A Novel Approach for Detection techniques for CD and CC in digital forensics

Rahul B. Diwate Dhiraj D.Shirbhate Sonal Channe

M.E. Scholar Asst.Prof [CSE] M.E.Scholar

Abstract- Now a day's there is great attention in the accounting and cyber crime fields because of government regulations in the whole world [1]. Although these regulations force corporations to provide financial transparency, they still commit accounting frauds such as tax evasion. Moreover, companies have substituted paper-work with IT systems such as DBMS, EDMS, and ERP system. So there is a need to focus the attention on discovering financial information in a database server. However, frauds are difficult to observe and detect because the perpetrators did their best to conceal their fraudulent activities. In particular, there is need to consider the case of a covert database server. Secondly a network covert channel is a communication channel that allows two cooperating processes to transfer information on the network in a manner that violates the system's security policy. So for covert channel, in order to solve the problem one detection algorithm can only detect one kind of network covert channel, the detection approach hierarchical and density based cluster was purposed. Because the coding scheme of the covert channel would cause many similar data occurred repeatedly, the detection algorithm cluster based on density can be used to detect several kinds of the covert channels. Moreover, the detection approach cluster based on hierarchy and density is able to tackle of detection a noisy channel. The algorithm can work well to distinguish the covert channel from normal network traffic. This paper proposes a novel approach for detection of covert database server and covert channel,

Keywords: Covert Databases, ERP, EDMS, DBMS, network, network security, steganography, covert channel, cluster.

1. Introduction

In the world, corporations organize their information using database management system infrastructures such as Oracle Database, Microsoft SQL Server, MySQL, etc. The current database technology used is based on distributed network and administration automation. However, it is certainly difficult to notice a database system among many systems. In addition to this, a corporation might not operate correctly by hiding their database server. In this way the investigation might be deterred. Although accomplice hides database system intentionally, even though the perpretator could intentionally hide the database system, the investigators must demand computer data submission based on the reliable electronic evidence.

This paper has been organized as follows. Firstly, provided background knowledge giving an example of hidden database systems. Secondly, illustrated a scenario with covert database systems. Thirdly, outlined a suitable technique to find databases in several investigation environments. Finally, suggested a checklist for detecting concealed databases.

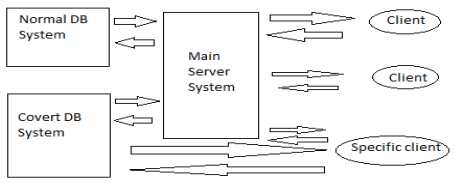

Figure 1: Structure of covert database server

2.Literature review & Related work:

2.1. An Example of Hidden Database System

Figure 1 illustrates a situation where clients usually have access to the normal database server through the main server system, which manages the database server [2]. However, they are not aware of the existence of the covert database server. The specific client usually has access to the normal database server or the covert database server through the main server system. This client can also directly access the covert database server in order to conceal confidential information. If a forensic practitioner looks into this scenario, the specific client will disconnect all the connections related to the covert database server.

Efficient methods to detect the covert database server in these infrastructures are needed. Digital forensic techniques are methods which efficiently analyze digital devices and precisely obtain the evidence.

3. Analysis of Problem

The use of effective and appropriate methods in facilitating projects enhances its effectiveness and efficiency. The method will be applied in System Analysis and Design method where an existing system is studied to proffer better options to solving existing problems.

Server systems could be hidden by manipulating DNS (Domain Name System) in the corporation. Networks could be usually hidden by manipulating IP address. We classify different response strategies according to different cooperation scenarios and the Korean law system.

Criminal uses the covert database for their criminal activities. This paper proposes a methodology for detecting covert database server & covert channel, which would be helpful for forensic investigators. Clients usually have access to the normal database server through the main server system, which manages the database server [3]. However, they are not aware of the existence of the covert database server. The specific client usually has access to the normal database server or the covert database server through the main server system. This client can also directly access the covert database server in order to conceal confidential information. If a forensic practitioner looks into this scenario, the specific client will disconnect all the connections related to the covert database server. Efficient methods to detect the covert database server in these infrastructures are needed. Digital forensic techniques are methods which efficiently analyze digital devices and precisely obtain the evidence.

A diverse range of individuals and groups has found reason to utilize covert channels for communication and coordination. Typically this is motivated by the existence of an adversarial relationship between two parties (e.g., government agencies vs. criminal or terrorist organizations).

4. Proposed Work and Objectives:

4.1. Detection techniques for covert database server

In this section, there is a method for detecting hidden database management systems. An investigator could apply various techniques that are suitable to different scenarios. Initially, we collect network information in the local area network where the covert database is supposed to be. Moreover, user can control the real presence of a database by means of three techniques. The first is by means of Net- BIOS [4], which allows applications on separate computers to communicate over a local area network. The second technique is the so-called ping sweep, which is used to determine which range of IP addresses is mapped to live hosts. Finally, the third technique is the port scanning, which might be used to search live hosts. Finally, we should validate these collected data by means of trace analyzer usually implemented by DB clients.

- Detection technique using ActiveX Data Objects,
- Detection Techniques which Use Traces of DB Client Programs.

If database server is linked to the network, we can attempt to access it by means of IP protocol, with Microsoft ActiveX Data Object technology (ADO). To gain the database access, a user generally needs to know the identifier, the password, the IP address of database server, the port number, and finally the database service name. Interestingly, ADO technology helps to detect and access database server even if we do not know user data (user ID and password).

User might collect registry information by means of registry monitor programs [5]. Lastly, the third method is targeted at gathering log, setting, and database files. Therefore, we might acquire

information about database state after having analyzed internal data of these files.

Interestingly, we might gather information about IP address [6], port number and database service name by using registry monitoring analysis. Moreover, as already pointed out, log file analysis is the method that acquires relevant information from files which are under control by means of file monitoring analysis[7]. We can implement at least one method for this scenario. We can use any of the technology for the detection of the covert database server.

Methods of Detecting Covert channels

In this paper, our focus is on creating a mechanism that can detect covert channels in the network traffic. It has been suggested in the research of Cabuk that detecting a covert timing channel by disclosure the similarity among the inter packet delays. However, the same idea can be used in detecting a covert storage channel. The designer may code the normal value of the field in network packet to covey covert
information, which would cause some same values of the one field in the network packet occurs repeatedly. As a result, programmer can monitor the similarity of the value of some fields in the network packet to decide if there are covert channels exist.

5.2.1. Density and Hierarchical based Clustering

Density-based clustering algorithms characterize the data distribution by the density of each data object. Clustering is the process of identifying dense areas in the object space. Conventional density-based approaches classify a data object as one of the cores of a cluster if it has more than n neighbors within neighborhood [8][2]. Clusters are formed by connecting neighboring core objects and those non-core objects either serve as the boundaries of clusters or become outliers. Since the noises of the data set are typically randomly distributed, the density within a cluster should be significantly higher than that of the noises. Therefore, density-based approaches have the advantage of extracting clusters from a highly noisy environment. In the following, there is a list the definitions of terminologies regarding to density-based clustering convenience of presentation.

However, the performance of density-based clustering is quite sensitive to the parameters of object density, namely, for a complex data set, the appropriate parameters are hard to specify. A density-based clustering algorithm tends to result in a large number of trivial clusters due to the noise. For solving the problem of trivial clusters, in this work there is an algorithm based on hybrid strategy between the hierarchy and density based approaches. Considering of some cluster points formed by the noise with a low density, hierarchical clustering algorithm can be used that generates a hierarchy of nested clusters according their density. Now, there is list the definitions of hierarchical clustering based on density.

It would be used as a criterion to measure the difference of the distribution of two clusters. By measuring the density for each data object, density and hierarchical based cluster captures the natural distribution of the data. Intuitively, a group of covert traffic will form a dense area, and these samples with the highest density within the group become the medoid of the cluster. Therefore, where a cluster is formed by noise objects, it has low density. So some trivial clusters according their densities can be merged.

5.2.2 Detection Algorithm

Detection algorithm of Covert channel is composed of two subroutines, they are clustering algorithm and detecting algorithm. When detecting, the detection algorithm will call the clustering component. Before introducing the detection algorithm, the data obtained from the network sensor should be investigated, which mode can be described as:

$S = T \times X + W$

In the algorithm two stages are given which are as follows:

Stage one: finding core objects &

Stage two: Merging clusters.

By using these stages the core objects & clusters can be found which would be useful in detecting the covert channel.

The algorithm can be suitably described as follows::

Stage one: finding core objects

- Step1: Select one sample Si randomly as the core from the data source S, if there was a core object Rj in the prev set R, satisfied $Si \in N(Rj)$ then go to step2. Or, Si would be taken as the new core object, R=R+{Si},go step3 ;

- Step2: Let F[Si]=Rj, count the sample of the core objects ki.
- **Step3:** If all the samples of the data source are correspond to a core object, go to step4, or go to step1 ; finding core object/outlier.
- **Step4:** Modify the core object set Rjk ,k=1...K ,.Updating scanning times, c=c+1. The initial value of c is 1. If the number of the scanning is little than c, go to step1,start to scan the data set again.
- **Step5:** When ki≥p , Rj is a available core object of one cluster, or otherwise it is invaluable.
- If this is the case then all the samples in the clusters responses to an invaluable core object & will be allocated to an available core object.

Stage two: Merging clusters

- **Step6:** According to the character of the data ,compute the density Vi of cluster Ri and the density Vj of cluster Rj.
- **Step7:** Compute the difference of these clusters cov(V(Rj),V(Ri)), and select the minimum of the them, descript as Cmin. If Cmin<σ, then go to step 8, else go to step 6.(here σ=density of cluster.)
- **Step8:**Computer distance between of Ci and Cj, merging Ci, Cj into a new cluster Ck(k=1,2,3…)
- **Step9:** Until all the data was processed, or go to step6.

5. CONCLUSION

- The clustering method could be robust to noise, outliers, and the parameters.
- It is well recognized that the covert traffic data are usually noisy and the rules behind the data are unknown[9].
- So hacker always takes new ways for hacking.
- The algorithm of cluster is a sub component of the detection algorithm.
- The algorithms can be used to detect various frauds such as e-transaction frauds , terrorist activities etc.

6. Acknowledgement

The making of the seminar needed co-operation and guidance of a number of people. I therefore consider it my prime duty to thank all those who had helped me through their venture. It is my immense pleasure to express my gratitude to Prof. V. T. Gaikwad as guide who provided me constructive and positive feedback during the preparation of this seminar.

I express my sincere thank to the head of department Dr. A. D. Gawande and all other staff members of CSE department for their kind co-operation.

I would like to thank Dr. S. A. Ladhake, Principal of our institution for providing necessary facility during the period of working on this report.

I am thankful to my friends and library staff members whose encouragement and suggestion helped me to complete my seminar.

I am also thankful to my parents whose best wishes are always with me.

References

[1] Law : Pub.L. 107-204, 116 Stat. 745

[2] G. Lee, S. Lee, E. Tsomko, and S. Lee, Discovering Methodology and Scenario to Detect Covert Database System, In Proceedings of IEEE Future Generation Communication and Networking, 2007.

[3] J. Hangdahl, Inside NetBIOS, Architecture Technology Corp. 1990.

[4] D. Sceppa, Programming ADO, Microsoft, 2001.

[5] M. Russinovich and B. Cogswell, FileMon for Windows v.7.04.

[6] Girling C G. Covert Channels in LAN's. IEEE Trans. Software Engineering. 1987, SE-13(2): 292-96.

[7] C. Rowland. Covert channels in the TCP/IP protocol suite. First Monday: Peer-reviewed Journal on the Internet, 2(5),1997.

[8]WAND Research group. NZIX-II trace archive, dataavailable
http://pma.nlanr.net/traces/long/nzix2.html.

[9] C. S. Jensen and R. T. Snodgras IEEE Transactions on Knowledge and Data Engineering, Vol. 6, No. 6, December 1994, pp. 954–974.