Volume 1 || Issue 1 || 21-24
https://doi.org/10.51916/aksh.2020.v01i01.004

Sensitive Information and Data Leakage Prevention Using Secured Encryption Algorithm

*[1]R Usha, [2]Pramod K, [3]D Surbhi
[1,2,3] Department of Computer Science Engineering, Priyadharshini Engineering College, Vellore, Tamilnadu
Email: revatiusha1@gmail.com

Abstract
The information leak of sensitive information on systems includes a serious threat to organization information security. Statistics from security companies, research and analysis establishments and government organizations shows the quantity of data-leak instances have increased speedily in recent years. The shortage of proper cryptography on files and communications due to human delusion is one amongst the most causes of data drop as shown by the statistics. Deliberately planned attacks, unintended leaks and human mistakes cause most of the data-leak incidents. Retransformations result in extremely unpredictable leak patterns. In this paper, the detecting advanced data-leak patterns and sequence alignment technique is employed. For detecting long and inexact sensitive information patterns this algorithmic rule is meant. This detection is paired with a comparable sampling algorithmic rule. This comparable sampling algorithmic rule permits one to match the similarity of two differently sampled sequences. Smart detection accuracy in recognizing transformed leaks is achieved by this method. It implements a parallelized version of our algorithms in graphics process unit to attain high analysis information. Within the case of collective privacy preservation, organizations ought to address some fascinating conflicts. The advantage of our technique is that it allows the data and information owner to soundly delegate the detection operation to a semi honest supplier without revealing the sensitive data to the supplier. However, once organizations share information during a cooperative project, the goal isn't solely to protect in person recognizable data but conjointly sensitive information represented by some strategic patterns to possess the high multithread measurability of the data leak detection.

Keywords
Information Leak Detection, Sensitive Data Patterns, Privacy, Security, Content Scrutiny

Introduction

To shrink the subjection of sensitive document or information, a company needs preventing clear text sensitive information from showing within the repository or transmission. In today's digital world, there's sometimes a pressure amongst safeguarding privacy and distributing the data. [1] Although, in general, sensitive information clearly needed to be private, data owners are typically galvanized, or pushed, to distribute the sensitive information or data. Privacy-Preserving Sharing of Sensitive information, and proposes one logical and reliable representation that functions as a privacy guard to secure parties from revealing apart from the specified depth of sensitive data. The context of simple database-querying execution with two parties: a server that having a database, and a client, executing simple disjunctive equality queries recognizing the subjection of sensitive information is difficult due to data conversion within the content. Conversions end in extremely multitudinous leak styles.[2] In this paper, we tend to deploy sequence alignment technique for acknowledge sophisticated data loss by uneven cryptography, modify for the creation of a verifiable corporation between a public key and therefore the identity different attributes of the holder of the relative private key, for uses like authenticating the individuality of a specific entity, ensuring the integrity of data, giving to stay up for non-repudiation, and beginning an encrypted communications phase.

Proposed System

The motivation of this projected work is to relinquish the approach functions as a privacy guard to secure parties from exposing over the requirement of minimum of their specific sensitive data. PPSSI (Privacy Preserving Sharing of Sensitive Information) preparation ends up in varied challenges that are inscribed in this project. In depth empiric results attest to the utility of earned privacy options and present that our approach incurs quite little overhead. For effective attack detection, huge information integrates attack graph analytical strategies into the intrusion recognition procedure.[3] We should note that the look of doesn't mean to form higher any of the present intrusion detection algorithms; so, involves a reconfigurable virtual network appeal to find and counter the aim to compromise VMs, so intercepting zombie VMs.

E–ISSN: 2582 7456

Volume 1 || Issue 1 || 21-24
https://doi.org/10.51916/aksh.2020.v01i01.004

The recommended technique has few of the benefits.
1. Avoiding the trespasser.
2. Privacy of the data should be preserved.
3. The procedure is strong to resist attacks.

The understanding of data mining's privacy has to know and understand how privacy may be infracted and therefore the possible suggests that for intercepting security crime. In general, one prime factor contributes to security crime in data mining: the incorrect use of data and information.[4] Users' privacy may be infracted in varied ways and with several intentions. The shortage of enough safeguards, offend the informational privacy. One amongst the sources of privacy violation is named data magnets. Privacy may be infracted if personal data is employed for different motive beyond the particular dealings between the individual and a company once the information was gathered. Data magnets are techniques and tools used for assortment of individual's information. Examples of data magnets include expressly gathering of data through on-line registration, acknowledge users through IP addresses, software downloads that requires registration, and indirectly gathering of data for secondary usage. In several cases, users may or may not be conversant in that the information is being gathered or don't aware that information is gathered. Specially, personal information may be used for secondary usage mostly exceeding the users' management and security laws.[5] This state of affairs has results to an uncontrollable privacy infraction not attributable to data processing itself, however basically because of the incorrect use of data.

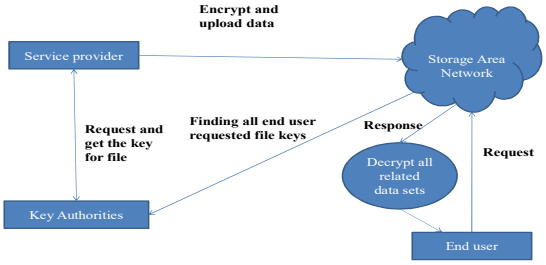

Figure 1: Projected Approach

- Individual Privacy Preservation: The key goal of data privacy is that the safety of individual recognizable information. In general, information is meant as in personally recognizable if it may be coupled, directly or indirectly, to an individual. Thus, once personal information is subjected to mining, the attribute values related to an individual are non-public and should be secured from exposing.[11] Miners are then able to gain the information from international models instead of from the attribute of a selected individual.
- Common Privacy Protection: Conserving personal information might not be enough. Sometimes, we tend to could need securing against learning sensitive information representing the activities of a group. We tend to talk to the protection of sensitive information as collective privacy protection. The goal here is way identical to it one for applied statistically databases, during which security management mechanisms offers combination information concerning teams (population) and, at the same time, prevents disclosure of confidential data about concerning people.[10] However, not like as is that the case for applied statistically databases, another objective of common privacy protection is to shield sensitive information that may provide competitive advantage within the business world. In the case of collective privacy preservation, organizations ought to address some fascinating conflicts. For example, once personal information undergoes analysis processes that offer new facts concerning users' searching patterns, hobbies, or preferences, these facts can be employed in recommender systems to predict or have an effect on their future searching patterns. In general, this state of affairs is useful to each individual and organizations. However, once organizations assign information during a cooperative project, the aim isn't solely to shield in person recognizable information however also sensitive information pictured by some strategic patterns. To grow the protection level this projected theme overcomes the restriction of hybrid cryptography algorithmic rule projected.[9] The projected enhanced theme includes Triple DES, SHA1, and RSA. Triple DES (Variant of DES) strengthens the protection of data transmission. The key size is accrued in Triple DES to confirm extra security through cryptography capabilities. The aim behind for choosing triple DES instead of Double DES is that in double DES algorithmic rule the key used for cryptography and decoding is suspected to meet-in-middle attack. By running the triple DES algorithmic rule in succession, it enlarges the dimensions of the key with 3 completely different key. It produces forty-eight passes through the algorithmic rule.[6]

E–ISSN: 2582 7456

Volume 1 || Issue 1 || 21-24
https://doi.org/10.51916/aksh.2020.v01i01.004

This could be troublesome to implement because of the resultant secret's 168 bits therefore there's choice provided of 2 keys in triple DES that execute through a method known as Encrypt-Decrypt-Encrypt (EDE).
1. Encrypt – cryptography is enforced to the data by using one key.
2. The encrypted information is decrypted using second key.
3. Encrypt- Ultimately, the decrypted information from second step gets encrypted using second key.
Double DES does not truly provide us that much more security than DES and therefore the key distribution problem and additionally to triple DES, SHA1 to verify the integrity of the information. SHA1 is way safer than MD5. However, MD5 don't offer the protection against collision. So, we've got used secure hash algorithmic rule together of cryptographic algorithmic rule. [7,8]

Figure 2: Flowchart

Research Methodology
ALGORITHM FOR ENCRYPTION
The steps are as follows
1. Take a file [F]
2. Perform the Encryption the plaintext blocks using single DES with DE key Y1.
3. Decrypt output of step1 DO by single DES with key Y2.
4. Encrypt the output of step 2 using single DES with key Y3.
5. Output of step 3 is the cipher text. (T)
6. 160 keys by SHA1 (SHA) DK
7. T= N Y3 (DO Y2 (DE Y1))
8. BLOCK= T+KD\
9. BLOCK is sent to battalion receiver.

Triple DES as an encrypt process, it is possible to use a 3TDES (hardware) implementation for single DES by setting Y1, Y2, and Y3 to be the same value. This provides backwards compatibility with DES. Second variant of Triple DES (2TDES) is identical to 3TDES except that Y3is replaced by Y1. In other words, user encrypt plaintext blocks with Y1, then decrypt with K2, and finally encrypt with Y1 again. Total BLOCK Length is 224 bits values returned by a hash function are called secure hash Algorithm or simply hash value. It is a 160-bit hash function.
ALGORITHM FOR DECRYPTION
Received BLOCK= T+DK
1. Whichever the key
2. c = E3 (DK1 (Y1 (N))) = E3 (N)
3. c = E3 (DK3 (Y2 (N))) = Y2 (N)
4. The secret keys as 1st and 2nd or 2nd and 3rd secret keys remain same.
5. It is possible to use 3DES cipher with a secret 128 bit key.
6. In this case 1st and 3rd secret keys are the same.
7. c = Y1(DK2(Y1(N))
8. If key match data decryption
9. km = D1 (Y2 (DK3(c)))

Result and Discussion
Analysis of the previous system and its constraint on implementation and careful designing involves within the implementation stage with analysis of shift strategies and planning of strategies to attain the shift. The theoretical style is clad into an operating system within the implementation stage. So most important stage may be thought-about in giving the user confidence that the new system can work, and be effective and accomplish successful efficient new system. We tend to reason three causes for sensitive information to seem on the outward-bound traffic of a company, as well as the legitimate information use by the workers.

E–ISSN: 2582 7456

Case I: Unintended Data Leak:
In the outward-bound traffic the sensitive information is fortunately exposed by a certified user. This paper focuses on preventing this kind of accidental information leaks over supervised network channels. Because of human errors like forgetting to use cryptography and without showing responsibility forwarding an inner information, email and attachments to outsiders unintended and information leak could happen.
Case II: Malicious Data Leak:
A lurking software, and malicious or scamp insider could steal sensitive personal or structured information from a host. As a result of the malicious antagonist will use steganography technique to disable content-based traffic scrutiny, so these forms of leaks (including covert channels) are out of the scope of our resolution. Host-based defenses ought to be deployed instead.
Case III: Authentic and Deliberated Data Transfer:
The sensitive information is sent by an authentic user supposed for admissible functions. In this paper, we tend to assume that legitimate information transfers use encoding like SSL, which permits one to tell apart it from the unplanned information leak. Therefore, in what follows we tend to assume that plain text sensitive information showing in network traffic is barely because of unplanned information leaks.

Conclusion

In this approach the privacy is achieved secured with empirical security. The data is shielded from insider and outsider attacks. The experimental results shows that this approach is sufficiently efficient for real-world applications. We have got used extremely secured protection by using combination of algorithmic rule. The security and performance analysis shows the projected schemes are incontrovertibly secure and extremely efficient.

References

[1] K. Ramya, D. Ramya Dorai, Dr. M. Rajaram “Tracing Illegal Redistributors of Streaming Contents using Traffic Patterns” IJC A 2011.
[2] A. Asano, H. Nishiyama, and N. Kato, “The Effect of Packet Reordering and Encrypted Traffic on Streaming Content Leakage Detection” Proc. Int’l Conf. Computer Comm. Networks (ICCCN ’10), pp. 1 6, Aug. 2010.
[3] O. Adeyinka, “Analysis of IPSec VPNs Performance in a Multimedia Environment,” Proc. Fourth Int’l Conf. Intelligent Environments, pp. 25- 30, 2008
[4] S. Amarasing and M. Lertwatechakul, “The Study of Streaming Traffic Behavior,” KKU Eng. J., vol. 33, no. 5, pp. 541 553, Sept. /Oct. 2006.
[5] M. Dobashi, H. Nakayama, N. Kato, Y. Nemoto, and A. Jamalipour, “Traitor Tracing Technology of Streaming Contents Delivery Using Traffic Pattern in Wired/Wireless Environments,” Proc. IEEE Global Telecomm. Conf., pp. 1 5, Nov. /Dec. 2006.
[6] Mr. Sagar Prasad, Ms.Malti Nagle, Mr.Tarique Zeya Khan” Prevention of data content leakage with secured encryption algorithm”, International Journal of Advanced Research in Computer Engineering & Technology (IJARCET)”, Volume 5, Issue 12, December 2016.
[7] Y. Chu, S.G. Rao, S. Seshan, and H. Zhang, “Enabling Conferencing Applications on the Internet Using an Overlay Multicast Architecture,” Proc. ACM SIGCOMM, pp. 55 67, Aug. 2010.
[8] LAKSHMANA CHARI, Dr G. RAMA SWAMY” Content Leakage Detection for Trusted Delivery Networks using DRM Technology,” International Journal of Computer Engineering In Research Trends”, Volume 2, Issue 11, November-2015, pp. 872-876
[9] Pooja Pawar, Supriya Palwe, Shweta Munde, Priyanka Gadhave, Mrs. Shikha Pachouly, ”Privacy Preservation and Detection of Sensitive Data Exposure over Cloud”, International Journal of Advanced Research in Computer and communication Engineering Vol. 5, Issue 3, March 2016.
[10] Hiroki Nishiyama, Desmond Fomo, Zubair Md. Fadlullah, and NeiKato, Fellow,” Traffic Pattern Based Content Leakage Detection for Trusted Content Delivery Networks” IEEE Transaction on Parallel and Distributed System, Volume 25, No 2 Feb 2014.
[11] Xiaokui Shu, Jing Zhang, Danfeng (Daphne) Yao, Senior Member, IEEE, and Wu-Chun Feng, “Fast Detection of Transformed Data Leaks”, IEEE Transactions on Information Forensics and Security, VOL. 11, NO. 3, MARCH 2016
[12] Shradha V. Raghorte, Dr. Rahila Sheikh, “Security privacy preserving for content leaks” International Journal of Engineering Research in Computer Science and Engineering, Volume 4, August 2017.

E–ISSN: 2582 7456