ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

Guilty Agent Detection by Using Fake Object Allocation

Keerthi Pagadala[1], Martha Sheshikala[2], D.Rajeswara Rao[3]

[1]M.Tech Student, Dept of CSE, SR Engineering College, Warangal, A.P, India

[2]Assistant Professor, Dept of CSE, SR Engineering College, Warangal, A.P, India

[3]Associate Professor, KL University, Vijayawada, A.P, India

Abstract: Users among the organization's perimeter perform numerous actions on this information and will be exposed to sensitive info embodied among the information they access. In an attempt for determinative the extent of harm to a corporation that a user will cause exploitation the knowledge she has obtained, we tend to introducing the construct of misuse ability Weight. For hard the M-Score, A misuse ability weight live, this one calculates a score that represents the sensitivity level of the information exposed to the user and by that predicts the power of the user to maliciously exploit the information. By distribution some score that represents the sensitivity level of the information that a user is exposed to, the misuse ability weight may be confirm the extent of harm to the organization if the information is victimised.By exploitation this info, the organization will then take applicable steps to forestall or minimize the injury.

Index Terms: data misuse, Data leakage, security measures, misuse ability weight.

1. INTRODUCTION

Sensitive info like client or patient knowledge and business secrets represent the most assets of a company. So,such info is a lot of essential for the organization's subcontractors, employees, or partners to perform their tasks. At identical time, limiting access to the knowledge within the interests of conserving secrecy may harm their ability to implement the actions that may best serve the organization. Thus, the knowledge run and data misuse detection mechanisms square measure essential in distinguishing malicious insiders. The main focus of this paper is on mitigating run or misuse incidents of knowledge hold on in databases (i.e., tabular data) by associate business executive having legitimate privileges to access the data. There are various things for addressing the malicious corporate executive situation. These strategies that are devised area unit usually supported user behavioural profiles that outline traditional user behavior associate in an issue to alert whenever a user's behavior considerably deviates from the traditional profile. The common approach for representing user behavioral profiles is by associate in the SQL statement submitted by an application server to the information (as a results of user requests), and we tend to area unit extracting varied options from these SQL statements. And another approach focuses on analyzing the particular information exposed to the user, i.e., the result-sets. However, none of the projected strategies think about the various sensitivity levels of the information to that Associate in corporate executive is exposed. This issue has the good impact in estimating the injury that may be caused to a corporation once information is leaked or victimized.

Security-related knowledge measures together with k-Anonymity, l-Diversity. Anonymity is especially used for privacy-preserving and isn't relevant once

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

the user has free access to the info. Therefore, we have a tendency to square measure presenting a replacement thought, the Misuse ability Weight that assigns a sensitivity score to the info sets, thereby estimating the amount of damage which may be inflicted upon the organization once the info has leaked.

In this paper, we tend to developed algorithmic program of knowledge allocation methods for locating the guilty agents that improves the possibilities of distinctive the source. We tend to conjointly take into account the choice of adding pretend objects to the distributed set. Such object don't corresponds to real entities however seem realistic to the agents means pretend objects act as a sort of watermarks for the complete set, while not modifying any original knowledge. If it seems that N agent was given one or a lot of pretend objects that were leaked, then the distributor are often a lot of assured that agent was guilty.

2. PROBLEM SETUP AND NOTATION

A. Entities and Agents

Let the distributor information owns a collection S=1 that consists of knowledge objects. Let the no of agents be A1, A2,...,An [6][10]. The distributor distributes a collection of records S to any agents supported their request like sample or specific request.

- Sample request Little R= SAMPLE (T, mi): Any set of mi records from T are often given to Ui.
    - Ui receive all T objects that satisfy condition.

The objects in T may be of any kind and size, e.g. they will be tuples during a relation, or relations within the information. Once giving objects to agents, the distributor discovers the set S of T is leaked. This implies that another third party known as the target has been caught in possession of S. as an example, the target could also be displaying S thereon computing machine, or maybe as a part of a legal discovery method, then the target turned over S to the distributor. Since the agents (A1, A2... An) have a number of the info, it's cheap to suspect them unseaworthy the info. However, the agents will argue that they were innocent, which the S information is obtained by the target.

B. Guilty Agents

Guilty agents square measure the agents World Health Organization had leaked the information. Suppose the agent say Ai had leaked the information wittingly or unwittingly [1]. Then mechanically notification are going to be the send to the distributor process that agent Ai had leaked the actual set of records that additionally specifies sensitive or non sensitive records. Our goal is to estimate the probability that the leaked information has came from the agents as critical alternative sources.

C. Data Allocation Problem

This paper in the main target information the allocation problem: however the distributor should showing intelligence provide data to agents for up the probabilities of police work a guilty agent. There are unit four instances of this drawback, reckoning on the kind of information requests created by agents and whether or not “fake objects” area unit allowed. Agent makes 2 styles of requests, referred to as sample and specific.

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

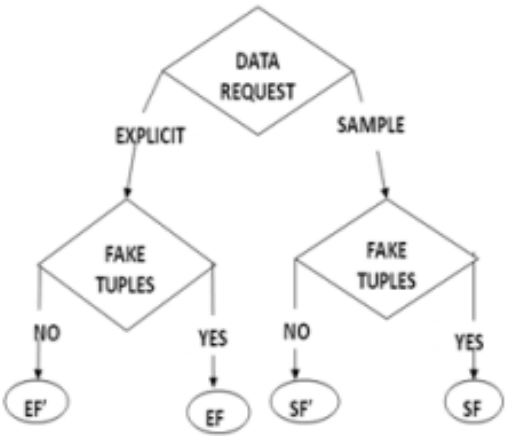

Fig 1: Leakage problem instances

D. Fake Objects

Fake objects ar objects generated by the distributor that aren't in set S. The objects are designed to seem like real objects, and these are distributed to agents at the side of the S objects, so as to extend the possibilities of detective work agents that leak knowledge.

3. RELATED WORK

The guilty agent detection approach we've conferred associated with the info birthplace problems: tracing the lineage of associate degree S object implies basically the detection of guilty agents. It provides an honest summary on the analysis conducted during this field. Suggested solutions square measure domain specific, like lineage tracing for information Warehouses, and assume some previous data on the method an information read was created out of knowledge sources Our downside formulation with objects and sets square measure a lot of general and simplifies lineage tracing, since we have a tendency to don't contemplate any information transformation from Rhode Island sets to S.As so much because the information allocation methods square measure involved, our work was largely relevant to watermarking that was used as a way of the establishing original possession of distributed objects. Watermarks were at the start employed in images, video and audio information whose digital illustration includes right smart redundant. Our approach and watermarking is analogous within the sense of providing agents with completely different types of receiver distinguishing data. However, by its nature, a watermark modifies the item being watermarked. If the article is to be watermarked can't be changed, then the watermark can't be inserted.

4. PROPOSED SYSTEM

The distributor's knowledge allocation to agents has one Constraint and one objective. The distributor's constraint is to satisfying agents' requests, by providing them with variety of objects they request or with all on the market objects that satisfy their conditions.

His objective is in a position to observe associate agent. United Nations agency leaks any portion of his knowledge. The main objective to maximize the possibilities of detection a guilty agent that leaks all his knowledge objects during this paper we have a tendency to develop a model for assessing the "guilt" of agents is developed.

DATA ALLOCATION STRATEGIES:

In this section, we tend to describe allocation ways that categorizes the traditional and licensed agents supported their requests given to the distributor. We tend to square measure coping with each the specific knowledge requests and sample knowledge requests of the agents.

A. Explicit Data Requests

The approved agents send the request for accessible records that contain each sensitive and non sensitive information within the distributor owed set id est. Ri=EXPLICIT (cond1), then the request is alleged to be express information request to the distributor. The distributor cannot take away or alter Re information

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

to decrease the overlap between requests from all alternative agents. that the distributor adds pretend objects at the side of the requested information that don't influence the condition mentioned in agent's request id est. R=. If the distributor is ready to form additional pretend objects, he might more improve the target. we have a tendency to gift the algorithms for express information requests allocation, agent choice for e-random and e-optimal as follows.

Algorithms 1: Explicit Data Request Evaluation

1. Calculates total pretend records as add of faux Records allowed.
2. Whereas total pretend objects > zero.
3. Choose the agent which will yield the best Improvement within the add objective i.e. i=argmax((1\|Ri|)-(1\|Ri|+1))sigma j Ri∩ Rj.
4. Produce pretend record
5. Adding this pretend record to the agent and conjointly to pretend record set.
6. Decrementing pretend record from total pretend record Set.

B. Sample Requests

With sample information requests, agents aren't interested in particular objects. Hence, object sharing isn't expressly outlined by their requests. The distributor has "forced" to portion the sure objects to multiple agents providing range the amount the quantity} of requested objects mi exceeds to the number of objects in set T. There is a lot of information objects the agents request in total, and therefore the a lot of recipients, on average, associate degree object has; and a lot of

Objects are shared among completely different agents; the more difficult it's to notice a guilty agent.

Algorithms 2: Sample Data Request Evaluation

1: Initialize Min_overlap ← 1, the minimum out of the maximum relative overlaps that the allocations of different objects to Ui.

2: for k €{k |tk € Ri} do

Initialize max_rel_ov ← 0, the maximumrelative Overlap between and any set that the allocation of tk toUi

3: for all j = 1,..., n : j = i and tk € R do

Calculate absolute overlap as abs_ov ← | Ri∩ Rj| + 1

Calculate relative overlap as

rel_ov ← abs_ov / min (mi, mj)

4: Find maximum relative as

max_rel_ov ← MAX (max_rel_ov, rel_ov) If max_rel_ov ≤ min_overlap then min_overlap ← max_rel_ov

ret_k ← k

Return ret_k

5. Results

A knowledge distributor has given sensitive data to a group of purportedly trustworthy agents (third parties). a number of the info is leaked and located in associate degree unauthorized place (e.g., on the online or somebody's laptop).

Fig.2 Login for agent & distributor

The distributor should assess the chance that the leaked knowledge came from one or a lot of agents, as hostile having been severally gathered by

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

alternative suggests that.

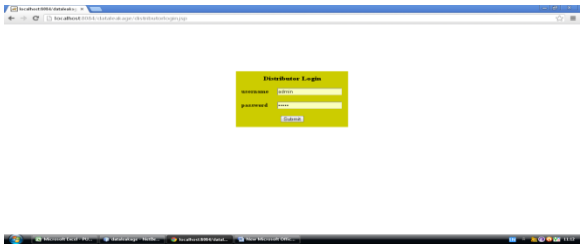

Fig.3 Authentication for Distributor

The distributor should assess the chance that the leaked knowledge came from one or a lot of agents, as against having been severally gathered by alternative suggests that.

Fig.4 Distributor home page

We propose knowledge allocation ways that improve the chance of characteristic leakages.

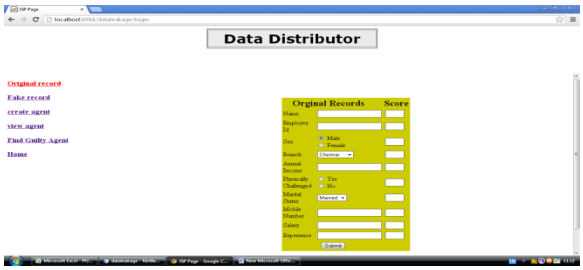

Fig.5 Data allocation stratagies

Fake objects are objects generated by the distributor so as to extend the probabilities of sleuthing agents that leak knowledge. The distributor is also ready to add pretend objects to the distributed knowledge so as to boost his effectiveness in sleuthing guilty agents. Our use of pretend objects is galvanized by the utilization of "trace" records in mailing lists. In some cases we will conjointly inject "realistic however fake" knowledge records to more improve our possibilities of sleuthing outpouring and distinctive the guilty one.

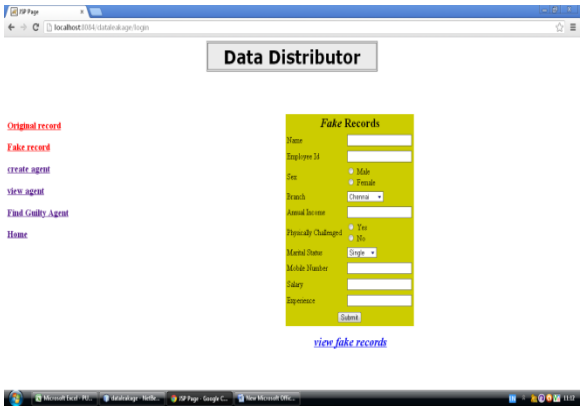

Fig.6 Adding fake records to original

Our goal is to sight once the distributor's sensitive knowledge has been leaked by agents, and if attainable to spot the agent that leaked the info.

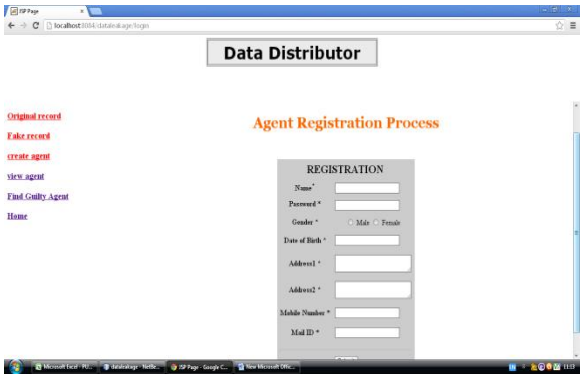

Fig.7 Agent registration process

The distributor should assess the probability that the leaked information came from one or additional agents, as critical having been severally gathered by different means that.

ISSN-2320-7884 (ONLINE)
ISSN-2321-0257 (PRINT)

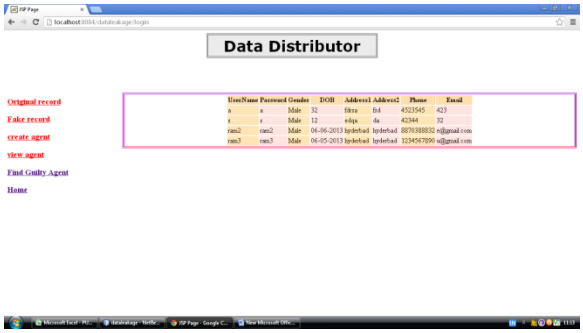

Fig.8 View agent details & requests

The optimization Module is that the distributor's information allocation to agents has one constraint and one objective. The distributor's constraint is to satisfy agents' requests, by providing them with the quantity of objects they request or with all out there objects that satisfy their conditions.

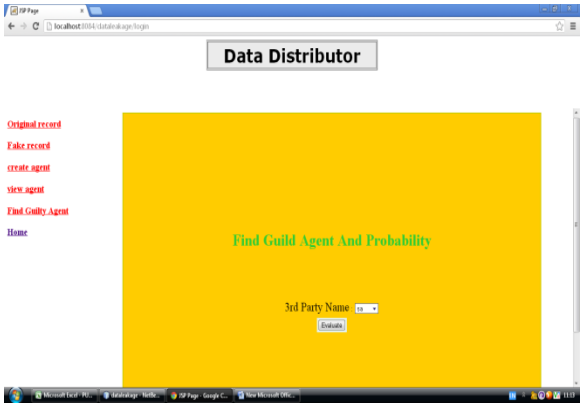

Fig.9 Finding guilty agent & probability

The distributor main objective is to be able to detect an agent who leaks any portion of his data.

Fig.10 Find guilty agent

In this the information distributor goes to search out the guilty agent and conjointly shrewd the M-Score price.

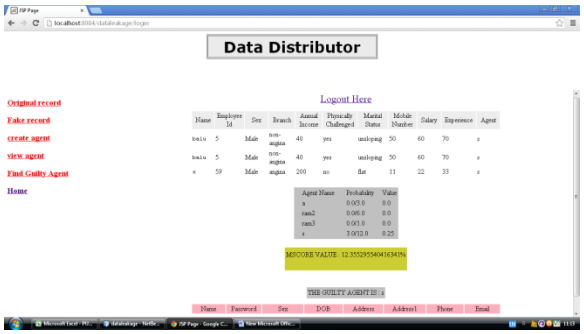

Fig.11 Calculating M-Score value

6. CONCLUSION

From the higher than study we have a tendency to conclude that in a very good world there's no have to be compelled to reach sensitive knowledge to agents World Health Organization might inadvertently or maliciously leak it. In spite of those difficulties, we have a tendency to bestowed that it's attainable to assess the probability that associate agent is accountable for the leak, supported the likelihood that objects may be known by different place. during this planned system we have a tendency to re presenting associate application that is employed to search out the guilty agents and conjointly calculative the likelihood of the leaked knowledge.

ACKNOWLEDGMENT

We wish to acknowledge the efforts of **Pantech Solution Pvt ltd., Hyderabad,** for guidance which helped us work hard towards producing this research work.

7. REFERENCES

[1] N. Sandhya, G. Haricharan Sharma, K. Bhima, "Exerting Modern Techniques for Data Leakage Problems Detect," International Journal of Electronics Communication and Computer Engineering (IJECCE), Vol. 3, Issue (1) NCRTCST, ISSN 2249 –071X.

[2] Sandip A. Kale C1, Prof. S. V. Kulkarni C2, "Data Leakage Detection: A Survey," IOSR Journal of Computer Engineering (IOSRJCE) ISSN: 2278-0661 Vol. 1, Issue 6 (July-Aug 2012), PP 32-35www.iosrjournals.org

[3] P. Saranya, "Online Data Leakage Detection And Analysis,"IJART, Vol. 2 Issue 2, March 2012

[4] P. Papadimitriou, H. Garcia-Molina, "Data Leakage Detection," technical report, Stanford University, 2008.

[5] Shivappa M. Metagar, Sanjaykumar J. Hamilpure ,B. P. Savukar, "Water Marking Technique: An Unique Approach For Detecting The Data Leakage," Volume 2, Issue 5, Sept 2012.

[6] Archana Vaidya, Prakash Lahange, Kiran More, Shefali Kachroo & Nivedita Pandey, "DATA LEAKAGE DETECTION," Vol. 3, Issue 1, pp. 315-321.

[7] Jagtap N.P., Patil S.S. And Adhiya K. P., "Implementation Of Guilt Model With Data Watcher For Data Leakage Detection System," Volume 4, Issue 1, 2012.

[8] Rohit Pol, Vishwajeet Thakur, Ruturaj Bhise, Prof. Akash Kate , " Data leakage Detection," International Journal of Engineering Research and Applications (IJERA) ISSN:2248-9622 ,Vol. 2, Issue 3, May-Jun 2012, pp. 404-410.

[9] Sujana Dommala & M.SreeDevi,"Data Leakage Detection Using Fake Objects," Internati-onal Conference on Computer Science and Information Technology, ISBN: 978-93-81693-86-5, 10th June, 2012-Tirupati.

[10] R. Arul Murugan, Kavitha .E, Nivedha .M, Subashini .S, "Data Leakage Detection And Prevention Using Perturbation And Unobtrusive Analyzes,"International Journal of Communications and Engineering, Volume 04– No.4, Issue: 03 March2012.

[11] Naresh Bollam, Mr. V. Malsoru, "REVIEW ON DATA LEAKAGE DETECTION," International Journal of Engineering Research and Applications (IJERA),Vol. 1, Issue 3, pp.1088-1091

Keerthi.P has received the B.Tech degree in CSE from Vaagdevi college of Engg&Tech, Warangal, A.P, India. Currently pursuing M.tech(CSE) in SR Engineering college, Warangal, A.P,I ndia.

M.Sheshikala has received B.Tech Degree in Computer Science and Engineering from Kamala Institute of Technological Sciences, Huzarabad, AP,I ndia and M.Tech(CSE) from Jayamukhi institute of technological sciences, Narsampet, Warangal, AP, India. Currently working as an Associate Professor(Dept of CSE) in SR Engineering College,Warangal.AP,India.

D. Rajeswara Rao, He is an Associate Professor in K.L. University Vijayawada, India ,
mail: rajesh.dovvada@kluniversity.com