ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 9, Issue 6, June 2022

Detecting Data Leakage in Cloud Computing

S.MENAKA[1], S.GOWTHAMI[2],K.SOWMIYA[3], P.SOWMIYA[4], A.THONDRAL NAYAGI[5],D.VANITHA[6]

1,2 Assistant Professor, Department of Computer Science and Engineering .
3,4,5,6 Under Graduate Students , Department of Computer Science and Engineering .
Vivekanandha College of Technology for Women, Tiruchengode, Namakkal, Tamil Nadu,India.

ABSTRACT:

In a virtual and widely distributed network, the process of transferring sensitive data from a distributor to trusted third parties is a constant occurrence in this modern world. It needs to protect the security and durability of the service based on the need of users. The idea of changing the data itself to detect leaks is not a new approach. Typically, sensitive data is leaked to agents, and a particular agent responsible for leaked data should always be available in advance.Thus, the detection of data from the distributor to agents is mandatory.[1] Hence, the system can be applied in any environment for the prevention and detection of any data leakage.[2]

KEYWORDS: Data Leakage, Cloud Computing, Cloud Security, Information Security.

INTRODUCTION:

Nowdays, utilizing the cloud computing technology has become the norm in enterprise, education and government sectors.This is because cloud computing provides much faster, more flexible performance and has more self-help features compared to a standard server platform. Although cloud computing offers many advantages, it has some potential for security issues.The very core technology in cloud computing is virtualization, because its liberate applications, servers, storage and desktop some becoming dependent to physical hardware layers, by abstracting resources in isolated virtual computing environments. In cloud platform, the host running on the platform is called as Virtual Machine (VM) and the common operations performed on VM are replication and migration. During such activities, the risk of data leakage may occur as a result of malformations, software interruptions (e.g hypervisor) or malicious management behavior. In addition, the risk of data leaks could also occur during verification in a communication session when cloud users access a self-care cloud site or dashboard.

LITERATURE SURVEY

Watermark is a secure, unobtrusive, and secure signal embedded in the original content such as an image, video, or audio signal, which generates a marked signal and describes information that can be used for proof of identity or error. The watermark can be found even in a small group of watermarked relation as long as the sample contains other marks.The watermarking software introduces small errors into the object being watermarked[1].Water marking is a technique where a bit pattern is added to the data at a particular position on the tuples and subset of the data. The tuple and subset and their attributes are coded in such a way that they are controlled by a key that can only be accessed by the owner. We do not need access to the original data or watermark pattern in order to obtain the watermark [2] Perturbation is a very useful method in which data is processed and made “less sensitive” before being given to agents. To improve user data

security the model has an encrypted algorithm that provides secure data transfer between the user and TPA to achieve, current algorithms for distributing items to users, in a way that enhances our leaky identification opportunities. With the addition of false data items to the distributed set provide incorrect information if the data is accessed by any other third party [3]. Discrete cosine transform (DCT) is a method of converting signals into components of a basic frequency. The basic idea of DWT is to separate most of the image rotation into a small image of a different landscape and independent waves.Watermark embedded in LSB pixels. The insignificant fraction of the pixel density of the main cover image is coded to provide the watermark. in digital information security products. Multi-resolution data integration is embedded when the image and watermark are both used and converted to a different wavelet form.Watermark is embedded at each wavelet level. To save multimedia data such as audio, video and image, a watermark should be added to key signal components if it is to strengthen the normal signal distortion and malicious attack.But, the modifications of these components may lead to degradation of the data signal.[4]. The analysis on the packet capture is conducted based on detection parameters which are applied as display filters in Wireshark tool. The detection parameters are any username, password, or Active Directory data and other data strings.The methodology adopted for this research work is the Object-Oriented Analysis and Design methodology (OOAD). The existing system from the point of view of objects and similar objects is organized into categories and their features are treated as structures + while their behavior is treated as actions or modes within a mass of the same object. This method is preferred because it is best used to manage a system where different objects are present.It is a structured approach that is used in analyzing and designing a system. It applies object-oriented concepts and develop a set of graphical system model during the development lifecycle of the software.

CLOUD COMPUTING INTRODUCTION

Cloud computing contains three different types of computer services delivered remotely to customers online.Clients typically pay a monthly or annual service fee to providers, to gain access to systems that deliversoftware as a service, platforms as a service and infrastructure as a service to subscribers.The connection of a large number of visible system is known as the cloud. Connected systems may be private or public and the data stored on them may also be categorized as private and public access. For example, a single drive powered by Microsoft is one example of cloud computing.For example, a single drive powered by Microsoft is one example of cloud computing. In one drive every user is given a specific user id and password and can access the data in the cloud anywhere by simply connecting to the internet and using their unique ID and password.

2.DATA LEAKAGE IN CLOUD COMPUTING PLATFORM

A Cloud leak is when sensitive business data stored in a private cloud instance is accidentally exposed to internet.The cloud is part of the internet.The difference is that “the cloud” offers pockets of privatized space that can be used to carry out enterprise scale IT operations.Enterprise data sets,often handled by third party information analytics companies,are often stored unencrypted in the cloud,with exceptation that their data lives within one of these private pockets.Here the pictorial representation of how data leaks in cloud platform shown in fig 1

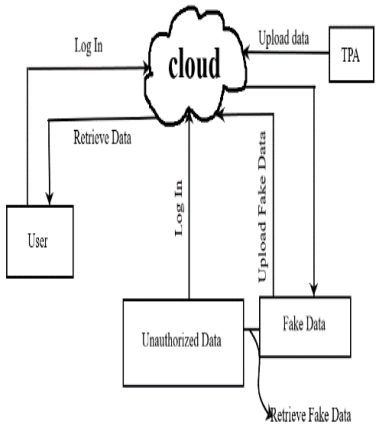

Figure 1. Architecture

PROPOSED SYSTEM

Our goal is to find out when distributor sensitive data is leaked to agents, and if possible to identify the leaked agent.Perturbation is a very useful way in which data is processed and made "slightly sensitive" before being handed over to agents. . we develop unobtrusive strategies for detecting leaks in a collection of items or records. In this section we create a “case” test model for agents. We also introduce distribution algorithms to agents, in a way that enhances our chances of identifying a leaky person. Finally, we also consider the option of adding "fake" items to a distributed set. Such items are not compatible with real organizations but appear to be real to employees.In a way, counterfeit items act as a kind of watermark for the whole set, without changing any individual members. If it turns out that the agent has been given one or more forged leaks, then the distributor may be more certain that the agent was guilty.

METHODALOGY

The new user registers itself and its data is then uploaded to local server with administrative permissions given by admin. The admin simultaneously governs to put the pre-defined fake data into the local server for security purpose. Now if any data is sent to any recipient, the fake data which was clubbed or attached to the original data with fake key will be sent along.Suppose an access breaches into the communication channel, then if that person is having the right key to decipher the data (which is obviously the recipient) will be allowed to have access to original data. But if the user is unauthorized then due to wrong key match the data being given to the user would be the fake one. The admin knows what the original and fake data were. The person having the most of the fake data would be the one responsible for leaking the data.It has the following modules

MODULE 1: DATA ALLOCATION MODULE

This Model focuses on safe and clever way of allocation of data by distributor to its stakeholders to dart out the guilty employees.Logged in user edits and updates its data and save it into cloud with administrative permissions. The desired data is sent to the user by admin. The authentication data are communicated between agents and user via mails (e.g: user login details, etc.).

MODULE 2: FAKE OBJECT MODULE

The data distribution body will add falsehoods to real data, to get a leak point.Fake objects are the objects which looks similar as of the real data being sent but are not the real object records.A fake object puts the data leaker in delusion that he got the original record.As soon as any record is downloaded with the wrong key, it will display the fake record and also sends the mail. The wrong key here refers to that duplicate key which being

displayed as the main key is known as Public Key.It secures the private key which is the actual main key. [5] proposed a secure hash message authentication code. A secure hash message authentication code to avoid certificate revocation list checking is proposed for vehicular ad hoc networks (VANETs). [6] discussed that the activity related status data will be communicated consistently and shared among drivers through VANETs keeping in mind the end goal to enhance driving security and solace. Along these lines, Vehicular specially appointed systems (VANETs) require safeguarding and secure information correspondences.

MODULE 3: OPTIMIZATION MODULE

This module is purely a mandatory module for the data distributor. The distributor has two objectives to achieve. These are: (a) providing the user requested data and making customizable requests for users. (B) Detecting data leaks and identifying the responsible agent.

MODULE4: DATA DISTRIBUTION MODULE

Data distributor takes up the job of sending some sensitive data to some of his entrusting stakeholders or employees. He must ensure that any piece of this data must not be leaked.But supposedly, if it is leaked, admin will come to know which file has been leaked.Also for circling out the guilty employee distributor must assess and examine by considering that the data is leaked by its own one or more agents. The data which is communicated can be of any type and size.

REFERENCES:

[1] V.Shobana,M.E,M.Shanmugasundaram,Assistant Professor Department Of CSE ,Velammal EngineeringCollege /AnnaUniversity,Chennai,India Email:vshobana88@gmail.com,mshans@gmail.com

[2] SandilyaPemmaraju,V.Sushma&Dr.K.V.Daya. Sagar, K L University, India

[3] Chandu Vaidya, Prashant Khobragade, Ashish Golghate --Assistant Professor, Department of Computer Science & Engineering Rajiv Gandhi College of Engineering and Research, Nagpur

[4]Abhijeet Singh, Abhineet Anand Department of Computer Science and Engineering, Galgotias University, Greater Noida, India

[5] Christo Ananth, M.Danya Priyadharshini, "A Secure Hash Message Authentication Code to avoid Certificate Revocation list Checking in Vehicular Adhoc networks", International Journal of Applied Engineering Research (IJAER), Volume 10, Special Issue 2, 2015,(1250-1254).

[6 Christo Ananth, Dr.S. Selvakani, K. Vasumathi, "An Efficient Privacy Preservation in Vehicular Communications Using EC-Based Chameleon Hashing", Journal of Advanced Research in Dynamical and Control Systems, 15-Special Issue, December 2017,pp: 787-792.