ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

Insider Collusion Attack on Text File Using Privacy-Preserving Kernel Based System

Ms. Nilima V. Kayarkar [1], Prof. Subhash Malewar[2]

M.Tech Scholar, CSE Dept., WCEM, Nagpur, India[1]

Assistant Professor, CSE, Dept., WCEM, Nagpur, India[2]

ABSTRACT: The unauthorized transfer of important information from computer to outside word is called data leakage. We know that for business purpose, it is necessary to send important data to trusted parties. But during this transfer of data, information is reach at unauthorized place such as website or somebody's laptop. Therefore it is very challenging and necessary to find leakage and guilty person responsible for data leakage.

In this system we used distributor and agent. Distributor means owner of data and agents means trusted parties to whom we send data. The main aim of this system is to find data of owner which is leaked and detect agent who leaked data. In this system, the agent who is responsible for data leakage is the insider. Means agent is a person who is present within the organization. Insider attacks arise not from system errors but from staff working in the company.

KEYWORDS: Data leakage, insider attack, distributor, agent, trusted parties.

I.INTRODUCTION

Now a day data breaching problem due to insider attacks are one of the rapidly increase type of attack. Data breaching is an incident where information is stolen or taken from system without the authorization of the system's owner or we can say that accidental or intentional distribution of data to unauthorized entity is the data leakage. According to the"2015 Verizon Data Breach Investigations Report" insider attack is increases from 2013 to 2015. In 2013 it is 8% and in 2015 it is 20.6% means it is increases at triple rate. Therefore it is necessary to prevent this attack and also find the guilty person responsible for this attack. Many data mining applications contain large amounts of important personal information therefore extensive research has mainly focused on dealing with privacy breaches.
In this system we used two term

1. Distributor: It is the organization or owner of data who send important information to agent.
2. Agent: It is the members within the organization, means it is insider and is semi-trusted. They may leak their own data to outsiders. The insiders leak the data content.

In proposed system we find the guilty person who is responsible for data leakage of text data. To find guilty person we used fake object, means distributor add fake object in the original data before distributed to agent. Depends on this fake data we find guilty person.

II. RELATED WORK

In [1] authors propose an insider collusion attack that carried out on most data mining systems. It discusses how many insiders are sufficient to do this type of attack. In this system insiders within organizations collude with outsiders. This paper introduced several proposed privacy-preserving schemes to counter this attack.
In [2], user collect his personal information like weight, height, food, sports etc. and send this information to the cloud server for data analysis. According to this analysis the app provide best diet plan and related suggestion to user.
In [3] many applications like large-scale financial and medical data analysis, real-time monitoring and social network need to collect large amount of data from different data sources. Then this data is encrypted as well as outsourced.
In [4] authors discussed how to protect important information from the security threats brought by data mining. For this we used user-role based methodology. It uses four different user roles that are commonly used in data mining applications, i.e. data provider, data collector, data miner and decision maker.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

In [5] it proposed Data Annotation Step Agent Interaction analysis (DASAI). DASAI is a logic rule-based static analysis approach used when an attack can take place on a process. DASAI determines in which ways attack can be performed on the process and who are the insiders responsible for this attack.

III. METHODOLOGY OF PROPOSED SYSTEM

In this system model is developed for finding the guilty agents. We add fake objects using kernel-based data mining method while distributing objects to agents. We inject realistic but fake data records to further improve our chances of detecting leakage and identifying the guilty person. Here guilty person is insider. In this system we used automatic fake data generator which generate and add fake data in the original data. This system is used to find out attack carried out on text data. In this system we used two functions.

a) Distribution of file: The methodology of distribution of file is shown in following flow diagram. While distributing file to agent distributer can add fake data in the file. We generate fake data by using automatic fake data generator. For the generation of fake data we used kernel-based data mining algorithm. According to this algorithm, it scans the original file and divides it according to how many percentages of fake data is added. Then it generates fake data, add it in the original file and send file to the agent.

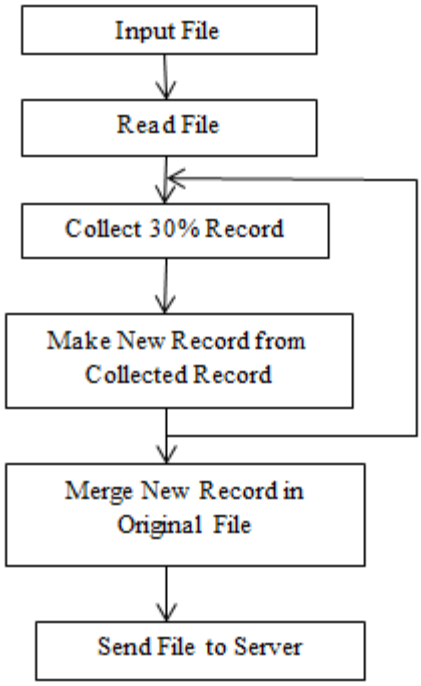

Fig1: Flow diagram of distribution of file

b) Detection of leak file: The methodology of detection of leak file is shown by following flow diagram. Here we give leak file as input. The system read file and match fake record with file rows. It finds percentage of fake record and according to that it decides guilty person responsible for the leakage of text file.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

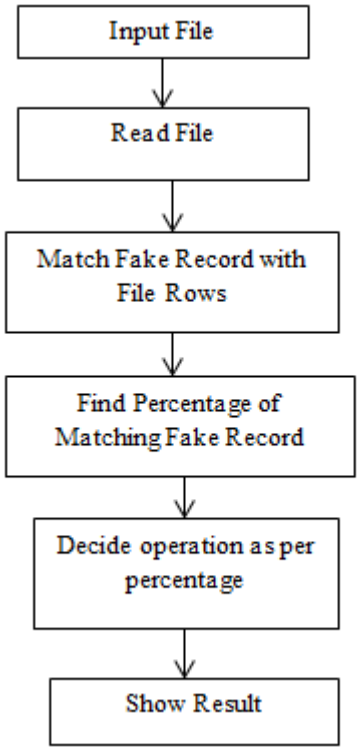

Fig 2: Flow diagram of detection of file

IV. EXPERIMENTAL RESULT

The experimental result of this system can find the guilty person responsible for leakage of text data file. Here we used excel as well as access file for the demonstration. Because in many organizations excel or access file is used for storing the data.

Fig 3: Home Page Window

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

In this window, we have admin login, distributer login, agent sign in and agent registration. According to requirement we make login. First we make distributor and the make login. Inside distributor we make number of agent.

Fig 4: Send File Window

In this window distributor send file to agent. Here distributor select file name, type of file and then add fake data. Finally distributor select agent name to which file is to send and send file. Here to generate fake data we used automatic fake data generator.

Fig 5: Record Detection Window

In this window we have to give file name which is leak and it gives us the name of agent who is responsible for leakage of file. It also finds percentage of matching.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

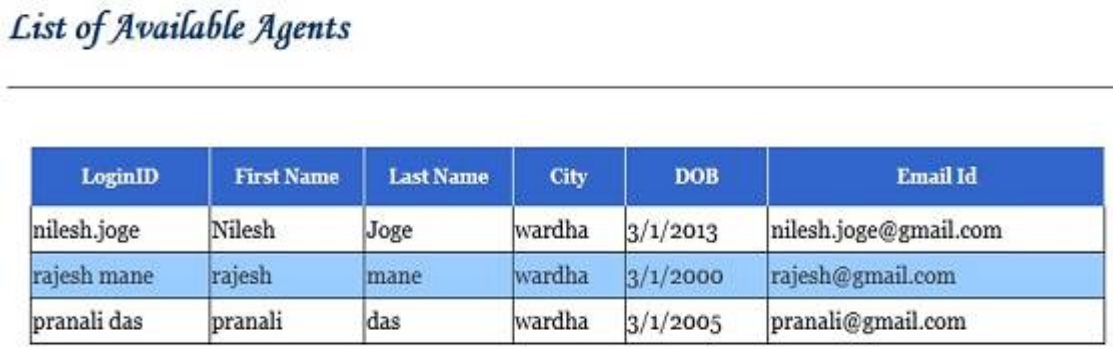

Fig 6: window shows list of agent

This window shows list of all available agents.

Fig 7: Available file window

This window shows list of available file in agent. From this window agent can view the content of file. Here file contain fake record.

V. CONCLUSION AND FUTURE WORK

The propose system can identify the attack carried out only on text data. Using this system we can increase the security and also find the guilty person who is responsible for attack. Here the guilty person is insider. In this system we cannot avoid the attack means the exact recovery of data is not possible.

Therefore in future, we can implement the system which can find guilty person responsible for attack carried on text, image and video file because now a day attack on image and video file get increases. Means we increase the security level of the system.

VI. ACKNOWLEDGMENT

Initially, we would like to thank our almighty in the success of completing this work. We would extend our gratitude to all the experts for their critical comments.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Website: www.ijircce.com

Vol. 5, Issue 3, March 2017

REFERENCES

[1] Peter Shaojui Wang, Feipei Lai, Hsu-Chun Hsiao, And Ja-Ling Wu, "Insider Collusion Attack on Privacy-Preserving Kernel-Based Data Mining Systems", *IEEE Trans*, Apr.2016.

[2] P. S. Wang, S.-W. Chen, C.-H. Kuo, C.-M. Tu, and F. Lai, ``An intelligent dietary planning mobile system with privacy-preserving mechanism," in *Proc. IEEE Int. Conf. Consum. Electron. (ICCE)*,Jun. 2015, pp. 336_337.

[3] F. Liu, W. K. Ng, and W. Zhang, ``Encrypted SVM for outsourced data mining," in*Proc. IEEE Int. Conf. Cloud Comput. (CLOUD)*, Jun./Jul. 2015, pp. 1085_1092.

[4] L. Xu, C. Jiang, J. Wang, J. Yuan, and Y. Ren, ``Information security in big data: Privacy and data mining," *IEEE Access*, vol. 2, pp. 1149_1176, Oct. 2014.

[5] A. Sarkar, S. Köhler, S. Riddle, B. Ludaescher, and M. Bishop, ``Insider attack identification and prevention using a declarative approach," in*Proc.IEEE Secur. Privacy Workshops (SPW)*, May 2014, pp. 265_276.

[6] W. R. Claycomb and A. Nicoll, ``Insider threats to cloud computing: Directions for new research challenges," in*Proc. IEEE 36th Annu. Comput.Softw. Appl. Conf. (COMPSAC)*,Jul. 2012, pp. 387_394.

[7] S. Goryczka, L. Xiong, and B. C. M. Fung, ``*m*-privacy for collaborative data publishing," *IEEE Trans. Knowl. Data Eng.*, vol. 26, no. 10, pp. 2520_2533, *Feb*. 2012.

[8] K. Chen and L. Liu, ``Geometric data perturbation for privacy preserving outsourced data mining," *Knowl. Inf. Syst.*, vol. 29, no. 3, pp. 657_695, Dec. 2011

[9] K. Liu, H. Kargupta, and J. Ryan, ``Random projection-based multiplicative data perturbation for privacy preserving distributed data mining," *IEEE Trans. Knowl. Data Eng.*, vol. 18, no. 1, pp. 92_106, Jan. 2006.

[10] C. C. Aggarwal and P. S. Yu, ``A general survey of privacy-preserving data mining models and algorithms," *in Privacy-Preserving Data Mining Models and Algorithms*. USA: Springer,2008, pp. 10_52.

[11] S. Hartley, *Over 20 Million Attempts to Hack into Health Database*. Auckland, New Zealand: The New Zealand Herald, 2014.

[12] K.-P. Lin and M.-S. Chen, ``Privacy-preserving outsourcing support vector machines with random transformation,"in *Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2010, pp. 363_372.

[13] M. Gönen and E. Alpayd_n, ``Multiple kernel learning algorithms," *J. Mach. Learn. Res.*, vol. 12, pp. 2211_2268, Jul. 2011.

[14] S. R. M. Oliveira and O. R. Za_ane, ``Privacy preserving clustering by data transformation," *J. Inf. Data Manage.*, vol. 1, no. 1, p. 37, 2010.

[15] D. Vizár and S. Vaudenay, ``Cryptanalysis of chosen symmetric homomorphic schemes," *Studia Sci. Math. Hungarica*, vol. 52, no. 2, pp. 288_306, 2015.

[16] J. Vaidya, H. Yu, and X. Jiang, ``Privacy-preserving SVM classification," *Knowl. Inf. Syst.*, vol. 14, no. 2, pp. 161_178, Feb. 2008.
[17] J. Que, X. Jiang, and L. Ohno-Machado, ``A collaborative framework for distributed privacy-preserving support vector machine learning," in *Proc. AMIA Annu. Symp.*, 2012, pp. 1350_1359. [Online]. Available: http://privacy.ucsd.edu:8080/ppsvm/
[18] P. Gaonjur and C. Bokhoree, ``Risk of insider threats in information technology outsourcing: Can deceptive techniques be applied?" in *Proc.Int. Conf. Secur. Manage. (SAM)*, Las Vegas, NV, USA, Jun. 2006.
[19] S. Furnell and A. H. Phyo, ``Considering the problem of insider IT misuse," *Austral. J. Inf. Syst.*, vol. 10, no. 2, pp. 134_138, 2003.
[20] G. B. Magklaras and S. M. Furnell, ``The insider misuse threat survey: Investigating IT misuse from legitimate users," in *Proc. Austral. Inf.Warfare Secur. Conf.*, Perth, WA, Australia, 2004, pp. 1_9.
[21] Cloud Security Alliance (CSA). (2010). *Top Threats to Cloud Computing, Version 1.0*. [Online]. Available: https://cloudsecurityalliance.org/contact/
[22] L. Sweeney, ``Achieving *k*-anonymity privacy protection using generalization and suppression," *Int. J. Uncertainty, Fuzziness Knowl.-Based Syst.*, vol. 10, no. 05, pp. 571_588, Oct. 2002.

BIOGRAPHY

Nilima Kayarkarhas received her B.E. degree in Information Technology in 2012. She is pursuing Master in Technology in Computer Science and Engineering from Wainganga College of Engineering and Management, Nagpur. Her areas of interest include security and data mining.

Subhash Malewar is Assistant Professor at Wainganga College of Engineering and Management, Nagpur. He has completed his M.Tech from G.H.Raisoni Academy of Engineering and Technology. His research areas is cloud computing.