International Journal of Science and Engineering Research (IJ0SER),

Vol 6 Issue 4 April -2018
3221 5687, (P) 3221 568X

Effective Framework For Leakage Pattern Identification

M.Pavithra, Dr P.Tamil Selvan,
M.sc Computer Science, Assistant professor,
Department of computer science,
Karpagam Academy Of Higher education

Abstract:Information leakage leads severe threats to organization which handles sensitive information's like money transaction details, client's financial details, credit card details and patient information's so on. Currently water marking technology plays a vital role in organization to provide ownership of the organizational data and to provide security against modification issues. But that fails with leakage pattern discovery. In this research we propose sequence alignment technique for predicting leakage patterns even in large scale data distribution environment. In large scale data distribution environment detection of data leakages with pattern discovery is very difficult to analyze. Unlike existing approaches, our proposed algorithm efficiently tracks data leakage along with the leakage probability of patterns.

Keywords: ***data leakage, privacy preservation, sequence alignment algorithm***

A.INTRODUCTION

In a distributed computing environment the data is publically accessed so some people may misuse and can anonymously access the data. Hence there is a need to detect leakage patterns. Cryptographic schemes were proposed to secure user data but there is a need to build effective leakage detection framework. Generally, watermarking was used to prevent the ownership of data. But those schemes only provide authority but that fails to detect unauthorized leakage patterns. This research proposes an effective framework for identifying leakage patterns in a distributed environment by using sequence alignment algorithm.

B.RELATED WORK

Data leakage detection technologies were surveyed [1] based on the algorithms this survey evaluated Evaluation of explicit data request algorithms, Evaluation of sample data request algorithms. Data Allocation strategy[2] for the allocation of data that need to be sent, adds fake objects to the data and optimization strategy were discussed but it states that those technique fails in distributed environment. Rupesh mishra and DK chitre proposed guilt agent detection [3] in which an admin does not provide with the needs of the agents to improve the chances of finding the malicious agent. Allocation of data's is given more importance. For effective identification distribution algorithm was proposed which can predict the leakage patterns even in distributed environment. Data leakage detection was utilized in cloud [4] that effectively analyzed the security implications that can provide security against active and passive attacks by considering leakage scenario. So, to trace leakage data patterns another approach was proposed which is also a cloud based detection scheme approach and it categorizes identification patterns among multi client architecture.Data Loss/Leakage Prevention (DLP) [5] that is found to prevent from unauthorized data leakages. Unobtrusive Techniques [6] is proposed for Outsourcing-BPO field.

C.PROPOSED WORK

With the quick growth of info business on cyberspace, the info is also unsafe when passing through the unsecure network. The data can be processed and shared legally and illegally with intra and inter networks. So, the data can be illegally accessed by unauthorized authority. The data suppliers and data consumers truly with completely different interest ought to have different roles of rights. So there is a need to shield and verify the info becomes terribly important here. Within the course of doing business, typically sensitive data should be handed over to purportedly trusty third parties.

As an example, a hospital might offer patient records to researchers so they can devise new treatments. Our goal is to notice once the distributor's sensitive information is leaked by agents and if attainable to spot the agent that leaked the info. However, in some cases, it is important to not alter the first distributor's information. For example, if nursing outsourcer is doing our payroll, he must have the precise remuneration and client checking account numbers. If medical researchers are going to be treating patients (as opposed to simply computing statistics), they want correct data for the patients.

Historically, outflow detection is handled by watermarking, e.g., a singular code is embedded in every distributed copy. If that duplicate is later discovered in the hands of unauthorized party, the informant maybe identified. Watermarks maybe terribly helpful in some cases, but again, involve some modification of the first information. Furthermore, watermarks will typically be destroyed if the data recipient is malicious.

This research proposes a sequence alignment technique for predicting leakage patterns even in large scale data distribution environment. In large scale data distribution environment detection of data leakages with

pattern discovery is very difficult to analyze. Unlike existing approaches, our proposed algorithm efficiently tracks data leakage along with the leakage probability of patterns.

BLOCK DIAGRAM

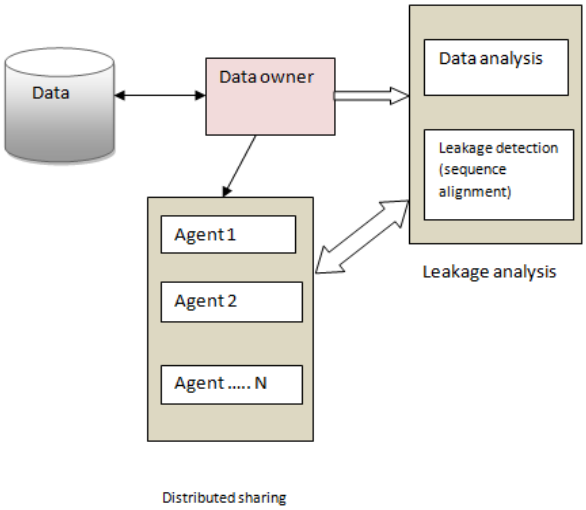

Fig 1: Data Distributed Sharing

In our proposed system we have utilized pattern based detection scheme along with 3DES based encryption to avoid unauthorized data leakage and also secured data sharing for Sensitive data. Two most important end users are included.

- ***Data Owner*** is responsible for Sensitive data he provides authentication to each end user.
- ***Server*** – It behaves like a supervised channel and avoids and prevents from unauthorized data leaks .server only maintains data and that are secured using 3DES based encryption algorithm hence nobody can access

- ***Data user*** can access the data which area all sent by data owner. And without the knowledge of data user, the authority can access

Algorithm: pattern fixing for users

Input: user details U1, U2, U3…Un, Patterns P1, P2…..Pn

Output: pattern fixing

For i=1….n do

If Ui >0 then

Pi->Ui

If Ui>thresh

Block Ui

Delete data content

Else Transfer to Ui

Figure 2 pattern fixing algorithm for sensitive rules

Advantages of proposed system:-

- When compared to Levenshtein distance technique pattern based technique effectively predicts data leak patterns
- Online patterns of client can be traced
- Data security is implemented using 3DES.

D.EXPERIMENTAL SETTING

The proposed leakage detection scheme was implemented using .Net. First users profile has to register by authority. When authority registers user's profile, the fake patterns are also allocated to the user without the knowledge of them. The fake object starts to increases its counts whenever user shares organizational data illegally. This proposed model is implemented by utilizing LAN as a distributed network. The results given below represent connected hosts details in distributed network architecture. Figure 2 represents device configuration details of clients so that admin can monitor users device details and available user details.

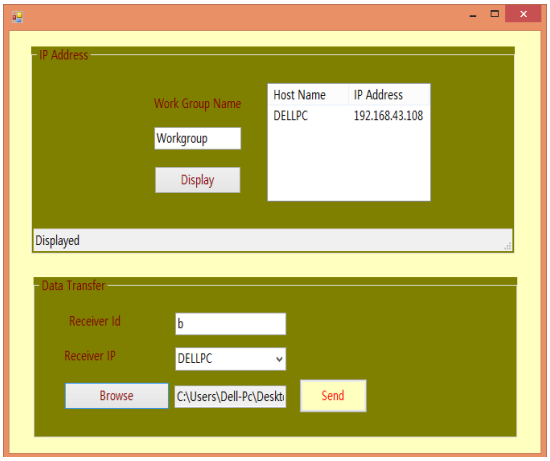

Fig2: Device configuration details of client

Figure 3 analysis data leakage patterns by sequence alignment algorithm which can effectively analyze data leakage patterns in distributed environment.

Sequence alignment technique is used to predict leakage patterns even in large scale data distribution environment. In large scale data distribution environment detection of data leakages with pattern discovery is very difficult to analyze. Unlike existing approaches, our

proposed algorithm efficiently tracks data leakage along with the leakage probability of patterns. In our proposed system we have utilized pattern based detection scheme along with 3DES based encryption to avoid unauthorized data leakage and also secured data sharing for Sensitive data. Figure 3 analysis data leakage patterns by sequence alignment algorithm which can effectively analyze data leakage patterns in distributed environment.

Fig3: *Data leakage detection*

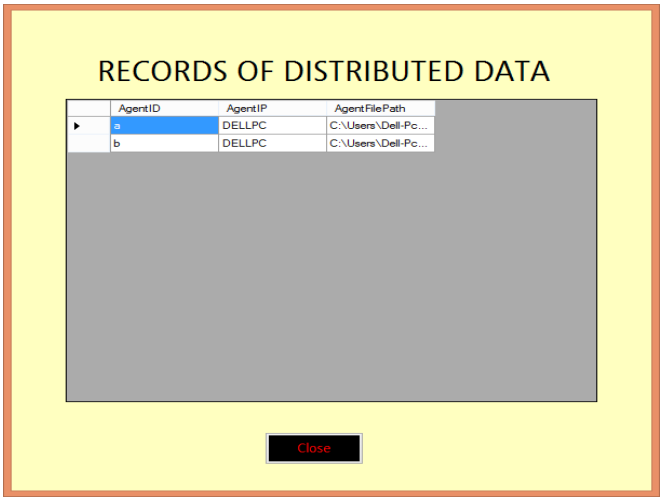

Fig: 4 *details of data records*

Figure 4 describes the details of data records of each and individual data users in a connected authorized network Along with the file path and their identity details.

3DES provides secured data transactions than previous research works. We evaluated that pattern based detection technique performance is well suitable for solving real world large scale and large data oriented problems in a privacy preserving shared manner .They can be used to obtain efficient, privacy-preserving implementations for many dynamic programming algorithms over distributed datasets

E.DISCUSSION

In this section, the result analysis of effective pattern mining method is evaluated. The proposed effective mining method is compared with the existing two methods namely; Fake objects methodology [2], cloud based approach [8]. The performance of proposed framework is evaluated along with the following metrics.

Table 1 Comparison Table

Methodology	**Fake object based leakage detection**	**Cloud based approach**	**Pattern based approach**
Encryption	Not provided	provided	Highly secured using 3DES
Detection accuracy	Less accuracy	Not provided	It efficiently evaluates leakage patterns
Privacy factors	Medium	high	Highly utilized
Probability analysis	Provided	Not provided	Provided

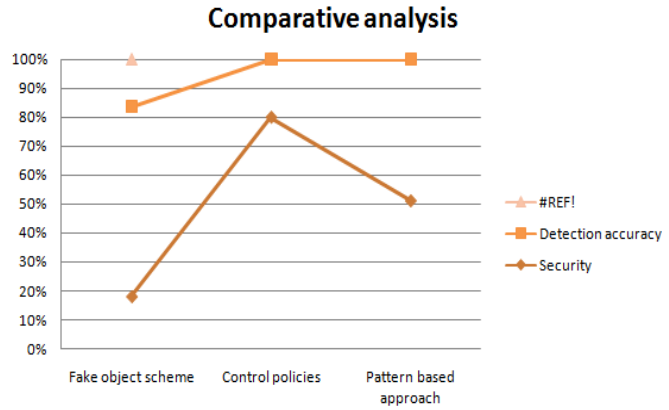

Fig5: *Comparative analysis*

Figure5 represents the comparative analysis of proposed methodology and existing methodologies such as fake object scheme [2], and control policies [18].but in existing secured cryptographic mechanism is not discussed. Control policy scheme only provides attribute based access control but it does not effectively used to evaluate the leakage patterns. By considering all drawbacks of existing system this research is proposed which can work under

even in distributed environment

F.CONCLUSION

Proposed pattern based leakage detection scheme proposed hierarchical architecture for generating leakage detection reports of each data user's .We implemented fake pattern based on assigning and checking leakage patterns based users registered patterns based encryption technique. In large scale data distribution environment detection of data leakages with pattern discovery is very difficult to analyze. In our proposed system we have utilized pattern based detection scheme along with 3DES based encryption to avoid unauthorized data leakage and also secured data sharing for Sensitive data. Pattern based sequence alignment algorithm effectively analyzed data leakage patterns in distributed environment. In future we will concentrate to implement this research in real time large scale chatting applications

G.REFERENCES

[1] Sandip A. Kale C , Prof.S.V. Kulkarni C "Data Leakage Detection: A Survey" IOSR Journal of Computer Engineering (IOSRJCE) ISSN : 2278-0661 Volume 1, Issue 6 (July-Aug 2012), PP 32-35

[2] Rupesh Mishra, D.K. Chitre "Data Leakage and Detection of Guilty Agent" International Journal of Scientific & Engineering Research, Volume 3, Issue 6, June-2012.

[3] Neeraj Kumar, Vijay Katta, Himanshu Mishra & Hitendra Garg, "Detection of Data Leakage in CloudComputing Environment", in Sixth International Conference on Computational Intelligence and Communication Network, 2014.

[4] Alex Ofori Karikari, Joseph Kobina Panford, James Ben Hayfron-Acquah, Frimpong Twum "Detecting Data Leakage in Cloud Computing Environment (A Case Study of General Hospital Software)" International Journal of Scientific Engineering and Applied Science (IJSEAS) - Volume-1, Issue-3, June2015 ISSN: 2395-3470

[5] Vijay Katta, "Detection of Data Leakage in Cloud Computing Environment International Conference .

[6] Mr. Ajinkya S. Yadav1 , Mr. Ravindra P. Bachate2 , Prof. Shadab A. Pattekari3 "Detection of Data Leakage Using Unobtrusive Techniques" OSR Journal of Computer Engineering (IOSR-JCE) e-ISSN: 2278-0661, p- ISSN: 2278-8727Volume 8, Issue 4 (Jan. - Feb. 2013), PP 79-84.

[7] X. Shu, D. Yao, and E. Bertino, "Privacy preserving detection of sensitive data exposure," IEEE Trans. Inf. Forensics Security, vol. 10, no. 5, pp. 1092–1103, May 2015.

[8] K. Borders and A. Prakash, "Quantifying information leaks in outbound Web traffic," in Proc. 30th IEEE Symp. Secur. Privacy (SP), May 2009, pp. 129–140.

[9] S. Jha, L. Kruger, and V. Shmatikov, "Towards practical privacy for genomic computation," in Proc. IEEE Symp. Secur. Privacy, May 2008, pp. 216–230.

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

[11] P.Buneman, S.Khanna, and W.C.Tan, "Why and Where: A Charaterization of Data provenance," Proc.Eighth Int'l Conf. Database Theory(ICDT '01'),J.V. den Bussche and V.Vianu,eds.,pp.316- 330,Jan.2001

[12] P.Buneman and W.C.Tan,"Provenence in Databases",Proc ACM SIGMOD, pp.1171-1173,2007

[13] Y.Cui and J.Widom, "Lineage Tracing For General Data Warehouse Transformations," The VLDB J.vol.12,pp.41-58,2003.

[14] J.J.K.O.Ruanaidh, W.J.Dowling, and F.M.Boland," Watermarking Digital Images For Copyright Protection", IEE Proc.Vision,Signal and Image Processing,vol.143,no.4,pp.250-256,1996.

[15] F.Hartung and B.Girod,"Watermarking of Uncompressed and Compressed Video," Signal Processing, vol.66, no.3,pp.283-301,1998.

[16] S.Czerwinski, R.Fromm,and T.Hodes,"Digital Music Distribution and Audio watermarking," http://www.Scientificcommons.org/43025658,2007.

[17] S.Jajodia, P.Samarati, M.L.Sapino,and V.S. Subrahmanian,"Flexible Support For Multiple Access ControlPolicies,"ACMTrans.DatabaseSystems vol.26.no.2.pp.214-260,2001.

[18] P.Bonatti, S.D.C.di Vimercati,and P.Samarati,"An Algebra For Composing Access Control Policies,"ACM Trans.Information and System Security,vol.5,no.1,pp.1-35,2002.

M.Pavithra completed her in M.sc Computer Science,from Karpagam University in 2018.She has presented in papers in National Conference.

Tamil Selvan P. completed his Ph.D in Computer Science from Karpagam Academy of Higher Education in 2017. He is working as Assistant Professor in Department of Computer Science, Karpagam Academy of Higher Education, Coimbatore. His experience is 10 yrs. He has presented a paper in International Conference. His research interests are Data mining and warehousing.