Loss Prevention and Fraud Detection System

A Thesis

submitted in partial fulfillment of the requirements for the award of the degree of

Master of Engineering

in

Computer Science and Engineering Department

by

Lalit Sharma

(801632022)

Under the supervision of

Dr. Prashant Singh Rana

Assistant Professor, CSED

THAPAR INSTITUTE OF ENGINEERING AND TECHNOLOGY

PATIALA, PUNJAB - 147004

July 2018

Certificate

I hereby certify that the work, which is being presented in the thesis, entitled "***Loss Prevention and Fraud Detection System***", in partial fulfillment of the requirements for the award of the degree of **Master Of Engineering** in Computer Science and Engineering submitted in Computer Science and Engineering Department of **Thapar Institute of Engineering and Technology**, Patiala is an authentic record of my own work carried out under the supervision of ***DR. Prashant Singh Rana***. I have also cited the reference about the text(s)/Figure(s)/Table(s) from where they have been taken.

The matter presented in this thesis has not been submitted elsewhere for the award of any other degree or diploma from any institution.

Signature: Lalit Sharma

Lalit Sharma

This is to certify that the above statement made by the candidate is correct to the best of our knowledge.

Signature: 

Dr Prashant Singh Rana

Assistant Professor

CSED

Acknowledgement

This research work would not be possible without acknowledgement of the people who supported and guided me for the successful completion of this task.

First of all, it's my pleasure to express thanks and credit to **Mr. Anil Kumar Pillai**, Senior Manager, Oracle RGBU and **Mr. Atul Gupta**, Senior Software Developer, Oracle RGBU, Bangalore, India for their valuable guidance and mentorship that helped me to overcome every challenge I faced as I moved on in this work.

I want to thanks to **Dr Prashant Singh Rana**, Assistant Professor, Computer Science and Engineering Department, Thapar Institute of Engineering and technology, Patiala for recognition my work and to give his valuable time and guidance and also to his continuous encouragement throughout research work. It won't be possible without his intellectual guidance which triggered my knowledge in right direction and in equal proportion.

I am grateful to **Dr. Maninder Singh**, Hon'ble Head of Computer Science and Engineering Department, Thapar Institute of Engineering and technology, Patiala for all the help and to maintain a healthy environment which also plays a major role in research work.

I am also thankful to the Institution and all faculty members of Computer Engineering Department, Thapar Institute of Engineering and technology, Patiala for providing right direction and their motivation.

Lalit Sharma

801632022

Abstract

The project aims at providing a customer friendly solution for POS (point of service) application. This integration helps accumulate data from POS application at different locations under one centralized cloud system. This way the service providers can target different categories of loss and frauds that can emerge in retails. It also helps in providing business analysis. In this project, integration consists of two major modules i.e. loss prevention and fraud detection application with any POS (point of service) or online customer retail application. This integration helps the two applications to communicate where one application uses the data from the other .The UI framework has been designed using the OJET (Oracle Java Extension Toolkit). This helps in displaying the generated reports which provide data analysis. This data analysis helps in the decision making process and to increase the profits. The business analysis module is used to give an insight into the data pattern and then formulate policies and decisions accordingly.

Table of Content

List of Figures

Chapter 1

Introduction

There are a number of POS applications present at various stores which provide a huge amount of customer data that needs to be handled. Also, there are different categories of customers which leads to return and refunds for the products which the purchased. Returns and refunds also handled at POS application but which return and refund is honest that we need to consolidate and Figure out. Hence, here the Loss prevention and Fraud detection come into the picture. Loss Prevention and Fraud Detection System is the most widely used for loss prevention and point of service (POS) data analysis tool. It mainly contains two different modules:

1.1 Modules of Loss Prevention and Fraud Detection System

1.1.1 Loss Prevention and Asset Protection

The Loss Prevention module and Asset Protection leads to generate reports and dashboard dynamically for detecting patterns and anomalies. It leads to use huge bulk of data from various POS. It also allows customer to generate reports based on given data in the real time system. It leads to fetch data and keep it in data warehouse from which it fetches data periodically by the lag of 15 minutes.

1.1.2 Sales and Productivity

Sales and Productivity module offers robust and highly configurable reporting across all levels of the retail organization hierarchy (Salesperson, Store, District, Region, and so on), merchandise hierarchy (item, class, dept., and so on), and/or by geographic attributes. Through a comprehensive set of grid and graph reports, documents and interactive dashboards, users can compare same store sales to past performance and custom goals,

measure sales members' productivity, and evaluate the impact of merchandise characteristics on productivity.

Customers can purchase the Loss Prevention module separately, or have both modules bundled in the application. It shows that both the modules are independent and can be separately used.

Figure 1.1 Fetching of Data from POS

So application how fetch data from various POS is referenced in Figure 1.1. As customer purchasing data from store all transaction stores into POS database. Data stored into POS database used by Loss Prevention and Sales and Productivity system for generating various reports and dashboards. Based on generated reports and dashboards it leads to detect fraud and losses.

1.2 Data Flow of Application

Loss prevention module fetches data from various POS and then leads to generate various reports. So basically how to fetch data from POS will be referenced in Figure 1.2

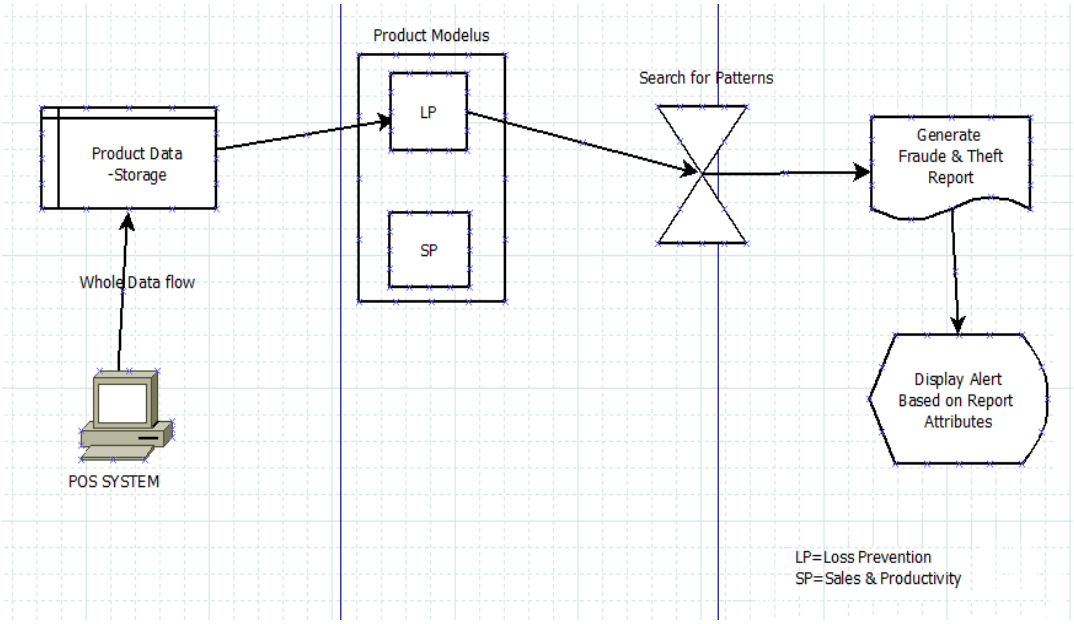

Figure 1.2 Loss Prevention Data Flow

We have various modules in the system, data from POS also goes for sales and productivity modules which is referenced in Figure 1.3

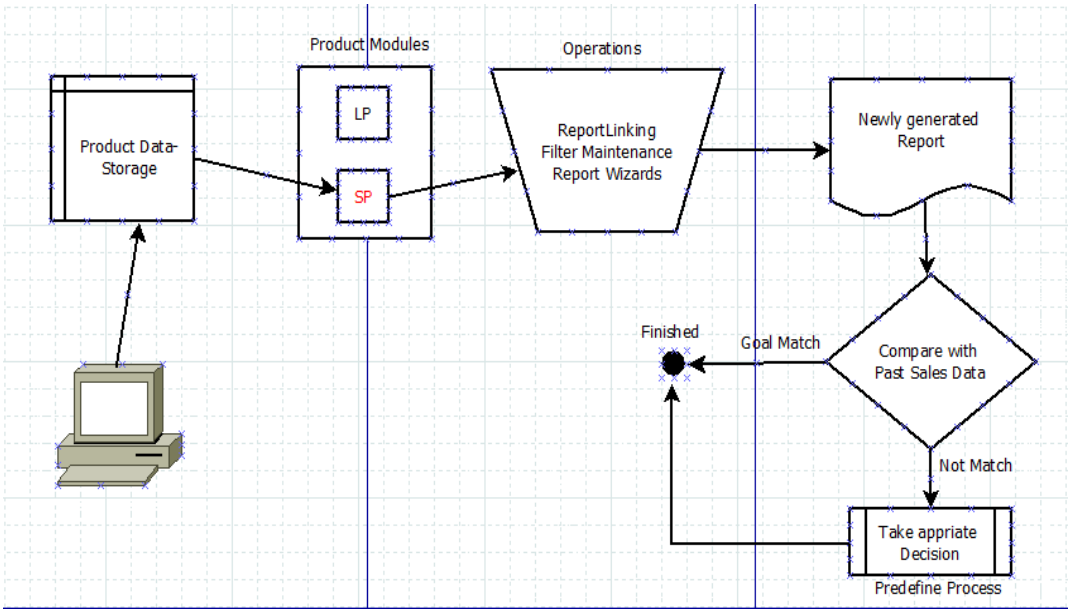

Figure 1.3 Sales and Productivity Data Flow

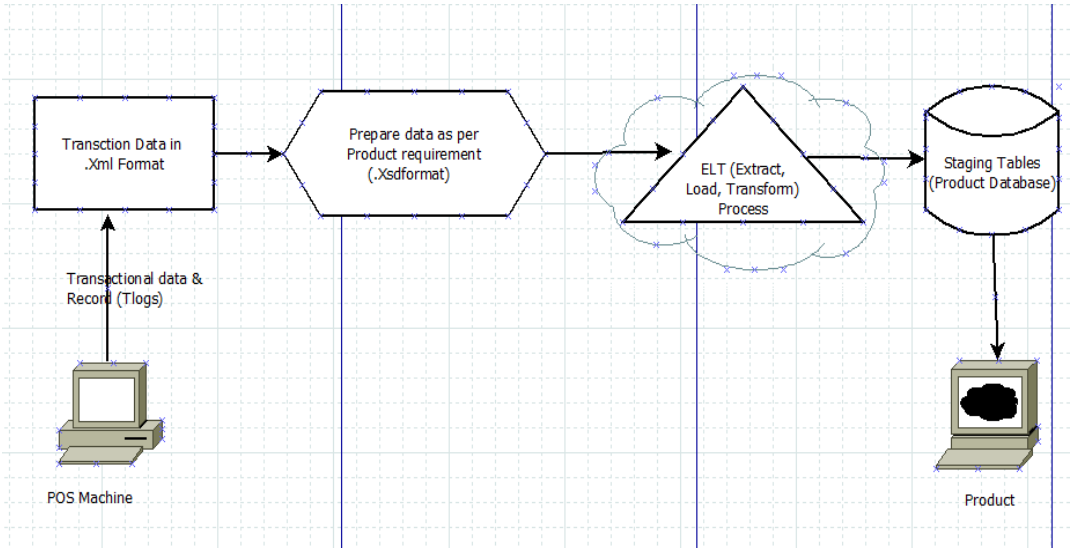

Figure 1.4 POS to Application Data Flow

Data works for the system is referenced in Figure 1.4. It shows Application data flow.

1.3 Problem Description

From past 20 years, in retail industry 14% losses are only incurred due to fraud and losses. Customer purchasing items from store and online sites and they returned it by formatting some causes and reason. User try to perform some malpractice As a total of 92000+ scenarios are there that needed be tested regressively thus time plays a crucial role in losses. Manual Testing was difficult for getting desired results. Hence, loss and fraud detection leads to generate dynamically reports and detect losses and to increase sales.

1.4 Objectives

Here are some objectives of loss prevention and fraud detection and also sales and productivity module:

- To shrink merchandise shrinkage.
- To ensure effective use of CCTV camera, Alarms and other security instruments.

- To increase the sales and productivity by creating report/dashboard from previous data.
- To generate reports/dashboards for detecting losses and fraud and to increase sales.
- Reduce losses% in fraud and losses in retail industries.
- To provide security against theft exceptions in retail store.
- Securely data transfer between two parties.

1.5 Thesis Organization

- **Chapter 1:** This chapter introduces loss prevention and fraud detection system and all the basic modules of it. It also covers flow of data from POS to fraud detection system. It also covers problem definition, main objectives and also thesis organization.
- **Chapter 2:** Describes the fundamentals required for this thesis. It explain all the surveys which required as literature survey for retail industries. It also explain authorization traditional approach which used earlier.
- **Chapter 3:** This chapter defines architecture and implementation. It explain about the architecture, flow and also working of the system. It also explain about ETL data flow and how to proceed with it.
- **Chapter 4:** This chapter explains about all the features and functionality of the product on the deeper side. It covers all the features of loss prevention module and all the features of sales and productivity module.
- **Chapter 5:** In this chapter the results of thesis are discussed and also the main reasons behind the not so good quality of applications are listed.
- **Chapter 6:** This chapter states the conclusion and the scope for future work.

Chapter 2

Literature Review

2.1 Survey of Retail Industries Losses

Now a day most of the transactions are happening through electronic payment, like credit card, debit card, e-wallet etc. Main focus of any retailer is to increase the profit in their business through increasing sales of item/goods. There is other side to increase the profit by reducing loophole in existing system [17]. As per my Survey retail industries facing losses into internal theft, Shoplifting, fraud in the system, paperwork error etc. [1]
Below chart show the theft and fraud in Retail industries.

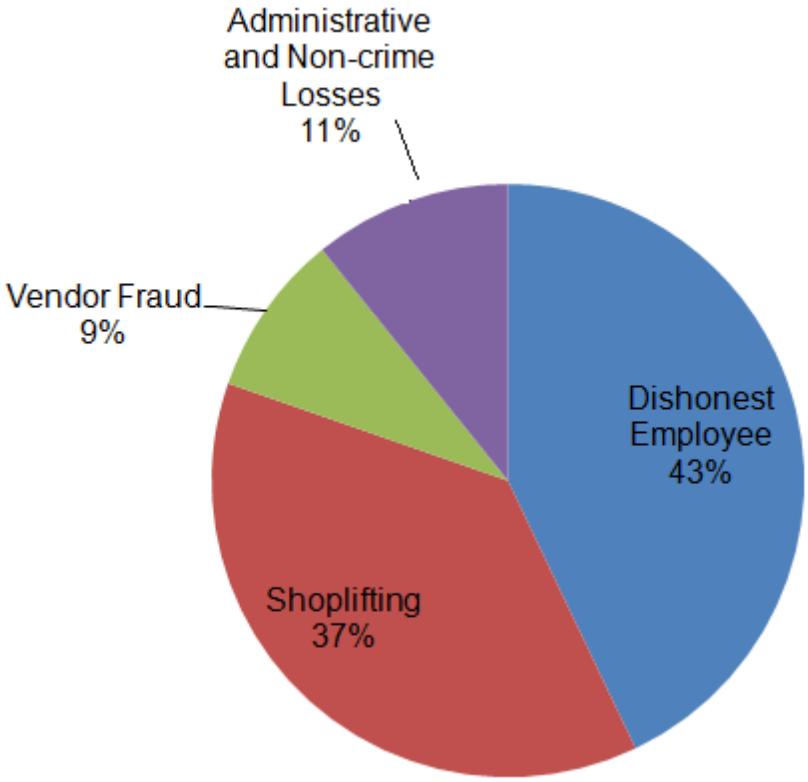

Figure 2.1 Chart Diagram of Retail Losses

Theft and fraud in retail industries losses in last year referenced in Figure 2.1. To handle this much amount of data, system should have edibility and scalability. Cloud give that functionality so this project will design in such a way so it can move to cloud and serve the need of customers dynamically [18].

2.1.1 Traditional Authorization method

Before JSON web Token introduce in the market, Industries uses tradition session approach to identify the user. In session approach server needs to maintain session for each and every user and store into its memory or Database system.

Steps to authorize the user using Session/cookies approach [2]:

1. User request for the resources by sending Username and Password
2. Server verify the login information.
3. Server create session and store it into DB
4. Server return Session Id (SID) to particular client and store into client cookies
5. Client request next Resource with SID
6. Server check Session ID is valid or not (by searching into DB)
7. Check the user role for authorization of requested resources
8. Finally User receive requested Resource.

Authentication and authorization for setting up token to get secure access is described in Figure 2.2

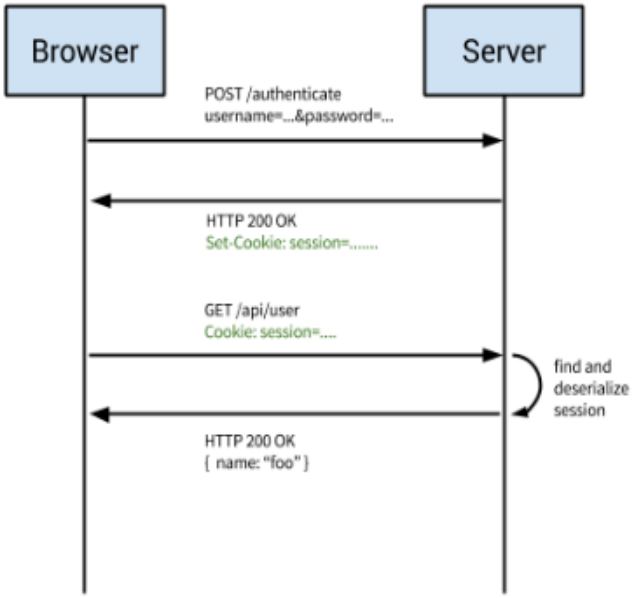

Figure 2.2 Cookies/Session based Authentication/Authorization

Chapter 3

Architecture and Workflow

This project follows a layered architecture and the user requests going from layer to layer. The architecture given below depicts the working of the projects in brief and how the data travel from user system to the database and how the reports get generated and communicated back to the desired results.

3.1 Implementation Architecture/Workflow Diagram

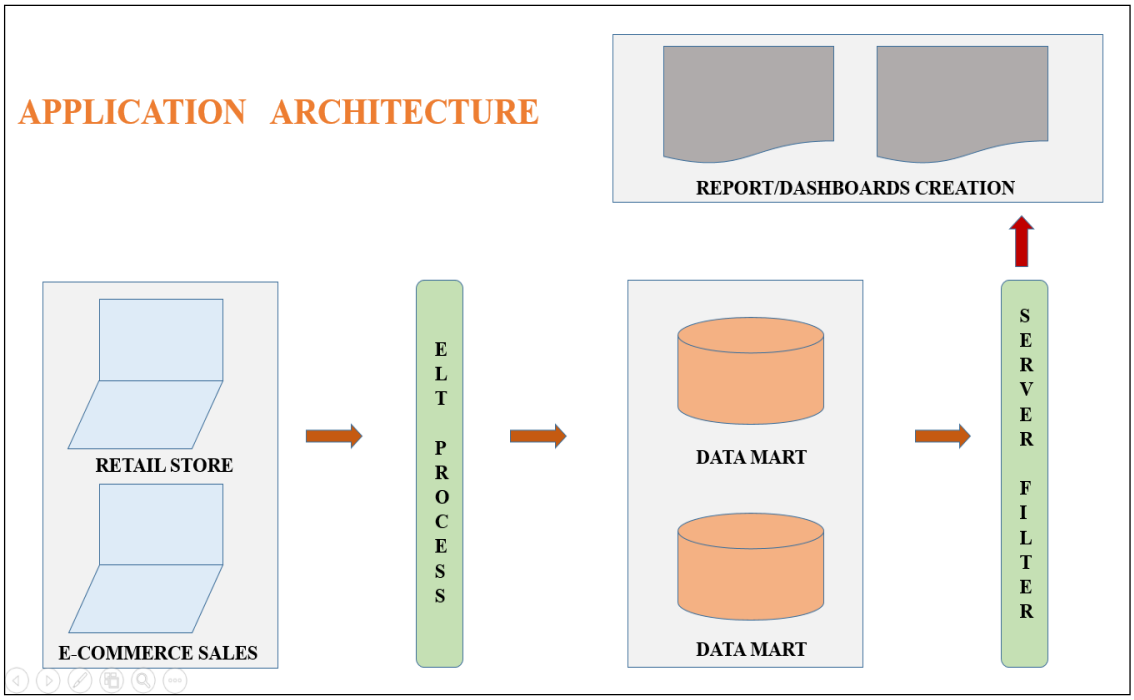

Figure 3.1 Application Architecture of Loss Prevention System

The architecture is used in the project is referenced in Figure 3.1 and depicts the request and response from the user end [10]. Also, it depicts the layered architecture as discussed above.

3.2 Application Workflow

The following points outline how the application data retrieval is done in the project and shows the workflow of the application [6].

1. All the customer transaction starts from bottom to up to convert in valuable report.
2. On the application architecture, the customer transaction generated at Any Retails Store and E-commerce Inventory Sales.
3. All customer related transaction are then transferred to the local database of our system through via path of cloud storage and with the process of ELT [7].
4. Transaction then stored into data mart and all related data (data about data) will store into Application Meta Data.
5. Then, all the data will transferred to Intelligence Server which leads to take inputs from Data Mart and Application Meta Data [8].
6. Now, all structure getting stored into Schema Object and all values stored into Metrics and then Intelligence Server apply some kind of Filters on data so that valuable information get collected.
7. Filtered information then leads to generate reports/dashboards on the basis of given previous data. Dashboard also implemented here as it is nothing, it just collection of various reports.
8. Then, generated reports and dashboards are available to display at various devices i.e. Mobile and Web [9].

3.3 ETL Dataflow

ETL is a method for data compression. It leads to extract data form various source of data into valuable format and then leads to transform and compress it. Transforming of data leads to store values into metrics and compress values (keep only those values which are valuable) [19]. ETL performed well only when the high-performance data engine used in

end system, like a cloud installation, data appliance and Hadoop cluster ELT Data flow take major 4 steps for completion which are referenced in Figure 3.2:

- In first step, it will extract all the data from the customer transaction.
- Then, it will lead to deliver data on Cloud through a protocol called SFTP [11].
- Data available on cloud loaded into local database of system.
- Loaded data now transformed to a valuable format.

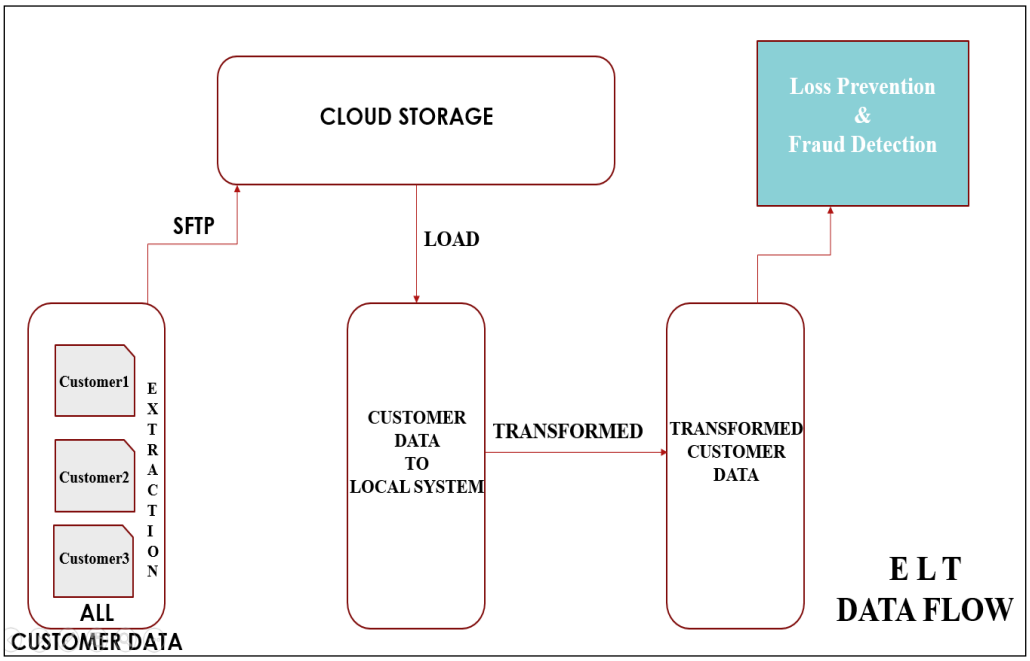

Figure 3.2 ETL Dataflow

Chapter 4

Features and Implementation

It consist of all those modules which describes product functionality for better understanding. Here, some of modules are described below:

4.1 Functionality and Features of the System

Here functionality of each modules is defined:

4.1.1 New Reports and Dashboard

This dashboard was designed to summarize controls and exceptions in order to provide Loss Prevention (LP) leadership with information on what exceptions are occurring, how they are being addressed, and who may be handling them [3]. The dashboard has a selection of date ranges to summarize on exceptions; looking at total generated, break down by status, type, where they are occurring, and who is assigned to investigate them.

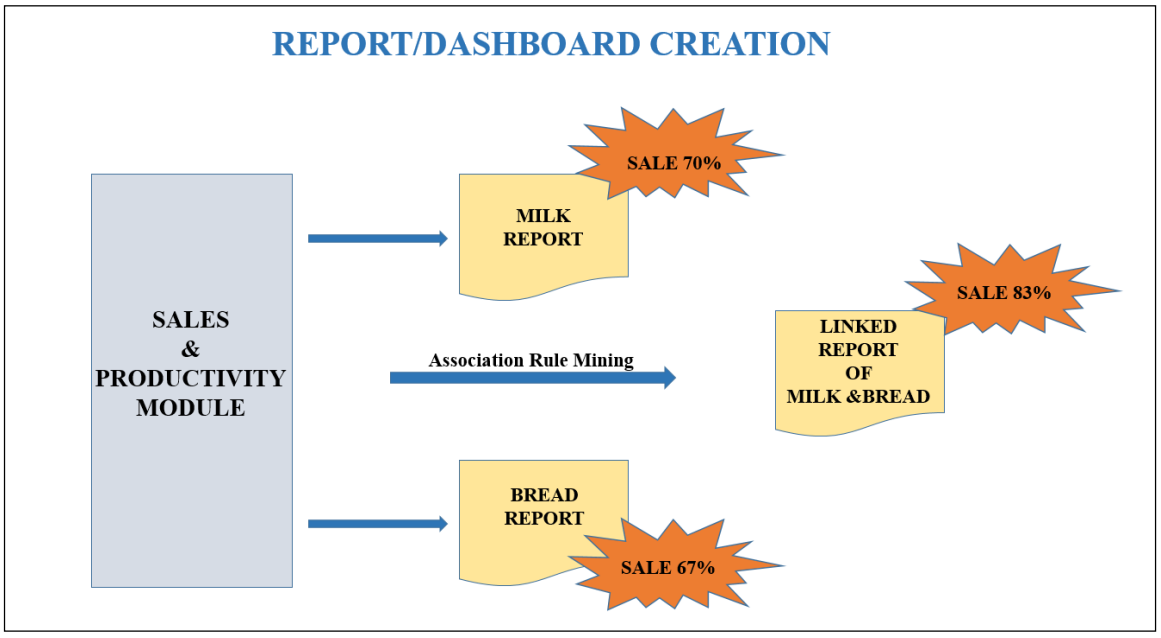

Figure 4.1 Report Dashboard linking diagram

To increase sales and productivity system generated linked reports are explained in Figure 4.1. It shows similarity between multiple products.

4.1.2 New Administration

New Administration is basically meant for different access rights at various user-levels. Here are some of instances of it:

4.1.2.1 Day Parts

This new Project Defaults screen lets users customize the day parts, or defined intra-day time periods (for example, Morning, Lunch, Afternoon, Evening) that are used in the Sales Flow by Period Dashboard and Sales Flow by Period reporting.

4.1.2.2 Discount Details

This report now includes the ability to display items. The discounts were not being associated with the item to which the discount was applied or against which it was prorated.

4.1.3 ETL Production

These all are generated by data driven from the process of ETL Production

4.1.3.1 Administration

Administration block contains various utilities like Add User, Add New Connection and Database Connection Check.

4.1.3.2 ETL Dashboard

ETL Dashboard are nothing but just collection of multiple reports. It leads to link various reports together and based on their linking results it represent them graphically.

4.1.3.3 ETL Report

ETL Report is basically collection of data results. It shows them graphically and also compare with all the previous data values for giving a bar or a standard.

4.1.3.4 ETL Tool

ETL Tool are mainly used for data extraction, data verification and validation and also for data compression. They also used for loading data into various repository.

4.1.4 Real Time Processing

It provides real-time processing to support intraday flash sales reporting for the new Sales and Productivity module. This is accomplished by processing data in specified time increments throughout the day rather than once at the end of day. Additional business logic supports the inclusion of post voids, clock in, clock out, and adjustments to normal business. Additional Loss Prevention analysis followed by no sale and no match transactions will be updated through the end of day process. It all explained in Figure 4.2.

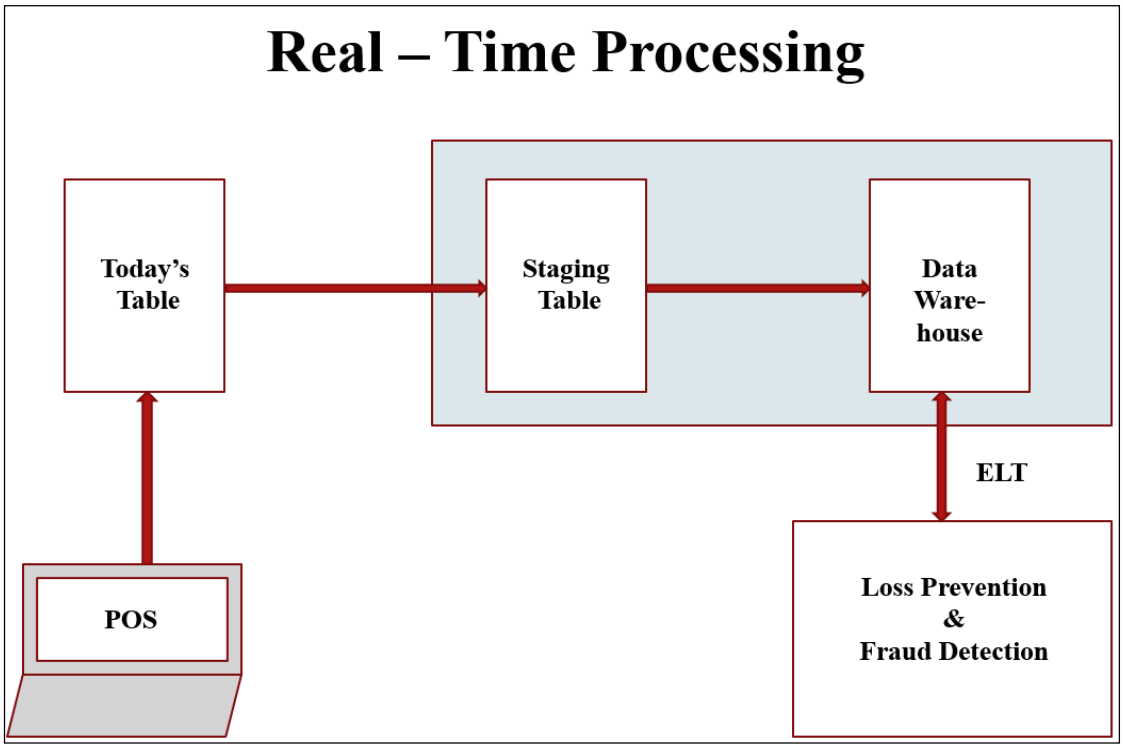

Figure 4.2 Report Dashboard linking diagram

4.1.5 Reporting

Reporting provide the facilities to create different types of report based on previous data. It's also vary from module to module [20]. Some of the Report types are mention below:

Cash Refund Summery Report

1. Discounts and Price Overrides Summary Report
2. Refund and Exchange Summery Report
3. Transaction Discount Report
4. Employee Discount Report
5. Coupon Transaction Report

4.1.6 Look and Feel to Visual Insight (VI) Module

There is a new look-and-feel that modernizes the appearance of Visual Insights to provide an improved user experience [4].

4.1.6.1 Dynamic Links

User can create dynamic links to external content in a Grid visualization.

4.1.6.2 Graph Value Option

User can now specify values outside the minimum and maximum value range to maintain a consistent scale over time, even if the data is republished and the minimum and maximum values change.

4.1.6.3 Attribute Thresholds

User can add thresholds to attributes in visualizations, allowing you to apply conditional formatting across headers. When applying a threshold to a parent attribute, the conditional formatting is applied at level of the visualization.

4.2 Features of Loss Prevention Module

Here all the features of product described here:

4.2.1 Video Linking

The Video Linking module provides lessons that show you how to link to a video clip of activity filmed during a specific POS transaction and how to archive video links, manage IP addresses for camera locations, and upload video _les to the database System.

The video link feature allows you to retrieve the digital video that corresponds to one or more transactions. You can execute a video link from any header/detail level report, or document that includes the Trans attribute. From the Video Queue page, you can modify the camera, start time, and end time before launching the video. After viewing the video clip, you can send the link to the Video Archive, where you can view the video again, make changes if needed, and upload it to the database for later retrieval.

4.2.2 Exception Control

You can add thresholds to attributes in visualizations, allowing you to apply conditional formatting across headers. When applying a threshold to a parent attribute, the conditional formatting is applied at level of the visualization.

A user's data security filter must be incorporated into the application of the alert filter. The security filter should be applied first before the alert filter or any report filter. For example, if a user is restricted to one store and an alert filter is used showing the top five cashiers for revenue, the report will include the top five cashiers in the store to which the user has access. Security filters restricting a user to their organization in a multi-tenant environment should be applied first as well.

4.2.3 Watch Status

Use the Watch List functionality to assign a watch status to stores or cashiers displaying questionable activities.

The Watch List detail screen contains the complete history of a specific cashier or store on the watch list as well as the tools used to change the status or add information regarding the status.

Cashier Watch List information includes:

1. Cashier or Store information
2. History: watch status history appears in chronological order, with the most recent information at the top of the list. Information within the History section cannot be changed or deleted.
3. Watch Status selection menu
4. Watch Notes: In the Watch Note field to enter comments regarding this watch status, the cashier, or any other information that may help track the item.

4.3 Features of Sales Productivity Module

4.3.1 Report/Dashboard Linking

Ability to link from one report to multiple reports that assist the end users with analysis or investigating the data. The linking functionality allows for reports to be link to the same report or different, allowing focus on a key field or set of fields while allowing expansion of the date range. The reports also need to link to and from dashboards that may contain one or multiple datasets. Dashboards also need to be able to link from one to other dashboards.

4.3.2 User Security Administration – Functionality Access

User Roles are defined to allow different levels of access application for functionality depending on the need of the users. In addition, the extended features can be controlled through the administrative interface.

4.3.3 Lookup Conversion

Ability to have lookup Tables that convert raw data to display descriptions that are more meaningful to the end users. The ability to have summarization of metrics at 10 different levels for analysis or comparison based on master Table hierarchies. The end user may want to compare the employee to the store, district and/or regional level as an average.

4.3.4 Report Wizards

The reports wizards are designed for the end user to build new reports with a limited access to what fields and metrics can be put on the report based on the Tables and joins. This provides a step by step process for the end user to reduce possible confusion.

This report wizards will designed to summarize controls and exceptions in order to provide leadership with information on what exceptions are occurring, how they are being addressed, and who may be handling them. The dashboard has a selection of date ranges to summarize on exceptions; looking at total generated, break down by status, type, where they are occurring, and who is assigned to investigate them [5].

4.4 Implementation of various modules

This application is running in distributed environment. Basically each system is connected to different server. It is very difficult to access the remote system. In this system, while executing report or run any other functionality of application it might get some error, exception or generate the log for the particular module.

So Loss prevention and fraud detection system provide the one unique functionality called log maintenance to access the different system's log remotely.

4.5 Log Maintenance

It is a functionality to maintain, view and download the logs/error of the different machine where the application is running. It will provide the remote access of the log file/table to the admin user only. In this application logs are store in two ways:

1. **File Logs**: File logs contains logs information in readable mode with the dot (.) log extension.
2. **Table Logs**: Table logs contain logs information in the tabular manner like: time of log, Log description, Log location etc. it will show the log size in number of rows in the table.

Application will deploy on Cloud so different instance of the application will created according to the number of user. It is very difficult to go (login) into each machine and show the log/error, so this functionality give User interface to the admin user to access the logs/errors of different machine in a single screen.

4.5.1 Log Maintenance Functionality

To control over the system's log, it provides major four functionality:

1. Add log References: Add reference to the log not actual log file. log References: Edit added references
2. Delete log References: Delete the log reference, actual log file is still there.
3. View/Download: It shows the actual log file contents base on Added reference. It also allow user to download the actual log file

Figure 4.3 can give the more realistic view of the functionality.

Example: If we want to access machines A logs (where the application is running.)

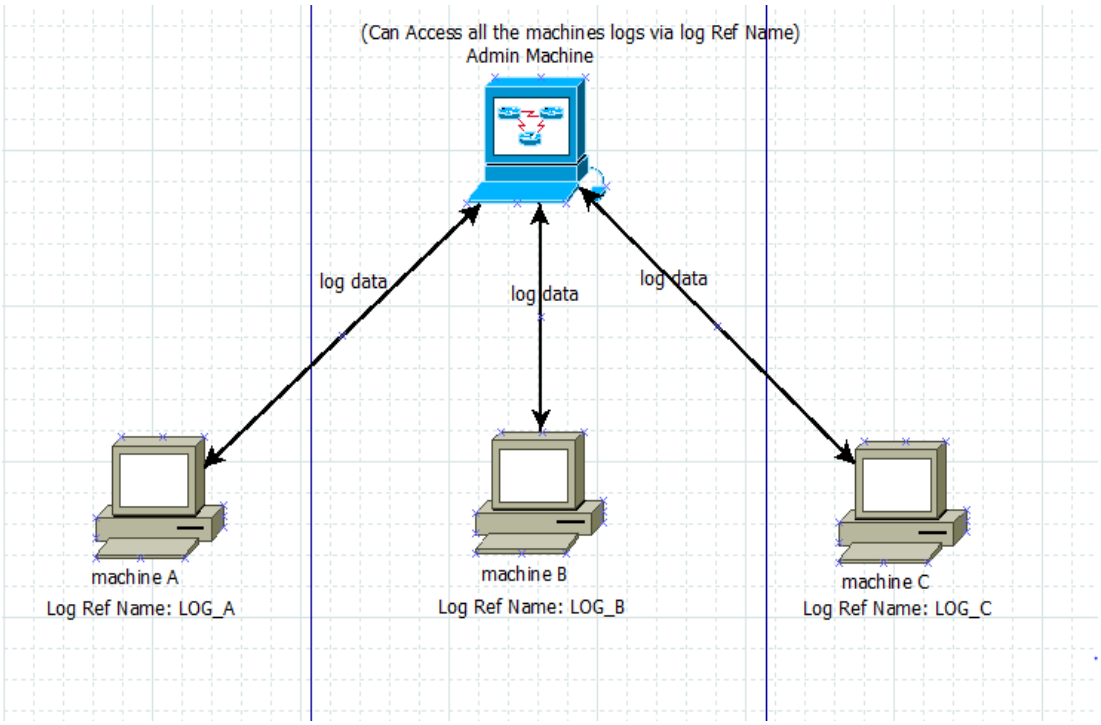

Figure 4.3 Log Maintenance Data Flow

1. First we need to add the one reference to the machines A log _le (If machine A store its log in location C:/log folder and we add the Reference to that location as LOG-A.
2. By using the log view functionality we can access that log by just giving log reference (LOG-A). Now it will show list of log which is in C:/log folder.
3. Same as add, we can edit the log reference from LOG-A to some other name.
4. While we delete the log reference (LOG-A) it just remove the reference which is pointing to C:/log folder but not an actual _le, actual _le is still in C:/log folder.

To handle this much amount of data, system should have edibility and scalability. Cloud give that functionality so this project will design in such a way so it can move to cloud and serve the need of customers dynamically.

4.5.2 Viewing the active system

It will help user to view the connected system to the particular server. We can filter out the connected system by using three parameters. You can refer to Figure 4.4 for viewing remote system.

1. Country
2. State
3. City

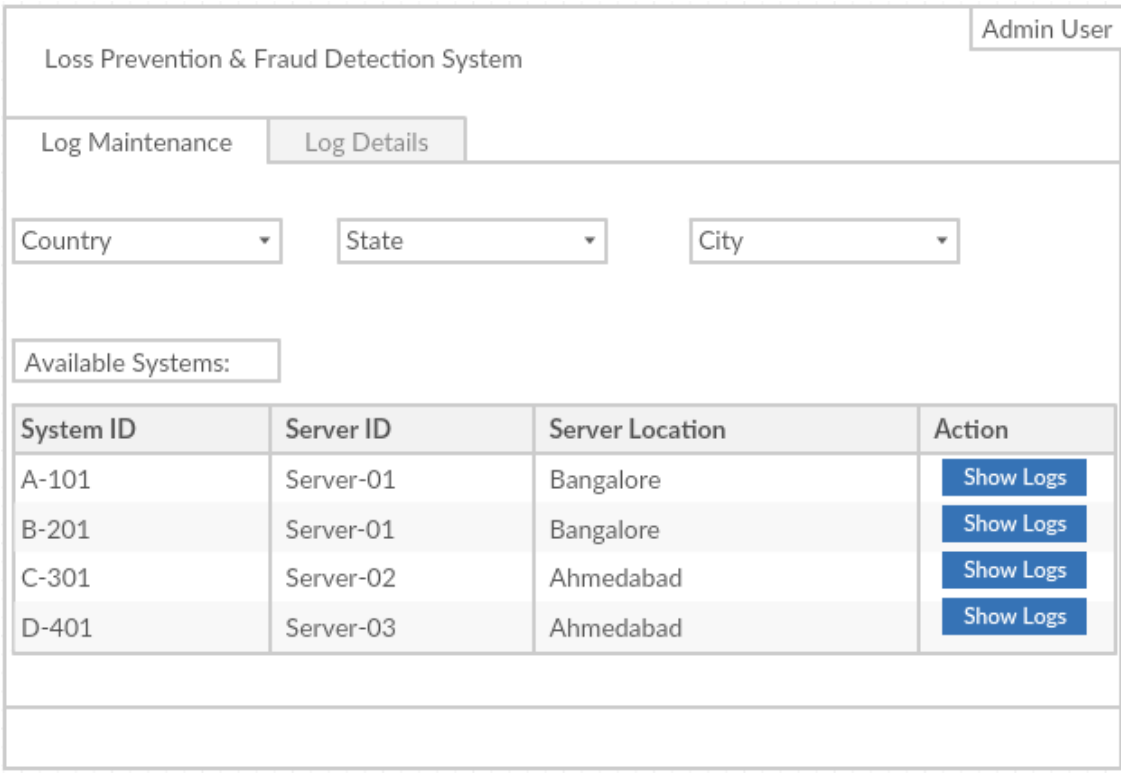

Figure 4.4 Viewing Remote System

After selecting appropriate location filter, it will show System ID, Connected server to that system (Server ID) and physical Location of server. It will contain show log action which will help to display logs of different server.

4.5.3 View/Download System logs

This functionality will provide access to the remote machine's log. it is very easy for the user to view the log of different machine (Server) by using show log action from the given panel. It will display log detail of selected system according to the date filter.

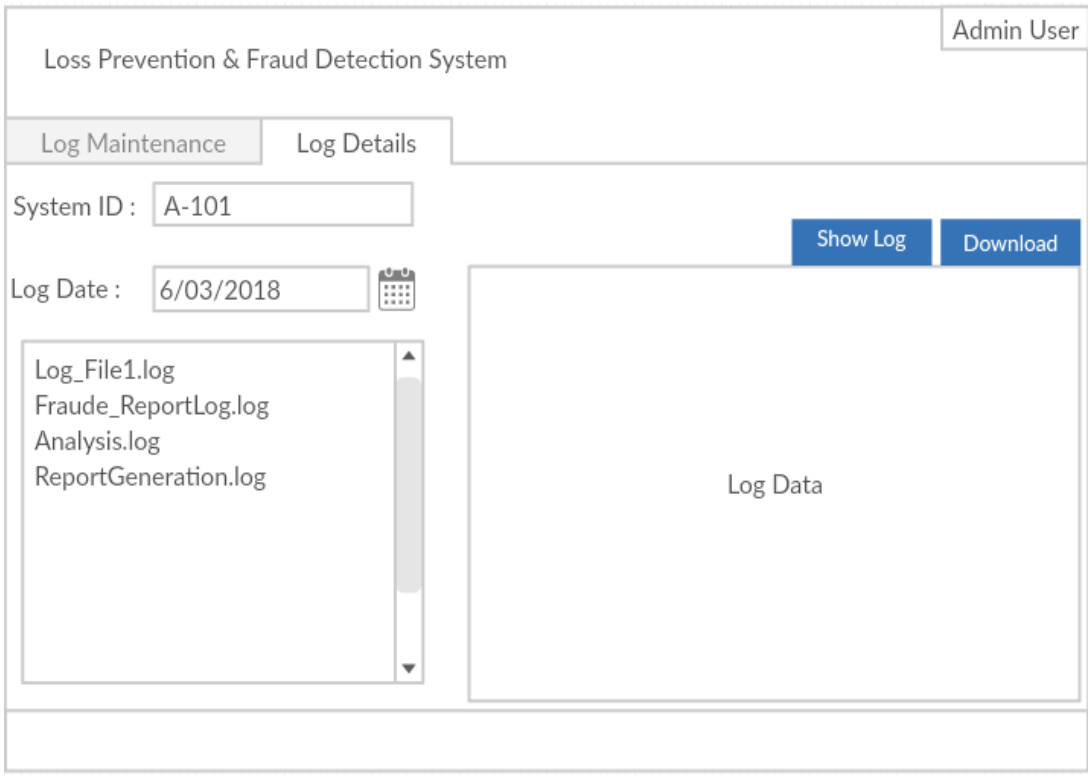

Figure 4.5 View/Download system's logs

This screen has major two functionality

1. Show Log: It will show the log within the application UI
2. Download Log: It will download the log to local system

While executing show/download action, request will send to server and it will search log according to the given filter (Date) and display that log in view area. We can also download the file log and table log by using the download action.

4.5.4 Log Configuration

This functionality is available for admin user only to create/add or Edit log reference to specific system. By providing System ID, Server ID, and Log location, User can add log reference which will reacted to viewing remote system screen.

Log Configuration Screen:

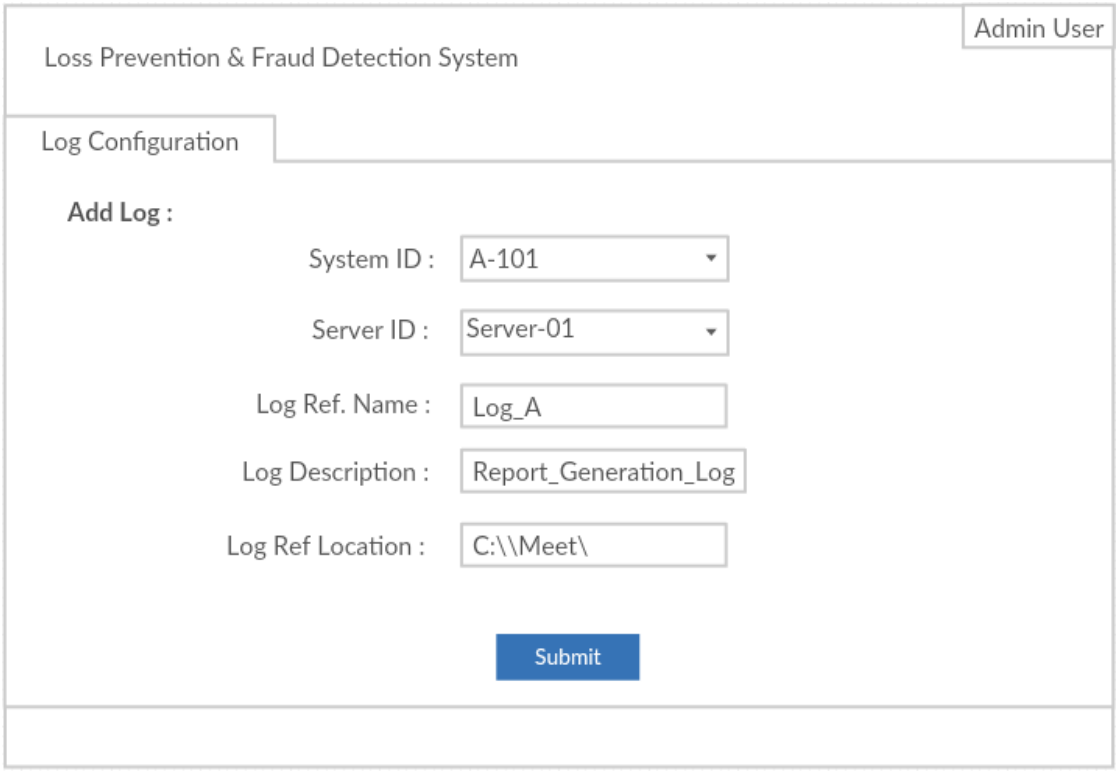

Figure 4.6 Add/Edit Log Configuration

Here user just add the log reference to particular server not the actual log file which is described in Figure 4.6. Log file will generated by server its self-according to the error or exception scenario.

4.6 Application Auditing

Application provide the interface between Users and sensitive data therefor application proper control on information flow are on top priority. To protect those data from user or keep eye on their action we require a frame work in such a way so it can monitor all required action from user. Purpose of the data auditing is used to capture the event or operation on

the application data, it involves profiling of user data. Some data of the application are very sensitive and restricted to the admin users only. It is very useful in future to track the modification and access of sensitive data [12]. Auditing in application can be done via two ways:

1. Store the data in flat file
2. Store the data in database table

In this application audit data are store in database rather than flat file. There are several parameters needs to be define in database to store the auditing data. Parameters are mention below:

1. User id: Id of user who is accessing the data
2. User role: role of user (ex: Admin, Super User)
3. Data and Time of auditing: Exact date and time of accessing data in predefined format
4. Performed action: It store the user action on while accessing particular data, action can be Create, Read, Update (modification of data) or Delete.
5. Severity of action: It indicate severity of user action. it will be like high, low, medium
6. Actual audit data: this data will be in JSON format which contain all the accessed data. Like in case of adding new record in system it will track of that record along with user details.

Below is the pictorial representation of flow and different parameters of audit data.

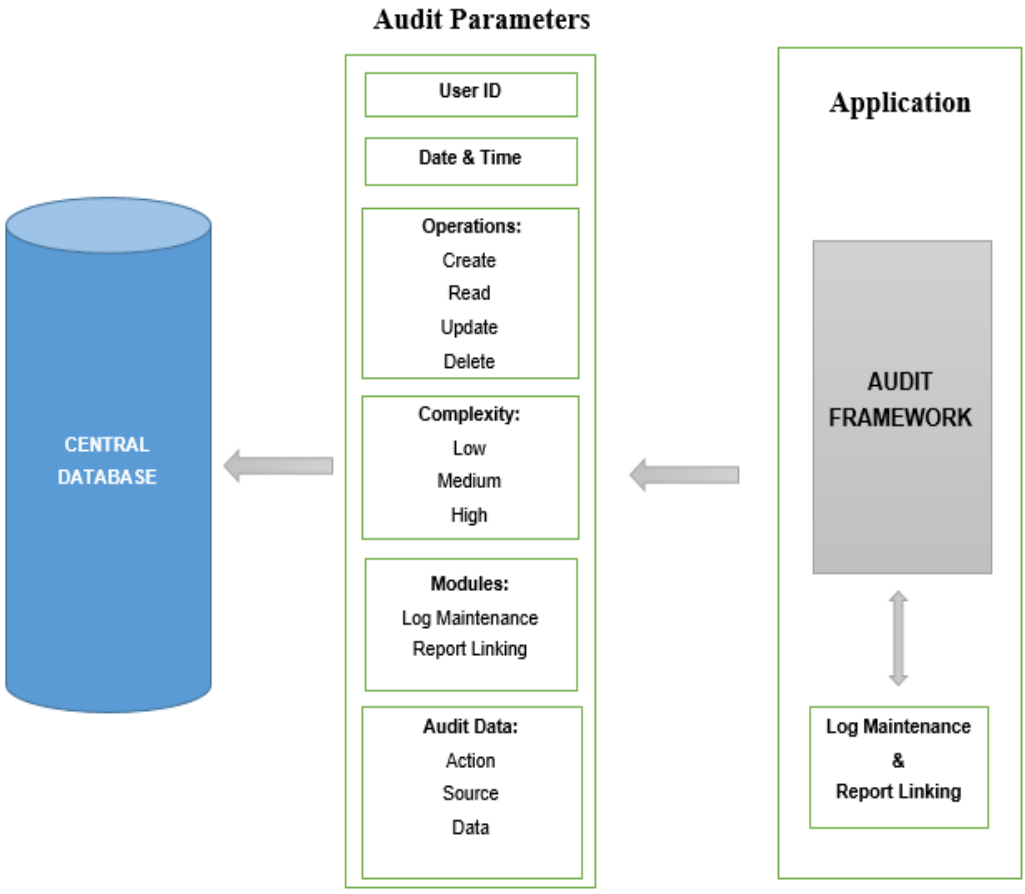

Figure 4.7 Audit Parameters

4.6.1 Data Auditing

In loss prevention and fraud detection system, we have different module to access the various functionality of application like report linking, Log maintenance, filter maintenance etc. Here I designed a frame work in such a way so any module can use this audit framework to make auditing happen in various module. Here is the explanation of how audit framework implemented in log maintenance features. There are several request will be initiate while serving output or result of requested module [13].

1. Choose one module and execute from System UI
2. Request will initiate for accessed module
3. Set required parameter to initiated request and REST call will execute to the server

4. Before serving output to System UI, It will call the audit framework to store the audit data in database
5. After successful auditing of data, output will displayed to User for a requested module

Flow diagram of audit data framework:

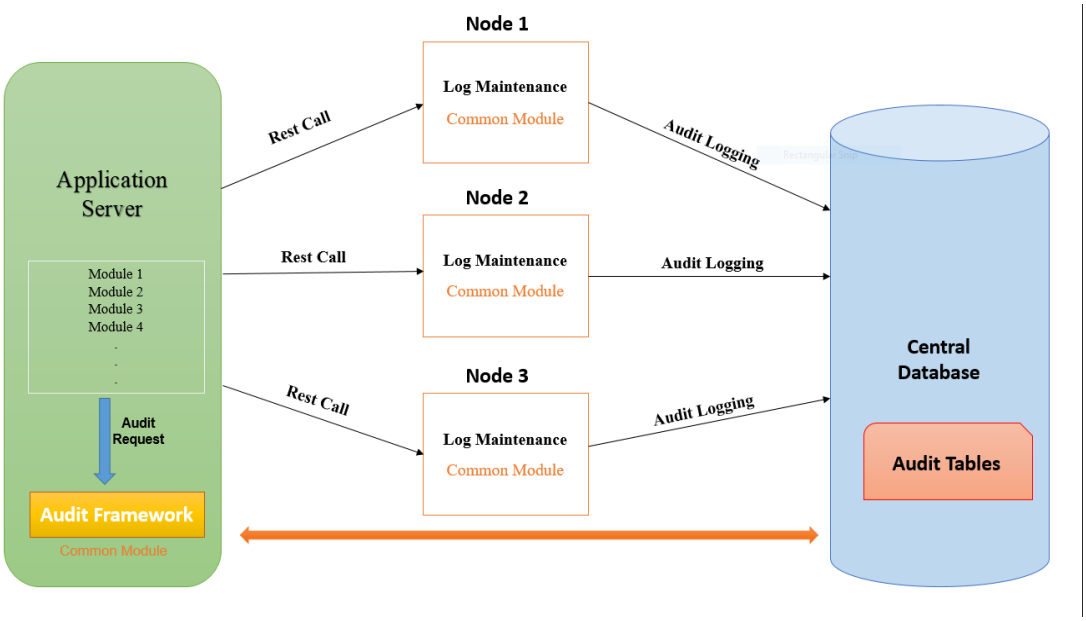

Figure 4.8 Audit Data Flow

4.6.2 Audit Database

Here the screen shots for LOGVIEWER-AUDIT-TABLE:

As here all the columns which are present in table are defined with their data type, with null able property (whether it will be null or not), Default data that will be present, COLUMN-ID assigned to each column of table and their respective comments [14].

LOGVIEWER_AUDIT_LOG

Columns | Data | Constraints | Grants | Statistics | Triggers | Flashback | Dependencies | Details | Partitions | Indexes | SQL

Actions...

	COLUMN_NAME	DATA_TYPE	NULLABLE	DATA_DEFAULT	COLUMN_ID	COMMENTS
1	USER_ID	VARCHAR2(128 BYTE)	No	(null)	1	(null)
2	LOG_DATE	DATE	No	(null)	2	(null)
3	LOG_TIME	VARCHAR2(20 BYTE)	No	(null)	3	(null)
4	CATEGORIES	VARCHAR2(20 BYTE)	No	(null)	4	(null)
5	SEVERITY	VARCHAR2(20 BYTE)	No	(null)	5	(null)
6	MODULE	VARCHAR2(20 BYTE)	No	(null)	6	(null)
7	LOG_DATA	CLOB	Yes	(null)	7	(null)

Figure 4.9 Audit Table

Properties are defined for various attributes of audit table due to certain assumptions/reasons [15].

1. USER-ID: Its data type should be VARCHAR2 of 128 bytes as the name of person can be too long. It cannot be null as it must contains user name on audit.
2. LOG-DATE: Log date data type must be of DATE type which defined for storing dates in DATABASE. It also cannot be null as it take system date.
3. LOG-TIME: It is stored in data type that is VARCHAR2. It takes system time and hence it must be in fixed format which leads to store in fixed size that is less than to 20bytes.
4. CATEGORIES/OPERATIONS: It defines CRUD operations which are fixed in length and of string type. Hence, VARCHAR2 of 20 bytes is suitable data type for it. It also cannot be null.
5. SEVERITY/COMPLEXITY: It tell about complexity whether it can be HIGH,MEDIUM,LOW and it will also be of string type with length than 20 bytes, hence VARCHAR2 of 20 bytes is suitable for it.
6. MODULE: It tell about user accessing the module. None module in application have name length more than 20 bytes.
7. LOG-DATA/AUDIT-DATA: It stores JSON object as its data. Hence it leads to store as in the format of CLOB. In this field data can be null if user just trying to access module but not trying to access sensitive data.

Column id are by default generated at the time of adding fields to table.

User who can access database can put comment for description against each column of table [16].

Screenshot displays how the data of an audit log stored into database:

Audit log 9th entry contains some user name, 19-APR-18 as date, 04:32:30 as time, read as operation, medium as complexity, log viewer as module name and JSON of filename, log name, logotype etc. as Audit data.

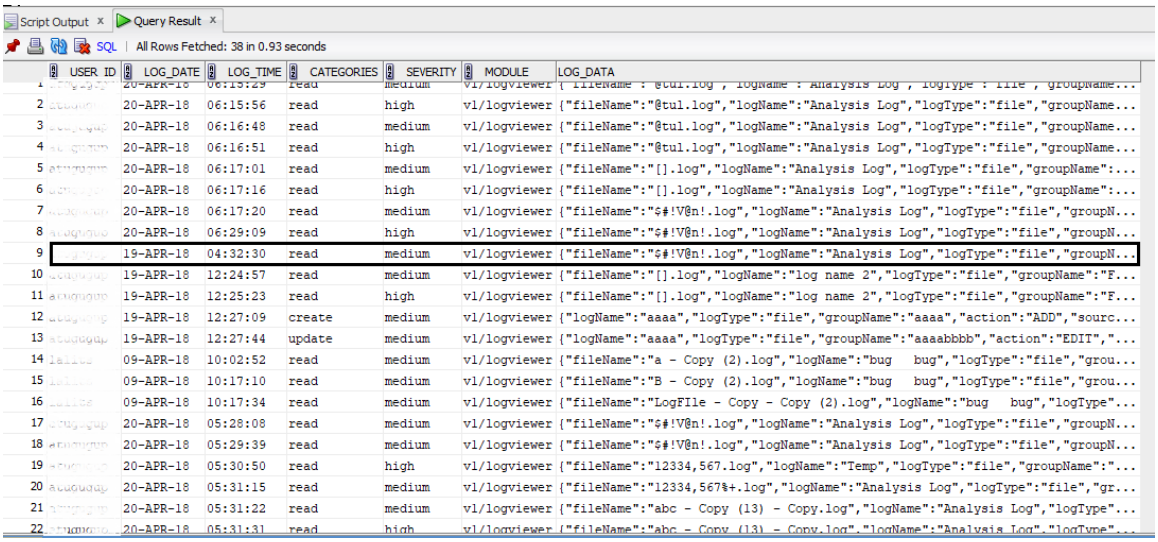

Figure 4.10 Audit Logs

4.7 Application Security Scan

Loss prevention and fraud detection system is a web application (web Project). It is frequently transferred data over the network.

Path manipulation errors occur when the following two conditions are met:

1. An attacker can specify a path used in an operation on the file system.
2. By specifying the resource, the attacker gains a capability that would not otherwise be permitted.

For example, the program may give the attacker the ability to overwrite the specified file or run with a configuration controlled by the attacker. In this case, the attacker can specify

the value that enters the program at particular field in and this value is used to access a file-system resource.

There are the several well-known vulnerability which could be found at a time of security scanning. Like:

1. Resource data-flow Vulnerability
2. Path manipulation Vulnerability
3. Application information leak vulnerability
4. Unreleased resource vulnerability

4.7.1 Fortify Scan

It is a tool for checking security vulnerability in the application. It prioritizes vulnerability by severity and importance. It helps to find the root cause of security issues. It will pinpoint to exact location/ line where attacker can expose the system. It has a capability to analyze the static codes in different language and also support various IDEs integration (ex: Eclipse, Net Beans). We can also build our custom rule for security scanning of application.

It categories security issues by its severity. It may be high, medium, low or by severity number (1, 2, and 3). It will highlight primary location, Issue name, Analysis type etc. on dashboard. Below is the dashboard of fortify tool:

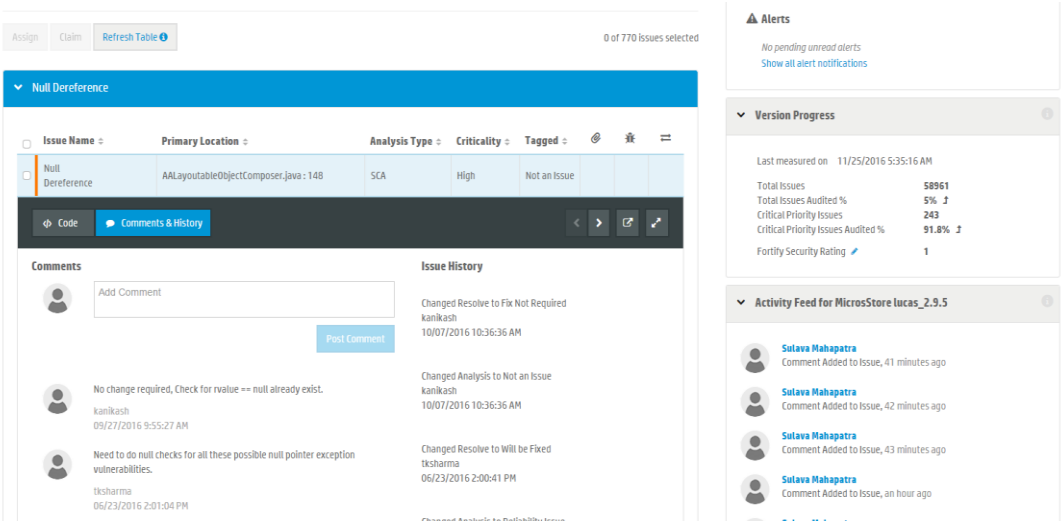

Figure 4.11 Fortify Dashboard

As standard by the industries, If scanning results have severity issues which is belong to high or medium category then application cannot be release to production environment but severity low category issues are _ne for them and it will gradually evolve as application enhance. Most of the type low category issues occur due to bad coding practice.

4.7.2 Analyze and Prevent Vulnerability

The best way to prevent path manipulation is with a level of indirection. Create a list of legitimate resource names that a user is allowed to specify, and only allow the user to select from the list. With this approach the input provided by the user is never used directly to specify the resource name. In some situations this approach is impractical because the set of legitimate resource names is too large or too hard to keep track of. Programmers often resort to blacklisting in these situations. Blacklisting selectively rejects or escapes potentially dangerous characters before using the input. However, any such list of unsafe characters is likely to be incomplete and will almost certainly become out of date. A better approach is to create a whitelist of characters that are allowed to appear in the resource name and accept input composed exclusively of characters in the approved set.

There are some way to Prevent Vulnerability:

1. If the program is performing input validation, satisfy yourself that the validation is correct, and use the HPE Security Fortify Custom Rules Editor to create a cleanse rule for the validation routine.
2. Implementation of an effective blacklist is notoriously difficult. One should be skeptical if validation logic requires blacklisting. Consider different types of input encoding and different sets of meta-characters that might have special meaning when interpreted by different operating systems, databases, or other resources. Determine whether or not the blacklist can be updated easily, correctly, and completely if these requirements ever change.
3. A number of modern web frameworks provide mechanisms for performing validation of user input. Struts and Spring MVC are among them. To highlight the invalidated sources of input, the HPE Security Fortify Secure Coding Rule packs dynamically re-prioritize the issues reported by HPE Security Fortify Static Code.

Analyzer by lowering their probability of exploit and providing pointers to the supporting evidence whenever the framework validation mechanism is in use.

4.7.3 Benefit of Fortify

1. It will reduce the development cost by finding vulnerability issues in early development phase.
2. It will enable to secure coding for the application.
3. It will help in remediation of security vulnerability
4. It provides correlation and priorities the results

Chapter 5

Result and Analysis

5.1 Loss Prevention and Fraud detection Analysis

Below screenshot illustrate a dummy example that how it enables to detect fraud for various industries on the annual basis. It leads to detecting fraud on some criteria that is extracting from the user given data and then leads to give reports result.

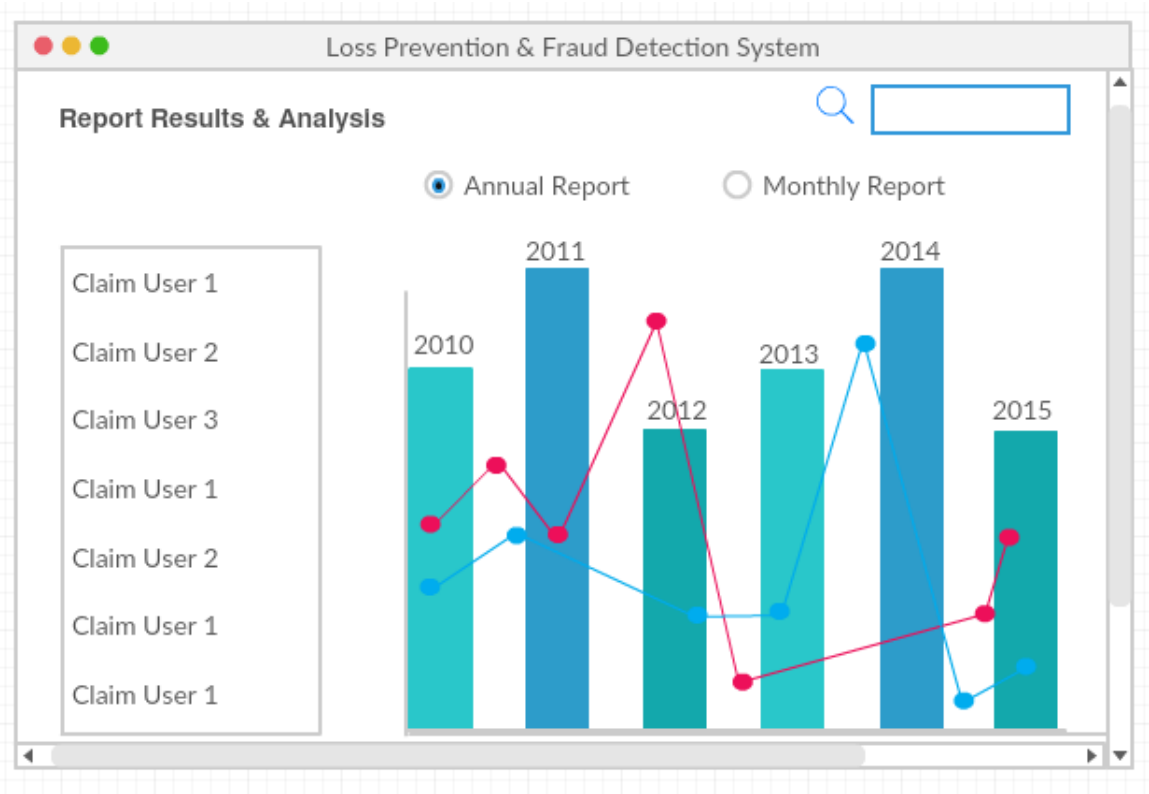

Figure 5.1Fraud detection report analysis

It showing that how the product is efficient in detecting report on the annual basis. User can also generate report on monthly basis. Detecting fraud for various companies need to acquire the whole data warehouse of that company in dynamically and in runtime manner.

Here some points are considered for detecting fraud from screenshots:

5.1.1 User Data

Customer data is must to detect the fraud and losses. Company's data warehouse required to go through the process of ETL (Extract Transform Load) for getting in the valuable format. Extracting data from company's warehouse lag runtime process by 15 minutes.

5.1.2 Report Generation

Report generation can only be done in two manner:

1. Monthly Basis: On monthly basis system can only detect fraud only for those companies which have high sales as high customer transaction is required for generating reports. Sometimes for companies which have lower sales, system get fails to detect fraud on monthly basis.
2. Annual Basis: On annually basis system leads to generate reports for huge data transactions. Hence, it will become very easy for system to detect fraud and losses. Companies which have less sales also get benefit by annual reporting.

5.2 Sales and productivity Analysis

Sales and Productivity analysis is mainly depends on algorithms which system using for finding similarity between different products. Data mining algorithms plays very crucial role to increase sales for any item individually or in association with other items. There are many algorithms which system can use for increasing sales. Here are some name of those algorithms:

1. Apriori Algorithm
2. K-means Algorithm
3. Page Rank Algorithm
4. C4.5 Algorithm

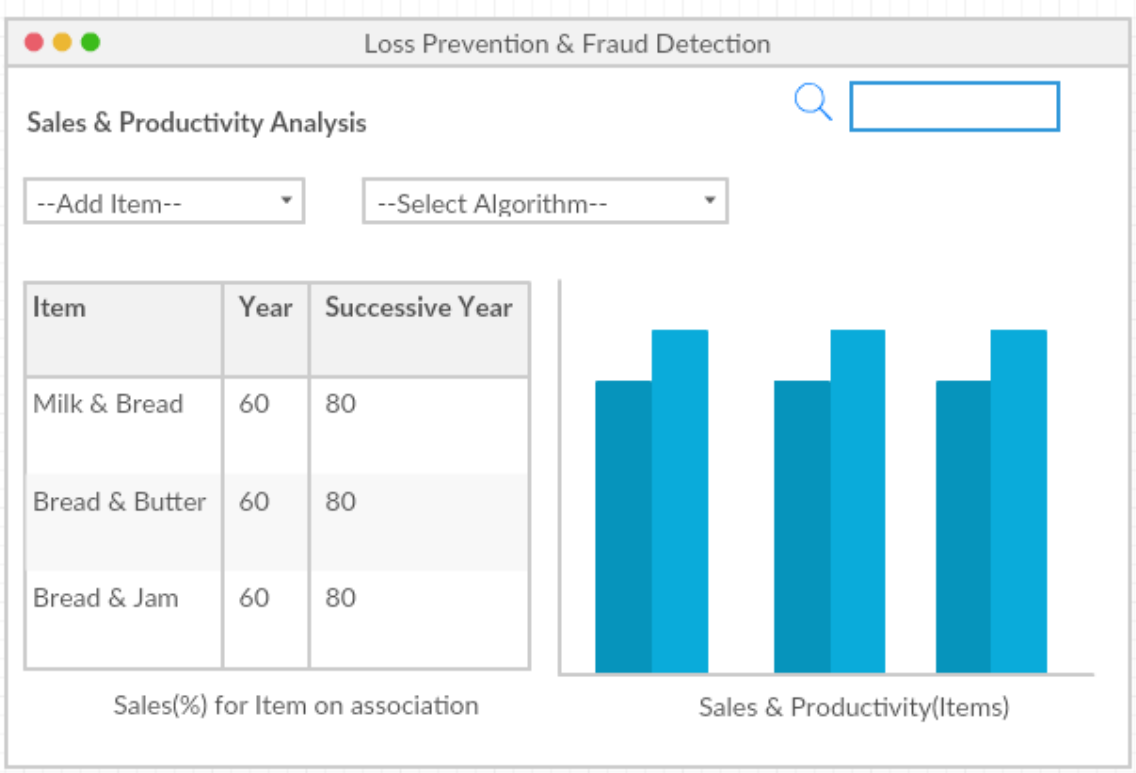

Figure 5.2 Sales and Productivity report analysis

Above Figure illustrate a dummy report analysis of sales and productivity for various associative items. Above report analysis depends on mainly two parameters:

1. Items sales Data: Item sales within all the companies is extracted and then transformed into a valuable format as a pre-process. The valuable format data then leads to loaded into sales and productivity modules for analysis. It will generate report by using any selected algorithms.
2. Similarity Algorithms: There can be many algorithms present in system for finding similarity between various products. Users are also allowed to add their own customized algorithms. In the above diagram association rules are used as algorithms and leads to detect the sales as merged items.

Chapter 6

Conclusion and Future Work

6.1 Conclusion

Since not all the customers returns reasons are not valid and genuine so catching of those customer which are making losses to our company and business get solved through these factors. Keeping track of such customers would be easy with the help of integration of various POS and with all the minute details of the customers. Hence, the customer harming company becomes easier to catch and detect now onwards.

The Sales and productivity modules helps in increases sales and manufacturing of different products based on various parameters and factors. It solve the problem of merchandise shrinkage and also the problem of removing products from the inventory which are not getting sales from long time.

6.2 Future Scope

In future we can improve the process of returning cash back to customer. We can make some reports which follows all the protocols of government on taxation so that returning amount is also include the amount of tax they pay while purchasing the item. Here are some points that can be consider for future scope:

1. Setting up threshold value for returned products.
2. Administration defaults for getting access.
3. Real Time Processing for intraday flash sales.
4. Watch List Implementation for CCTV-Camera.

References

[1] J. Hillmer, R. Jones, C. Gessner, C. Johnston, K. Lewis, and S. Deshpande, “System and method for detecting fraudulent transactions," Mar. 30 2004. *US Patent* 6,714,918.

[2] A. Kukic, “Cookies vs. tokens: The definitive guide," 2016. *US Patent*

[3] J. Shapland, “Preventing retail-sector crimes," *Crime and Justice*, vol. 19, pp. 263-342, 1995.

[4] B. Goodman, “Loss prevention and sales productivity cloud services," 2017. *US Patent*

[5] C. Welch, J. Rozmus, J. Whiteman, M. Negin, and W. Herd, “System and methods for preventing fraud in retail environments, including the detection of empty and non-empty shopping carts," Mar. 16 1999. *US Patent* 5,883,968.

[6] K. Cook, “Method and system for the detection, management and prevention of losses in retail and other environments," Apr. 20 1999. *US Patent* 5,895,453.

[7] M. Haekal and Eliyani, “Token-based authentication using json web token on sikasir restful web service," in *2016 International Conference on Informatics and Computing (ICIC)*, pp. 175{179, Oct 2016.

[8] P. Otemuyiwa, “Json web tokens vs. session cookies: In practice," 2016.

[9] S. Peyrott, “JWT Handbook”.2009

[10] P. Solapurkar, “Building secure healthcare services using oauth 2.0 and json web token in iot cloud scenario," in 2016 *2nd International Conference on Contemporary Computing and Informatics (IC3I)*, pp. 99{104, Dec 2016.

[11] R. Damphousse, “Build secure user interfaces using json web tokens (jwts)," 2015.

[12] L. K. Shar and H. B. K. Tan, “Auditing the xss defence features implemented in web application programs," *IET Software, vol. 6*, pp. 377{390, August 2012.

[13] S. Hiremath and S. Kunte, “A novel data auditing approach to achieve data privacy and data integrity in cloud computing," in *2017 International Conference on Electrical, Electronics, Communication, Computer, and Optimization Techniques (ICEECCOT)*, pp. 306{310, Dec 2017.

[14] S. Sasmal and I. Pan, “Mutual auditing framework for service level security auditing in cloud," in 2017 *Third International Conference on Research in Computational Intelligence and Communication Networks (ICRCICN)*, pp. 297{302, Nov 2017.

[15] L. K. Shar and H. B. K. Tan, “Auditing the xss defence features implemented in web application programs," *IET Software, vol. 6*, pp. 377{390, August 2012.

[16] Marin Vukovic, Damjan Katusic, Renato Soic, Mario Weber, "Rule-based system for data leak threat estimation", *Software Telecommunications and Computer Networks (SoftCOM) 2017 25th International Conference on*, pp. 1-5, 2017.

[17] Sherry Zhu, Eric Guo, Max Lu, Anna Yue, "An efficient data leakage prevention framework for semiconductor industry", *Industrial Engineering and Engineering Management (IEEM) 2016 IEEE International Conference on*, pp. 1866-1869, 2016.

[18] Arun Prabhakar, Savin P. S., K. Chandrasekaran, "On-the-Fly Encryption Security in Remote Storage", *Advances in Computing and Communications (ICACC) 2015 Fifth International Conference on*, pp. 163-168, 2015.

[19] Tobias Wüchner, Alexander Pretschner, "Data Loss Prevention Based on Data-Driven Usage Control", *Software Reliability Engineering (ISSRE) 2012 IEEE 23rd International Symposium on*, pp. 151-160, 2012.

[20] Liu Liu, Olivier De Vel, Qing-Long Han, Jun Zhang, Yang Xiang, "Detecting and Preventing Cyber Insider Threats: A Survey", *Communications Surveys and Tutorials IEEE*, vol. 20, no. 2, pp. 1397-1417, 2018.

Loss Prevention & Fraud Detection

ORIGINALITY REPORT

8%	7%	2%	5%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

PRIMARY SOURCES

1	groups.drupal.org Internet Source	2%
2	dspace.thapar.edu:8080 Internet Source	2%
3	Submitted to University of Sheffield Student Paper	1%
4	Submitted to ABV-Indian Institute of Information Technology and Management Gwalior Student Paper	1%
5	www.hpenterprisesecurity.com Internet Source	1%
6	Submitted to University of Glasgow Student Paper	<1%
7	www.washcost.info Internet Source	<1%
8	Submitted to University of Newcastle upon Tyne Student Paper	<1%

9	www.waset.org Internet Source	<1%
10	Submitted to Thapar University, Patiala Student Paper	<1%
11	www.airccse.org Internet Source	<1%
12	gdeepak.com Internet Source	<1%

Exclude quotes	Off	Exclude matches	< 8 words
Exclude bibliography	On		