A REVIEW ON FRAUD DETECTION IN CLOUD USING DATA MINING

Pranav Rangwala[1],Mrs.Vibha A. Patel[2]
[1,2]Department of Computer Engineering
Chhotubhai Gopalbhai Patel Institute of Technology, Bardoli
Uka Tarsadia University
E-mail: pranav.rangwala@gmail.com

Abstract— Cloud computing, undoubtedly, has become the buzzword in the IT industry today. Looking at the potential impact it's on varied business applications further as in our daily life. Cloud computing is an emerging computing paradigm in which the resources such as storage, processors and software applications are provided as services, remotely over the Internet. The software applications in the cloud may be targeted on different platforms. In this report, there is main focus on one of cloud services "Software as a Service (SAAS)". In SaaS model, the shopper needs to depend upon the service provider for correct security measures. The provider must ensure that multiple users don't get to see each other's data. The main approach of this report is to detect fraud in the cloud using data mining. The whole process related to security and its handling is organized by SaaS Provider. Data security and access control are the most challenging research work going on, at present, in cloud computing .For access control; Pattern Mining is useful when user pattern will be changed. Data mining is the process of extracting useful patterns or knowledge from large databases. Though, data mining also poses a risk to privacy and information protection if not done or used properly.

Keywords- cloud computing, Software-as-a-Service, pattern, cloud service, data mining.

I. INTRODUCTION

The computational world is becoming very large and complex. Cloud computing is emerged as a popular computing model to processing large volumetric data using clusters of commodity computers.

Cloud computing appears to be a highly disruptive technology, which gaining momentum. The concept of cloud computing addresses the next evolutionary step of distributed computing. The goal of this model is to make better use of distributed resources. It allows for the most cost-effective development of scalable web portals on highly available and fail-safe infrastructures

Cloud computing enables the sharing of applications and other system services. Cloud computing allows users to have access to resources, software and information using any networked client devices such as desktop computers, laptops, tablets, and Smartphone. cloud computing provides a means for acquiring computing services without any need for deep understanding of the underlying technology being used by the client [3].

According to Gartner, cloud computing can be defined as ''a style of computing, where massively scalable IT- enabled capabilities are delivered 'as a service' to external customers using Internet technologies.

Cloud computing refers to the many different types of services and applications being delivered in the internet cloud, and the fact that, in many cases, the devices used to access these services and applications do not require any special applications.

1.1. Service model of cloud computing [1]:

There are three types of cloud services infrastructure as a Service, platform as a Service, Software as a Service. In which SaaS is king of all the services.
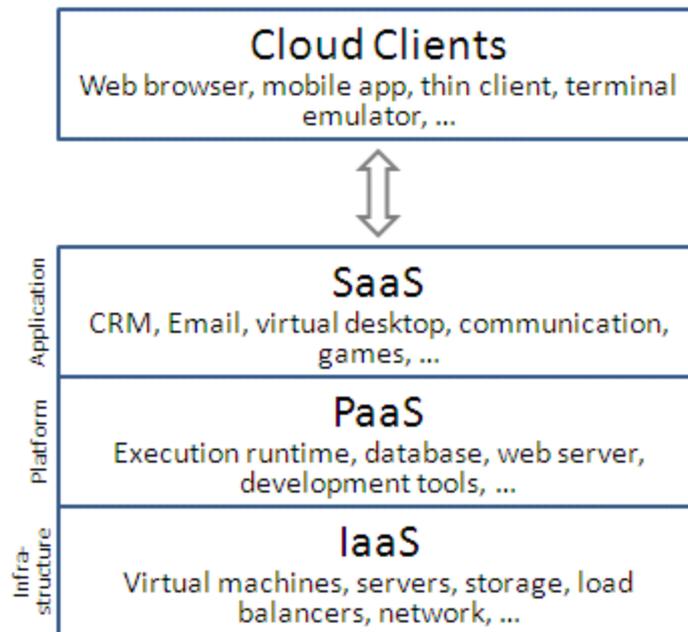
```
┌─────────────────────────────────────────┐
│            Cloud Clients                 │
│  Web browser, mobile app, thin client,   │
│        terminal emulator, ...            │
└─────────────────────────────────────────┘
                    ⇕

         ┌──────────────────────────────────┐
Application│           SaaS                  │
         │  CRM, Email, virtual desktop,     │
         │    communication, games, ...      │
         ├──────────────────────────────────┤
Platform │           PaaS                    │
         │  Execution runtime, database, web │
         │   server, development tools, ...   │
         ├──────────────────────────────────┤
Infra-   │           IaaS                    │
structure│  Virtual machines, servers,       │
         │  storage, load balancers,         │
         │   network, ...                    │
         └──────────────────────────────────┘
```

Figure 1.Layers of Cloud Computing [1]

IaaS:

- •Delivers computer infrastructure as an effectiveness service, generally in a virtualized domain.
- •Provides enormous potential for extensibility and scale.

PaaS:

- •Delivers a platform or solution stack on a cloud infrastructure.
- •Sits on a top of the IaaS architecture and integrates with development and middleware capabilities as well as queuing functions, messaging and databases.

SaaS:

- •Delivers the application over the Internet or Intranet via a cloud Infrastructure.
- •Built on underlying IaaS and PaaS Layer.

II.   DATA MINING

Data mining, in large databases is extracted for predictive hidden information, as it becomes powerful technology with great potential for the companies, to focus on important information from their data warehouses. Data mining tools predict future trends and behaviors, permitting businesses to knowledge-driven selections. Data Mining involves the method of discovering meaning patterns and relationships which will be hidden with massive one are several databases.

So there are many applications of Data mining in real world As, Hospital, Airline Reservation, Biometrics, Mathematics, Geographical, Forecasting, Student Management, Web Mining, Parallel Processing, Space Organization, Data Integrity, etc. So there are many applications in which the data mining term is very useful. So from many application we will be discussing on cloud computing, how data mining is used in cloud computing.

2.1. Data mining parameters

- **Association** - Looking for patterns where one event is associated with another event. It finds associated rules for frequently co-occurring items, used for market basket analysis.
- **Sequence or path analysis -** Looking for patterns where one event leads to another later event.
- **Classification -** Looking for new patterns, Most common techniques to predicting a Specific outcomes
- **Clustering -** Finding and visually documenting groups of facts not previously known

- **Forecasting** – Also known as predictive analysis. Discovering patterns in data that can lead to reasonable predictions about the future.

III. PATTERN ARCHITECHTURE

Software as a Service (SaaS) is a software distribution model in which applications are hosted by a vendor or service provider and made available to customers over a network, typically the Internet. SaaS sometimes referred to as "on-demand software" by ISVs or Application-Service-Providers, in cloud it's a software delivery model and software and associated data are centrally hosted on the cloud [3]. Web-browser acting as a thin-client for accessing the software remotely across the internet. Application delivery as a one-to-many model says single instance and multi-tenant architecture than to a one-to-one model.

SaaS is often referred to as On-demand-software and utilizing it is close to renting software rather than buying it. Not like traditional software applications that you would purchase the software and then install it onto your computer. The license for software's may also limited users and/or devices where the software can be deployed. Software as a Service users subscribe to the software usually on a monthly basis.

User might have to use software applications that don't required local installation and software maintenance. How software can be delivered to use over internet? [5]

SaaS applications delivered to use as a service generally through the internet via browsers or APIs. User's access application storage and computations are hosted in cloud without any local installations.
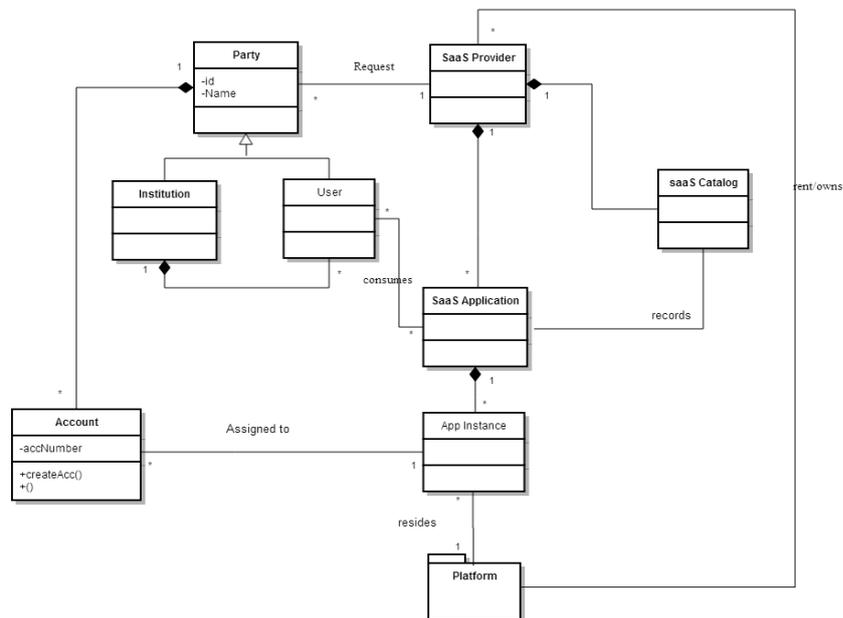

Figure 2. Class Diagram for SaaS Pattern [3]

In the above figure, there is a pattern shows how SaaS actually work. In above class diagram when user send request to SaaS provider first it check the user authentication. In SaaS catalog, it contains all the details of SaaS applications available in cloud and also contains the user details like what packages is purchased and its usage report. It also take a data of users whether it single user or an institute.

There is a main draw of the SaaS services is security. To define that user is authorized or not. Application used by the client is an authorized client or not. A key aspect of any service is to give the security and confidentiality to client.

Our design approach is to make a system which gives all security benefits to user by using efficient data mining techniques.

IV. PROPOSED WORK

As discussed above, User can access its applications provided by SaaS provider at any place over the internet. It can provide the benefits like pay for use, accessibility, transparency, on demand services, flexibility, simplifications and elasticity.

3

Propose work to introduce data mining technique for Software as a Service (SaaS) to detect fraud in cloud. Data mining classification algorithm will be used to detect the fraud. By using this approach we can prevent fraud in cloud. A SaaS Provider processes requests from Parties. A Party can be either a User or a group of users (Institution). A Party can have one or more Accounts. SaaS Catalog contains the list of the SaaS Applications that are offered to the users. There can be a single Application Instance of the application that is shared by different users, or single instance per user. Platform - The SaaS Provider offers a set of Software Applications that resides on a Platform. SaaS applications can be deployed to user using underlying Platform -as- a-Service or Infrastructure-as-a-Service offerings. The platform can be owned or rented to a third-party provider. The user log manager manages all the process of authentication and authorization. In this, manager can manage all data related to client like user Id, password, personal details, etc. At user log manager data mining algorithm will used to check user is authorized or not. When user enters user id and password at SaaS Providers page. Provider will send this data to User Log Manager. At User Log Manager, data mining algorithm will check that user is authorized or not. Log Manager will create a dataset of a fraud user's IP. This dataset can be store in data ware house for future detection.

At server side provider store the user's usage pattern, like user usage time and how long it consume the applications, mostly consuming time, frequent pattern of user to accessing the application.

This pattern mining approach in SaaS is very useful for client as well as providers for the security aspects. Another use of the architecture is to be a reference for security certification of services [7]. Knowing the misuse patterns that affect a particular service, a provider can show that his service can handle the corresponding threats by incorporating appropriate security patterns, which would increase customer trust in their use. Patterns provide a way to evaluate the security of complete systems by finding a matching security pattern to defend against each threat [7]. We will apply them to evaluate the security of cloud systems.

V. CONCLUSION

Pattern for Software-as-a-Service model with data mining algorithm has been done in the report. In Software as a Service model describes on-demand applications that are deployed on the cloud and accessed by users through the internet. The data mining in Cloud Computing allows organizations to centralize the management of software and data storage, with assurance of reliable and efficient secure services for their users. Using pattern mining approach SaaS can achieve the next level of security. Currently in cloud computing, research work has been carried out on data security and access control. Further pattern mining algorithm will be developed for efficiently mining user access pattern.

REFERENCES

[1] Bhagyashree Ambulkar and Vaishali Borkar, "Data Mining in Cloud Computing", Proceedings published by International Journal of Computer Applications (IJCA), April 2012.

[2] Naskar Ankita and Mrs. Mishra Monika R., "Using cloud computing to provide Data mining services", International Journal Of Engineering And Computer Science Volume 2 Issue 3, pp.545-550, March 2013.

[3] Hashizume, Eduardo B. Fernandez, Maria M. Larrondo-Petrie, " A pattern for Software-as-a-Service in Clouds", IEEE International Conference on Bio Medical Computing, 2012.

[4] Rashmi, Dr.G.Sahoo, Dr.S.Mehfuz, "Securing Software as a service model of a Cloud Computing : Issues and Security", International Journal on Cloud Computing: Services and Architecture (IJCCSA) ,Vol.3, No.4, August 2013.

[5] Panagiotis Kalagiakos and Panagiotis Karampelas, "Cloud Computing Learning", IEEE, 2011.

[6] T.R. Gopalakrishnan Nair, K.Lakshmi Madhuri, "data mining using hierarchical virtual k-means approach integrating data fragments in cloud computing environment", ieee, 2011.

[7] Ms. Sneha Sontakke, Dr. A. D. Gawande and Ms.Parag k. shelke, "Intrusion detection system for cloud computing", International Journal of Scientific & Technology Research Volume 1 Issue 4,pp.67-71, May 2012.