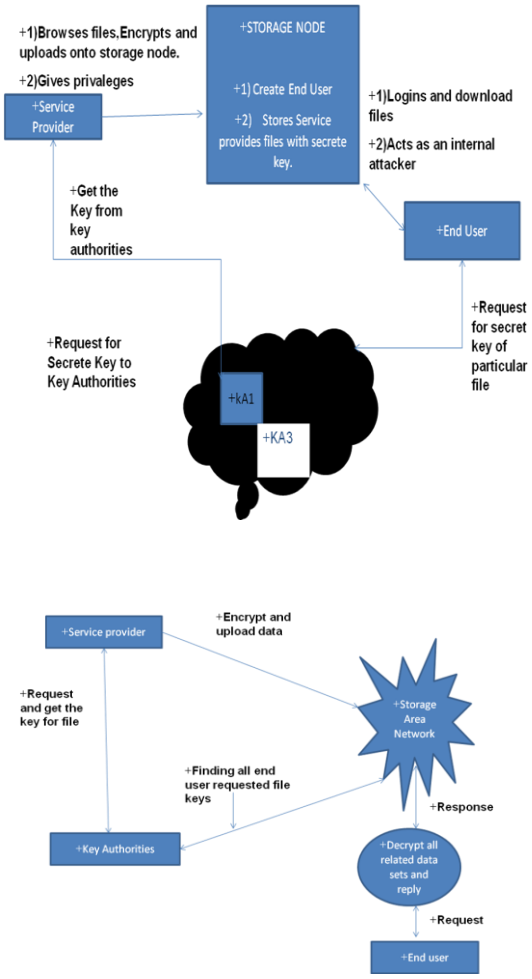

Figure1 System flow

To set up a system in practice, we need to implement the six principles covered in the previous section using mechanisms that tend to be similar from one system to the next.

- Threat Model
- Security Protocol
- Keys and passwords
- Key entropy
- Authentication
- Authorization
- Encryption
- The process of combining a piece of data and a key to produce random-looking numbers is called Encryption.
- It is useful only if a known key can be used to transform the random-looking numbers back to the original data.

CONCLUSIONS

Detecting multiple common data leak scenarios. The parallel versions of our prototype provide substantial speedup and indicate high scalability of our design. For future work, we plan to explore data-movement tracking approaches for data leak prevention on a host. Privacy guarantees are formally defined and achieved with provable security. Experimental results show that our approach is sufficiently efficient for real-world applications. To support efficient handling of multiple auditing tasks, we further explore the technique of bilinear aggregate signature to extend our main result into a multi-user setting, where can perform multiple auditing tasks simultaneously. Extensive security and performance analysis shows the proposed schemes are provably secure and highly efficient.

REFERENCES

[1] X. Shu, J. Zhang, D. Yao, and W.-C. Feng, "Rapid and parallel content screening for detecting transformed data exposure," in Proc. 3rdInt. Workshop Secur. Privacy Big Data (BigSecurity), Apr./May 2015,pp. 191–196.

[2] X. Shu, J. Zhang, D. Yao, and W.-C. Feng, "Rapid screening oftransformed data leaks with efficient algorithms and parallel computing,"in Proc. 5th ACM Conf. Data Appl. Secur. Privacy (CODASPY),San Antonio, TX, USA, Mar. 2015, pp. 147–149.

[3] (Feb. 2015). Data Breach QuickView: 2014 Data Breach Trends.[Online]. /reports/2014YEDataBreachQuickView.pdf, accessed Feb. 2015.

[4] Kaspersky Lab. (2014). Global Corporate IT Security Risks.[Online]. Available: ttp://media.kaspersky.com/en/business-security/Kaspersky_Global_IT_Security_Risks_Survey_report_Eng_final.pdf

[5] L. De Carli, R. Sommer, and S. Jha, "Beyond pattern matching:A concurrency model for stateful deep packet inspection," in Proc. ACMSIGSAC Conf. Comput. Commun. Secur., 2014, pp. 1378–1390.

[6] A. V. Aho and M. J. Corasick, "Efficient string matching: An aid tobibliographic search," Commun. ACM, vol. 18, no. 6, pp. 333–340,Jun. 1975.

[7] R. S. Boyer and J. S. Moore, "A fast string searching algorithm,"Commun. ACM, vol. 20, no. 10, pp. 762–772, Oct. 1977.

[8] S. Kumar, B. Chandrasekaran, J. Turner, and G. Varghese, "Curingregular expressions matching algorithms from insomnia, amnesia, andacalculia," in Proc. 3rd ACM/IEEE Symp. Archit. Netw. Commun.Syst. (ANCS), 2007, pp. 155–164.

[9] Y. Liu, Y. Guo, and C. Liang, "A survey on peer-to-peer video streaming systems," Peer-to-Peer Networking and Applications, Vol.1, No.1, pp.18- 28, Mar. 2008

[10] Y. Chu, S. G. Rao, S. Seshan and H. Zhang, "Enabling conferencing applications on the Internet using an overlay multicast architecture," in Proc. ACM SIGCOM, pp. 55-67, California, USA, Aug. 2001.