An exploratory analysis of insider attacks on financial information systems in Sri Lanka: a case study based on a database

A. L. D. S. Chathuranga
15020101
D. D. D. I. Perera
15020495
W. D. R. Fernando
15020207

This dissertation is submitted to the University of Colombo School of Computing in partial fulfillment of the requirements for the Degree of Bachelor of Science Honours in Information Systems.

University of Colombo School of Computing
35, Reid Avenue, Colombo 07,
Sri Lanka
January 2020

Declaration

I, A.L.D.S.Chathuranga (15020101) hereby certify that this dissertation entitled An exploratory analysis of insider attacks on financial information systems in Sri Lanka: a case study based on a database is entirely my own work and it has never been submitted nor is currently being submitted for any other degree.

20th February 2020 ……..

Date Student's Signature

I, D.D.D.I.Perera (15020495) hereby certify that this dissertation entitled An exploratory analysis of insider attacks on financial information systems in Sri Lanka: a case study based on a database is entirely my own work and it has never been submitted nor is currently being submitted for any other degree.

20th February 2020 ……..

Date Student's Signature

I, W.D.R.Fernando (15020207) hereby certify that this dissertation entitled An exploratory analysis of insider attacks on financial information systems in Sri Lanka: a case study based on a database is entirely my own work and it has never been submitted nor is currently being submitted for any other degree.

20th February 2020 ……..

Date Student's Signature

I, Kasun de Zoysa, certify that I supervised this dissertation entitled An exploratory analysis of insider attacks on financial information systems in Sri Lanka: a case study based on a database conducted by A.L.D.S.Chathuranga, D.D.D.I.Perera, W.D.R.Fernando in partial fulfillment of the requirements for the degree of Bachelor of Science Honours in Information Systems.

20th February 2020 ……..

Date Supervisor's Signature

I, Kenneth Thilankarathna, certify that I supervised this dissertation entitled An exploratory analysis of insider attacks on financial information systems in Sri Lanka: a case study based on a database conducted by A.L.D.S.Chathuranga, D.D.D.I.Perera, W.D.R.Fernando in partial fulfillment of the requirements for the degree of Bachelor of Science Honours in Information Systems.

20th February 2020 ……..

Date Supervisor's Signature

Acknowledgement

We would like to take this opportunity to express our gratitude to the people who have given their immense support for the preparation of this thesis and all the research work.

Foremost, we would like to express our sincere gratitude to our supervisors Dr. Kasun de Zoysa and Mr. Kenneth Thilakarathna for the continuous support of our final year research. We would like to thank them for their patience, motivation and immense knowledge given during the research period. Their guidance helped us a lot in proceeding our work and preparing the document materials including thesis.

Besides our supervisors, we would like to thank Mr. Primal Wijesekera for his valuable time, encouragement and comments given to direct us in the correct direction during the time of research. In addition to that, we would like to thank the personnel who helped us in conducting interviews and providing information about the banking environment.

Abstract

Despite the fact that security implementations are already placed in financial systems, there are numerous occasions around the world where financial systems were compromised by inside attackers all over the world. So there is a growing need to find ways and means to discover the potential threats to information security within banks. The research approach will be based on the discovery of potential information security issues and vulnerabilities existing in the system by analyzing patterns in data present in the back end system, procedures and relationships between data entities that may lead to insider threats and vulnerabilities with respect to a database. The aim of the research is to identify information security issues prevailing in information systems of a Sri Lankan bank as a case study, then identify the root causes behind their existence and to provide proper guidelines and solutions which will assist in making information systems in banks more secure in a similar context.

Table Of Contents

List Of Abbreviations 8

List Of Figures 9

List Of Tables 9

Chapter 1 - Introduction 1

1.1. Background of the Problem 1

1.2. Problem Statement 4

1.3. Research Questions 5

1.4. Goals and Objectives 5

1.4.1. Goal 5

1.4.2. Objectives 5

1.5. Significance of the Research 5

1.6. Research Approach 6

1.7. Limitations, Delimitations and Assumptions 6

1.7.1. Limitations 6

1.7.2. Delimitations 7

1.7.3. Assumptions 7

1.8. Contributions 8

Chapter 2 - Background 9

2.1. Information Security 9

2.1.1. Why has information security become a major concern for banks? 9

2.1.2. Information security principles and its implications on the financial sector 11

2.1.2.1. Confidentiality 11

2.1.2.2. Integrity 12

2.1.2.3. Availability 12

2.2. Financial Information systems 12

2.3. Regulations in Financial Sector 13

2.3.1. Information Security Management 14

2.3.2 Information Security Risk Management 14

2.3.3. Security Considerations 14

2.3.4. Legal Requirements 14

2.3.5. Regulatory Requirements 15

2.3.6. International Standards 15

2.3.7. Information Security Objectives 15

2.4. Insider Attacks in financial sector 16

2.5. Database Security and Analysis 18

Chapter 3 - Methodology and Design 20

3.1. Introduction 20
3.2. Research Questions 20
3.3. Research Design 20
3.3.1. Research Purpose 21
3.3.2. Research Approach 21
3.4. Data Collection 22
3.5. Data Analysis 23
3.5.1. Database Analyser Tool 23
3.5.2. Table Relationships 24
3.5.3. Data Visualization 27
3.6. Results Evaluation 28
3.6.1. Query Parser 28
3.6.2. Interviews 29

Chapter 4 - Results and Evaluation 31

4.1. Results 31
4.1.1. Direct Information Security Issues 31
4.1.1.1. Users have NULL values as their password 31
4.1.1.2. Session logout times and the authority level used are not recorded in certain user login sessions. 32
4.1.1.3. Relationships between tables can be recreated 33
4.1.1.3.1. Fluctuation patterns of cash and money deposits of the bank can be obtained 35
4.1.1.3.2. Transaction and behavioural profile of a bank account holder can be created 37
4.1.2. Indirect Information Security Issues 40
4.1.2.1. Essential data columns contain empty values 40
4.1.2.2. Primary keys contain two formats 41
4.1.2.3. Two date attributes for the same record 43

4.2. Evaluation **45**
4.2.1. Evaluating the precision of table relationships generated using the Database Analyser tool 45
4.2.2. Interview Findings 45

Chapter 5 - Discussion 48
5.1. Overview of issues identified 48
5.2. Challenges faced during the research 50
5.1.1. Challenges faced in reading LDF files 50
5.1.2. Challenges faced in relationship creating process implementation 51
5.1.3. Challenges faced in implementing the query parser 51
5.3. Reasons for the identified issues 52
5.4. Recommended mitigation techniques for information security issues 53

5.4.1. Application Validation 53
5.4.2. Use strong encryption mechanisms for passwords 53
5.4.3. Oracle Database Vault 53
5.5. Conclusion 54
5.6. Future Work 55

References **56**

Appendix A **63**
1. Code for relationship generator 63
2. Code for query parser 65
2.1. Parsing method 65
2.2. Refining method 67

Appendix B **68**
1. Refined table joins obtained from query parser 68
2.Table joins obtained from Database Analyser 70

Appendix C **72**
Interview Script 72

List Of Abbreviations

SWIFT	Society for Worldwide Interbank Financial Telecommunication
ATM	Automated Teller Machine
CBSL	Central Bank of Sri Lanka
VAPT	Vulnerability Assessment and Penetration Testing
InfoSec	Information Security
LDF	Log Data File
MDF	Main Data File
CIA	Confidentiality, Integrity and Availability
JSON	JavaScript Object Notation
MSSQL	Microsoft Structured Query Language
JDBC	Java DataBase Connectivity
NIC	National Identity Card
SQL	Structured Query Language
MSSMS	Microsoft SQL Server Management Studio
RDBMS	Relational Database Management System
SLIPS	Sri Lanka Interbank Payment System

List Of Figures

2.1. CIA Triad 11

3.1. High level architecture of the Database Analyser 24

3.2. JSON relationship structure format 25

3.3. Relationship generation and data retrieval process 25

3.4. Interface of database tool for checking related columns 26

3.5. Interface of database tool for relationship generation process 26

3.6. Sample chart created using Database Analyser 27

3.7. Query parsing process 29

3.8. Format of the relationship results obtained from query parser 29

4.1. Accounts with NULL passwords from PASS table 32

4.2. Empty IO_OUT values 32

4.3. List of unique table relationships obtained from Database Analyser 35

4.4. Total receipts and payments by each year in bank 36

4.5. Total receipts and payments by each month in bank 37

4.6. 2013 receipts & payments 39

4.7. 2013 receipts & payments 39

4.8. 2012 receipts & payments 39

4.9. 2011 receipts & payments 39

4.10. Empty NIC records and NIC stored in different formats in CUSTOMER table 40

4.11. Records with mismatching NIC numbers and date of birth values 41

4.12. The two formats of the primary key present in the PAYMENT table 42

4.13. Bulk insertion operations carried on 2008-09-30 43

4.14. Two date columns with time and without time in PAYMENT table 44

List Of Tables

4.1. List of database tables used to recreate table relationships 35

4.2. Interviews results 47

Chapter 1 - Introduction

In an era where technology and finance have become the leading drivers of every economy around the world, these two always go hand-in-hand to offer cutting edge services which simplify interactions between customers and financial service providers. Amounting to the impacts of globalization, Sri Lankan financial service providers also have adopted these technologies in order to stay on par with these modern trends in finance and technology. These continuous developments and innovations in technology have significantly impacted the operations of banks and other financial institutions which are having to face the challenge of adapting, innovating and responding to the opportunities faced by computer systems, telecommunications, networks, and other technology-related solutions to drive their business in an increasingly competitive domestic and global markets.

Introduction and adoption of the internet and the advancement of network infrastructure, opened doors to many businesses including banks and financial service providers to expand their product and service offerings while improving the efficiency in delivery. This very accessibility and dynamic nature of the internet could indicate both benefits and risks. So it's quite evident that banks and financial institutions rely highly on information technology and the internet to operate and to be competitive in their relevant markets, which would ultimately increase the potential risks, for both individual banks and the financial industry at large[1].

So it is important to identify the extent and impact of these potential risks to financial service providers originating from different threat sources. In order to be resilient to these potential risks and threats, these financial institutions follow various measures, guidelines, frameworks and also adopt risk assessment mechanisms along with tight internal controls. These guidelines are defined at various levels of management and also by the relevant authorities and governing bodies of respective countries. This study hopes to assess the above-mentioned measures followed by financial institutions within a selected context with the objective of exploring the effectiveness of laid out control mechanisms through the analysis of a transactional dataset of a financial institution.

1.1. Background of the Problem

Over the past couple of decades, there have been significant transformations in the banking and financial sector due to internal and external factors, including business model

transformation, adoption of advanced technologies, changing regulatory environments, etc. During this transformations, information security has become a secondary objective prioritizing business and functional goals. In this context, Information security has become a major challenge faced by banks and financial institutions in the current cyber ecosystem.

The country and society are moving and leaning towards completely interconnected systems across multiple services and platforms. With the evolving technology, people expect their lives to be smarter by having access to everything with a single touch. So with this evolution of technology and inter-connectivity between all information systems, the threat to information security is a very delicate subject. Financial institutions are placed in a special situation since any financial data is considered sensitive in nature compared to other data types and the industry is tightly regulated both in Sri Lanka and elsewhere as well.

Within the context of Sri Lanka, financial service providers are classified, regulated, controlled and monitored by the Central Bank of Sri Lanka (CBSL) as these institutions dominate the financial system within the country and also play a critical role within the Sri Lankan financial system, as they are engaged in the provision of liquidity to the entire economy, while transforming the risk characteristics of assets[2]. According to the CBSL, banks and financial institutions are also engaged in providing payment services, thereby facilitating all entities to carry out their financial transactions. On the other hand, banks can create vulnerabilities of systemic nature, partly due to a mismatch in the maturity of assets and liabilities and their interconnectedness. Therefore, the soundness of banks is important, as it contributes towards maintaining confidence in the financial system, and any failure may have the potential to impact on activities of all other financial and non-financial entities, and finally the economy[2].

According to Symantec's Executive Report on Financial Services[3], Banks and financial institutes are likely to remain top cybercrime targets, amounting to lucrative rewards when attacks are successful. Also, attacks against the financial industry are becoming increasingly sophisticated and highly targeted. 2016 Bangladesh Bank Heist is one of the prominent InfoSec breaches the industry has come across. The cybercriminals had targeted the Bangladesh Bank's (Central Bank of Bangladesh) account in the Federal Reserve Bank in New York[3]. They had processed 35 fraudulent fund transfer instructions via compromised access to the SWIFT(Society for Worldwide Interbank Financial Telecommunications) network which is a global payments network where most banks around the world are members of. The total of the 35 instructions amounted to close to US $1 billion. However, the attack was only successful in transferring a

sum of US $101 million through 5 instructions. Later investigation revealed that the perpetrators were suspected to have been aided by insiders within the targeted banks, who assisted in taking advantage of weaknesses in the banks' access to the SWIFT global payment network[5].

2019 data breach incident at the NorthAmerica, Canadian bank, Desjardins Group[4] resulted in the leak of sensitive data like names, addresses, birth dates, social insurance numbers (the Canadian equivalent of the social security number), email addresses and information on transaction habits of nearly 2.9 million members of the bank. Later it was discovered that the attack was carried out by an employee (an insider) in the IT department of the bank using the higher level of privileged access he had on the system and the access privilege of few other unsuspecting fellow employees. Noteworthy point is that security controls already implemented in the system could not detect this mass data gathering prior to the incident happening.

A local example of an attack on financial information systems is the massive fraud at Sampath Bank[6]. In this incident a branch manager along with the support of a regional manager of Sampath Bank had siphoned money from unsuspecting customer accounts through money transfers, practically operating as a private bank within the bank. The suspects have manipulated the victims cheques and slip amounts and the balances have been added as overdrafts to the victims accounts. The fraud has been carried out for nearly a year and a half and the fraud has only been discovered after the branch manager had received a transfer to another branch and a thorough audit has been carried out at the complaints of the victims at the receipt of notices to pay installments resulting due to the fraudulent transactions .

A very clearly distinguishable observation from the above-mentioned security incidents [4], [5], [6] is the involvement of insiders of the relevant financial institutions on these attacks. In 2016 out of all attacks carried out on financial institutions 58% of the attacks involved insiders. Out of which 53% were inadvertent actors while 5% included malicious insiders[28]. Due to the higher access privileges inside employees enjoy on the system and the intimate knowledge employees have on the system and its process, make an insider attack far more potentially damaging, whether the attack is done with malicious intent or not. The above security incidents[4], [5], [6] provide evidence on the financial loss and the damage to the value of the institution that may be caused by an insider attack. Therefore it is important to identify these insider threats faced by financial institutions and the underlying reasons for their existence accurately and as early as possible.

Preliminary literature review conducted on the insider threats research domain revealed that there is a lack of literature available on the domain. Majority of information security

research has been carried out on research areas such as outsider attacks on financial institutions, cyber threats on financial institutions, online financial applications, information security frameworks and security models on financial institutions [36], [37], [39], [40]. Only a few research have been carried out which explores the existence of insider attacks and which explains the underlying reasons behind the occurrence of insider attacks. Particularly in the Sri Lankan context, which represents financial institutions of the third world countries with relatively limited financial and technological resources yet with higher economical targets and pressure, the number of research that has been carried out is very minimal.

Insider threats on financial institutions is a very broad domain, consisting of numerous case specific applications, too numerous to be comprehensively studied in a single research. Therefore as the case study for this research a transactional database from a reputed Sri Lankan bank was used. The aim was to investigate the threats and vulnerabilities present in the financial information system of the bank that could be utilized by an insider to carry out an attack using a database dump.

1.2. Problem Statement

Financial institutions operate in an era where information technology has become a main pillar for economic growth and economic sustainability. Therefore in today's world financial institutions are totally dependent on the financial systems they use. But with very high rewards at stake if successful, for attackers financial information systems have become somewhat irresistible targets. When it is analyzed, it reveals that a considerable amount of these attacks involve insiders in carrying out these attacks[5], [6]. When the intimate knowledge of the system and the higher level of access to the system an insider has is taken into account, the damage and financial losses of an attack involving an insider is also higher. The frequency on which these attacks have taken place show that the underlying reasons behind recurrence of these insider attacks are not yet completely identified and addressed. Yet in the domain of information security (infosec) of financial institutions most of the research are focussed on security issues caused by outside attacks and solutions to prevent those outside attacks[36], [37], [39], [40]. Research carried on the implications of databases on infosec in financial institutions is even harder to find. Thus it is evident that there is a knowledge gap in the domain of insider threats on financial institutions. This research caters to this knowledge gap and lays the foundation for other researchers to carry out deeper research in the domain of insider threats on financial institutions.

1.3. Research Questions

1. Are there threats and vulnerabilities exposed to insider attacks, in systems and processes of Financial Institutions in Sri Lanka?

2. What are the root causes behind the existence of potential threats and vulnerabilities exposed to insider attacks, in systems and processes of Financial Institutions in Sri Lanka?

1.4. Goals and Objectives

1.4.1. Goal

Review, analyze and identify possible threats and vulnerabilities for financial institutes in Sri Lanka and provide best practices and guidelines to mitigate those threats and vulnerabilities.

1.4.2. Objectives

- Identify information security related threats which cannot be identified through automated vulnerability assessment and penetration testing (VAPT) tools
- Identify the root causes for the existence of InfoSec threats and vulnerabilities within the financial institutes.
- Discuss the best practice guidelines that must be introduced to financial institutes in Sri Lanka to mitigate the risks caused by the above threats and vulnerabilities.

1.5. Significance of the Research

Despite the fact that security implementations are already placed in financial systems, there are numerous occasions around the world where financial systems were compromised by both inside and outside attackers. Similarly in the Sri Lankan context, there are reports about certain InfoSec breaches and fraudulent dealings within certain commercial banks in the country. But only a handful of these reports reach the general public as these banks have the financial power to silence the media to uphold their reputation and trust among the general public. Recent ATM fraud in Sri Lanka, which involved skimming devices attached to ATM machines[8],

massive fraud at Sampath bank[6], hacking into the Commercial Bank of Ceylon internet banking facility in 2016[7] are some incidents that came into the light in recent years in Sri Lanka. This research will be conducted within the Sri Lankan banking context hence the findings of this research will be of significance when dealing with similar incidents in the future.

1.6. Research Approach

This research follows a nonexperimental exploratory research design. Research is mainly focused on identifying the vulnerabilities existing in financial information systems of financial institutions which can be utilized for carrying out insider attacks. A database dump of a Sri Lankan Bank is used for the case study of the research. The intention is to identify the vulnerabilities and threats that can be used to materialize an attack by an insider. Any manipulation in the data in the database will not present the real information security situation prevailing in the financial information system of the bank. Thus a non experimental research design is used in the research.

The database dump used in the case study is a Microsoft SQL Server 2000 database (MSSQL 2000) dump containing transactional data of the bank. The database used for the research consists of 440 tables out of which 120 tables contained data. The database contained transactional data spanning a period of 18 years. Interviews with the employees of the selected bank were used to obtain necessary information to evaluate the findings made in the research. As a result research involves both quantitative and qualitative data, therefore both quantitative and qualitative research approaches are used in this research, therefore the research falls under the mix method research category.

1.7. Limitations, Delimitations and Assumptions

1.7.1. Limitations

- Results of the research can not be generalized to all institutions in the financial information security domain since this research results are based only on a case study carried out on one single bank in Sri Lanka. Further research must be carried out using a larger sample to generalize these findings.
- The research approaches and methods used in this research were exclusively used to tackle the problems the researchers had to face in finding answers using a relational

database. Technical implementations and logics used may not be able to be used in its entirety, on other types of databases.

1.7.2. Delimitations

This research study is only focused on finding answers to research questions within the context of Sri Lanka. Even though the topic InfoSec is similarly applicable to every other industry and country around the world, data will only be available within the context of Sri Lanka to conduct this research.

The financial industry at large consists of many financial service providers. But this study will only consider banks and similar financial institutes in conducting research. Furthermore, the names of the financial institutions providing data will not be disclosed to uphold confidentiality. Through the systematic literature review, it was evident that there's satisfactory literature available with regard to InfoSec threats originating externally (Outsider threats). Hence, this study only accommodates analysis and findings of threats originating from within financial institutions (Insider threats).

With the dataset available, insider threats to the financial information are discussed with respect to the transactional data connected to customer accounts. We are not considering the whole database with its tables due to the limited time frame. Only the tables related to customer transactions and customer details are taken into the analysis.

1.7.3. Assumptions

- All deposits to the bank are saved in the RECEIPT table and the payments done are saved in the PAYMENT table.

1.8. Contributions

This research fills the existing knowledge gap by contributing as following,

- In Sri Lankan context, there is no proper research work carried out focusing on information security and financial sector.
- Majority of research work in the domain is focused on outsider attacks in the financial sector known as cyber attacks.
- Although there are research works focused on insider attacks globally, analyzing a database for vulnerabilities in a financial information system has not been done before as per to the best of our knowledge.

Further, we deliver the following to the research community for future research in this area,

- A novel method to find relationships between tables in a SQL server database in which the foreign key constraints have not been established.
- JSON based structure to store table relationships for a relational database.
- A query parser written using Java to extract the possible table joins from SQL statements used for creating SQL server database views.
- Database tool to find potential relationships within tables and extract data which can be reused with fewer configurations.

Chapter 2 - Background

2.1. Information Security

Information security (infosec) has become a challenging topic in the present context as a result of the advancement of technology in every industry around the world. Information is considered a key factor and has become the monetary model of most organizations. We have reported some incidents which showed us how valuable information is, such as Facebook data breach which led to stealing data in 50 million user accounts without any permission[10]. All organizations and individuals pay more attention to information security due to being a vital player for the existence of any entity in any industry. What is really meant by this term information security? It's a set of strategies for managing the processes, tools, and policies essential to detect counter threats and prevent any kind of information(digital or non-digital) ensuring its confidentiality, integrity, and availability[9]. Within this domain, information is considered as an asset. Organizations follow the required steps and measures to safeguard their information assets from unauthorized access and exposure to the outside world.

In 2016, the European Parliament and council introduced a regulation called GDPR(General Data Protection Regulation) in EU law, which is directly connected with information security. All companies operating in Europe must comply with these standards. Every organization has one or more information systems. The information must be stored safely. The exposure of a single tiny piece of data can lead to destroying a whole organization. Depending on the information circulate within the organization premises, this broad topic can be divided into several categories like application security, cyber security, cryptography, incident response, vulnerability detection, etc. So, our research is based on this information security domain applying it into the process going on in financial institutions within the Sri Lankan context.

2.1.1. Why has information security become a major concern for banks?

Financial institutions including banks are dealing with one of the most sensitive datasets in any country. It's financial records and transactions data. Information security plays a critical

role because banks and other financial institutions are committed to the security of their customer's financial and personal information. Risk of occuring financial losses and security breaches are very familiar topics discussed nowadays and not miracles anymore. In Sri Lanka also there were incidents where the attackers went through the holes in the existing security implementations. Massive Fraud At Sampath Bank[6], ATM skimming incidents[8] are some of the local incidents and Bangladesh bank heist[11] is a popular example for an international security breach. So these incidents show us that there's something wrong with the security controls established, even the technology has been improved so much by now.

As a responsible entity towards the customers, financial institutions must be able to guarantee that they are not leaking and misusing their customer data. That is the key point to win customer satisfaction. If any customer feels that his money is not going to be safe in any bank, will he invest or deposit money with that bank? Criminal profiles are shifting from random disorganized hackers to well funded organized crime groups gradually[12]. When it comes to the threats to financial institutions, they have evolved from infrastructure level to application, data levels. With the development in information systems, the possibility to become a victim of a threat has gone higher than ever.

Cyber security is a subset of information security which simply means protecting information from the internet. In banking environments this is the aspect of information security that the banks are focused on specially recently. The number of banking applications has been increased. Almost every bank in Sri Lanka also has a web presence and mobile applications. So, there’s a huge possibility to originate threats from global level via the internet since the institutions have gone public through their web applications. Thousands of online transactions are happening per day nowadays. People tend to use such a facility as it saves time in their busy schedules. They have the trust on financial institutions. Even banks are also a set of business organizations which need profits and are extremely dependent on investments of ordinary people. Therefore security controls must be able to detect and prevent both insider and outsider attacks without breaking the financial operations and without losing customer loyalty. Otherwise financial organizations will fail to serve as a reputed entity in society.

Information security is based on 3 main principles or can be called them as objectives. They are recognized as CIA triad. Confidentiality, Integrity, and Availability are those factors. Protecting these properties of information is the objective of information security. This is a well-known model for developing security policies in organizations. Especially information systems must be able to ensure these properties.

2.1.2. Information security principles and its implications on the financial sector

Information security is based on 3 main principles or can be called them as objectives. They are recognized as CIA triad shown in Fig 2.1. Confidentiality, Integrity, and Availability are those factors. Protecting these properties of information is the objective of information security. This is a well-known model for developing security policies in organizations. Especially information systems must be able to ensure these properties

2.1.2.1. Confidentiality

Confidentiality is ensuring the information is accessed by only the authorized people and can not be accessed by the people who are not authorized. The same information sent by the sender must be received by the receiver without any change. Cryptography, File permissions, ACL(Access Control Lists) and Encryption are some methods to protect the confidentiality of information.

In the context of banking and financial institutions, customers interact with these institutions implying that their personal as well as sensitive information will be kept confidential. This confidentiality is not limited to just account transactions of the customer but extends into all types of information that the bank possesses of the said customer[12]. The Financial Ombudsman of the country provides directives to banks and other financial institutes on protecting confidentiality of customers.

Fig. 2.1. CIA Triad

2.1.2.2. Integrity

Integrity is ensuring the accuracy of data and modification of information is restricted to unauthorized parties. Encryption, Hashing, Digital Signatures, Backup and Audit are some methods to ensure the property, integrity.

Typical information security attacks targeting banks and financial institutions are focussed on breaching this principle. Threat agents whose prime objects are financial gains, breach integrity of the Information System of the bank by acquiring authorized credentials through different distributed attacks. After successfully compromising the system, the attacker will be able to process fund transfers to desired destination accounts which was the case of the 2016 Bangladesh Bank heist[5], where the SWIFT network node at the Bangladesh Central Bank was compromised. The attackers who forwarded fund transfer instructions following this compromise were clearly not the authorized party to do so, so this amounts to breach of Integrity.

2.1.2.3. Availability

Availability is ensuring that the information can be accessed by authorized parties when requested. Redundancy processes, RAIDs, failover and high availability clusters can be introduced as some mechanisms to achieve the availability of information.

In the context of a bank or a financial institute, there are specific types of attacks that can affect the availability and continuity of the banking system. The financial institution would fail to conduct transactions with its customers, business partners and vendors when availability is denied[10].

2.2. Financial Information systems

Financial information systems are computer based systems used in financial institutions to collect data, interpret data, analyze data and reporting. These systems contain mainly 3 components called Financial Accounting (FI), Funds Management (FM) and Controlling (CO) handling separate functions in different areas[14]. Integrating all these components, systems are built to effectively manage heavy workload in banks and other financial institutes. Functionalities of each component are listed briefly below.

1. Financial Accounting (FI)
    - Record all details of transactions in ledger accounts as assets.

i.e. revenues, expenses and liabilities

- Create financial statements
- Maintain ledgers in sub sections

  i.e. General Ledger, Accounts Receivable Ledger, Accounts Payable Ledger,

2. Funds Management (FM)
    - Identify the sources for funds
    - Control spendings in fundings

      i.e. Operating Funds, Ancillary Operations, Restricted Funds, Capital Funds

3. Controlling (CO) - Track revenues and expenses based on reports

Financial information systems must have a proper governance and security culture which relies on certain standards and regulations to resist against potential threats. A considerable amount of significant threats and risks which are apparent to the information system of banks are present as a result of weaknesses in aspects of governance and security culture[16]. The same idea was stressed by Kooper in 2011[19] that in practice imperfect implementations of Information Technology Governance (ITG) such as delayed or aborted information security projects, service level remained unmonitored etc. This eventually exposes the organization to the existing and emerging information security risk. Ula, Ismail & Sidek[17] also stress on this fact by mentioning that InfoSec practitioners have a lack of consensus in the definition of information security governance. Veiga & Eloff[18] also comprehensively present their findings on how information security culture is the reason as well as they key to overcome many challenges and threats in the context of InfoSec.

2.3. Regulations in Financial Sector

Information Security Management Directives by CBSL

In terms of security posture, organizations in Sri Lanka represent a spectrum of capabilities. However, just as the strength of a chain lies with its weakest link, so does the strength of information security in the financial services sector lies with its weakest member, which in turn poses a threat to all other members, which may potentially lead to financial fraud.

Thus, the Central Bank of Sri Lanka has issued the 'Baseline Security Standard (BSS) for Information Security Management', Assurance Level 1, Version 1.0 under the Banking Act Directions No. 4 of 2014. The key directives and guidelines in the BSS have been outlined below [13].

The Central Bank of Sri Lanka (CBSL), the Sri Lanka Computer Emergency Readiness Team | Coordination Center(Sri Lanka CERT | CC) and the Sri Lanka Banks' Association (SLBA) worked towards the establishment of the Baseline Security Standard for Information Security Management (BSS), based on the globally recognized ISO 27000 series of international Standards for information security. The implementation of the standard will be supervised by CBSL and the subsequent revisions to the Standard will be proposed by Bank Computer Security Incident Response Team (Bank CSIRT) to CBSL for consideration. The standard outlines the following fundamentals with regard to Information Security so as to provide a better understanding for the users of the same;

2.3.1. Information Security Management

The preservation of Confidentiality, Integrity, and Availability of information by the appropriate and systematic application of security controls to manage the risk of exposure to a threat, which arises due to the existence of vulnerabilities in information assets.

2.3.2 Information Security Risk Management

Information security risk management is the systematic approach to ascertaining the impact and livelihood of an information asset being exposed to a threat.

2.3.3. Security Considerations

All organizations are required to derive their security requirements to conform to the laws in Sri Lanka including the regulatory requirements set by the respective regulators and the international best practices adopted globally. Additionally, security requirements are also governed by the business objectives set by the board of directors and the senior management of the organization.

2.3.4. Legal Requirements

All organizations are liable to comply with the laws applicable in this regard including the Computer Crimes Act No. 24 of 2007, the Electronic Transactions Act No. 19 of 2006,

Payments Devices Frauds Act No. 30 of 2006, and Intellectual Property Act No. 36 of 2003 of which any violations amounts to an offence.

2.3.5. Regulatory Requirements

Local industry regulations/directives set forth by the CBSL and other regulatory bodies must be complied with.

2.3.6. International Standards

In order to be recognized as competent online/e-banking service providers, organizations need to comply with internationally recognized industry-specific security standards, such as PCI-DSS (Payment Card Industry Data Security Standard).

2.3.7. Information Security Objectives

Information security objectives must be identified supporting fulfillment of key business objectives within the framework of the information security policies, statutory requirements, other requirements and business processes.

As per the Baseline Security Standard (BSS) for information security management, Risk Management is defined as a fundamental component of any cost effective information security management system. Further it elaborates on how risk can be considered as a basis for Information Security Management, and that assets contain vulnerabilities due to weak design, production, implementation, handling, management and a host of other activities. These vulnerabilities may be exploited to give rise to threats. The combination of the likelihood of a threat being realized and the impact of that exposure is called risk, and is an important measure of the relative urgency and need to impose control measures to mitigate that risk.

Further the CBSL instigate that the adoption of BSS and its successive revisions, will introduce security controls which would mitigate identified risks. Also specifies the guidelines for Risk Management should be done in accordance with ISO 27005:2011. The establishment of Guidelines, policies, procedural manuals, schemes and templates by the respective organizations will aid the implementation of the BSS.

The BSS standards outline the guidelines and directives for 14 Security Domains (applicable areas) in an organization. It clearly and directly specifies the objective of the adoption of the guidelines outlined;

- Organization of Information Security Management

- Information Security Policy
- Third Parties
- Information Asset Management
- Human Resource Security
- Operations Security
- Communications Security
- Physical and Environmental Security
- Access Control
- Internet and Email Security
- Information Systems Acceptable Use
- Information Security Incident Management
- Acquisition, Development, and Maintenance of Information Systems
- Business Continuity Management

2.4. Insider Attacks in financial sector

With regard to information security, the term “attacks” refers to any guided or planned activity that is targeted on information assets with the intent of potential gains which are numerous in nature; financial, sensitive information, confidential data used for profiling and social engineering. These attacks are fundamentally classified into 2 types based on their origin, namely: Insider attacks and outsider attacks. The paper work done by B. Bojinov(2016) says that infosec issues and threats mainly fall under four main categories listed as; staff (human factor), Internal processes, Systems and External factors[20]. Staff includes both unintentional and intentional actions performed by internal people.

Insider attacks refer to the kind of attacks targeted on a system or an asset where the human component of the activity is originated within the relevant organization or the system. Whichever attacks initiated externally to the system are known as Outsider attacks. According to Gokhan Kul [15], by 2016, insider attacks have posed a major threat to the Financial sector. The possible reason for this would be that in the current technological landscape, banks and financial institutions have invested deeply in information security and controls, to prevent external threats to their systems and assets. Thus, attackers find it costly and time-consuming to propagate attacks externally which can bypass robust control mechanisms set in place by the said banks and

institutions. This was evident with the recent infosec incidents that were targeted on the financial sector around the world, because the majority of those attacks had been successful with adequate help from insiders[11]. So the current focus of information security experts in the financial sector has taken a turn towards potential threats faced from insider attacks.

The work done by Damenu & Beaumont[16] in 2017 argues that InfoSec threats are more imminent to the financial sector given that the reward for attackers are usually enormous in value. Thus criminals put in hours of effort in crafting spyware and malware or trick customers to visit spoof sites to capture sensitive data. L. T. Khrais presents an explanation of insider attacks that are done with the help of insiders with respect to online banking. They are fraud and theft, back doors or trap doors, errors and omissions, employee sabotage[21]. These internal attacks can be more severe as internal employees have intimate knowledge of the system and therefore attacks can go undetected. In 2010, Veiga & Eloff[18] presented their findings on how insiders pose a greater threat to the protection of information within an organization. The research presents a framework and an assessment instrument on the information security culture in organizations by considering that improper InfoSec culture is one of the major challenges and threats.

Human error can make a huge impact on security of data in banks and other financial institutions. In 2011 Ula, Ismail & Sidek[17] points out that the following threats with regard to Information Security; physical destruction of premises and systems by natural disasters; unintentional damage due to human error, abuse of system and sensitive information by employees or agents of the bank; systematic collection of sensitive information by foreign intelligence services; and external attacks which compromise confidentiality, integrity and availability of information. According to them the most common technology risk or threat to banking and financial institutions is phishing attacks which are based on social engineering. The other forms or attacks are spyware, Trojan horses, and key loggers. In 2015, Fazlida & Said[23] in their research on information security risk governance and implementation setback, bring into light that InfoSec threats are mainly caused by outdated security controls and architecture. Other considering factors are; careless or unaware employees, cloud computing use, and unauthorized access.

Singh & Malpani[22] reveal that accidental entry of bad data by employees, accidental destruction of data by employees, the introduction of computer viruses to the system, natural and human-made disasters, employees' sharing of passwords, and misdirecting prints and distributing

information to unauthorized people are the most significant factors in breaching the security of any financial institution.

2.5. Database Security and Analysis

Computer software artifacts such as application software are complex pieces of work mankind has developed. Securing these artifacts against intelligent attackers who try to exploit flaws in software design and constructs has been continuously proven to be a greater challenge.

Application software developed for many industries nowadays are faced with troubles in terms of design flaws and implementation bugs which in turn gives rise to unacceptable security risks. As per I. A. Bubu[24], when developing application software, good software engineering practices would produce quality and secure software products. Further he argues how good software security practice involves thinking security early in the software lifecycle, knowing and understanding common problems (such as language based flaws and pitfalls), designing for security, and subjecting all software artifacts to thorough risk analysis, review and testing.

When considering database management systems (DBMS) which is a type of application software, the ultimate objective of a DBMS would be to retain and protect the data stored. It is the duty of the database administrator to ensure the security of the DBMS. Some of the key measures used in ensuring security in a banking DBMS are; database server enabling authorization and access control data views using data encryption, achieving authorization and data access control through passwords, implementing specific access rights to different user roles, restricting the actions that could be performed by subjects (users) on certain objects, user management through names, passwords, group names and access levels, utilizing views within the database to fulfill security features by allowing certain users to have access to a logical representation of a portion of the database according to their requirement access, regulating user rights by ensuring that the database administrator has full control in granting and cancelling rights of any user at any given time [24].

A significant amount of sensitive and valuable data is stored and processed in relational databases of banking institutions. Therefore numerous database-specific tools are developed for emergency database recovery. J. Wagner, A. Rasin and J. Grier in their research study [26], present a universal tool that seamlessly supports many different databases in rebuilding tables and other data content from any remaining storage fragments on disk or in memory. The tool defines an approach for reverse engineering storage is databases with minimal user intervention.

Also the authors verify the tool's ability to recover both deleted and partially corrupted data directly from the internal storage of different databases. Their work is classified under software-based restoration techniques in the context of relational database management systems. Further along the line of their work, they present how the tool can 'reconstruct volatile artifacts' of a database. In achieving the aforementioned, Firstly, the tool can recover the newly introduced data from inserts and updates. Secondly, the tool can recover recently performed user actions (i.e.,reconstructing the fact that data was inserted, deleted or updated). Thirdly, the tool can discover information about the changes that were canceled and undone (i.e., aborted transactions). The latter category is the most interesting, because this information would normally be unavailable to users even if the database were operating normally.

I. Drabikova, K. Matiasko and A. Lieskovsky, in their paper titled "Database Object Dependency Tool" [25], stress on the fact that retail banking and telecommunications sector are continually developing information systems for their specific purposes having efficient and elegant database tools. Similarly, they present a new approach for analyzing database objects with DB Object Mapper Tool that is being developed for mapping objects from large and complex database systems.

In the works of Andrew Ilyas, Joana M. F. da Trindade, Raul Castro Fernandez and Samuel Madden titled as 'Extracting Syntactic Patterns from Databases', the authors emphasize on the matter that, attributes in the enterprise databases are highly structured such that they follow simple syntactical patterns [45]. More common examples of structured attributes are dates, product identifiers, phone numbers, etc. Often these are stored as strings in the database, but if they could be labeled with richer structural information about the format of values, indexing, searching and comparing values, and finding exceptional or outliers values, could be done much more efficiently. In their paper, they present a new method to extract these syntactical patterns to better understand the attributes and correlate them with other attributes in different tables, entries in the same database to find relationships among each and every table. This method is introduced as the XSYSTEM.

Chapter 3 - Methodology and Design

3.1. Introduction

This chapter describes the methodology of the research that was followed in the process of obtaining the research objectives of this research titled - An exploratory analysis of information systems of financial institutions in Sri Lanka: a security perspective. Chapter provides a detailed description of the steps followed throughout the research process.

3.2. Research Questions

1. Are there threats and vulnerabilities exposed to insider attacks, in systems and processes of Financial Institutions in Sri Lanka?
2. What are the root causes behind the existence of potential threats and vulnerabilities exposed to insider attacks, in systems and processes of Financial Institutions in Sri Lanka?

3.3. Research Design

This research's aim is to review, analyse and identify possible threats or vulnerabilities present in information systems that can be exploited by insiders to carry out an attack on financial Institutions in Sri Lanka, identify the root causes behind the presence of those issues and provide a set of changes, new actions, and precautions that can be implemented by Financial Institutions in order to mitigate the risks they face due to those issues. This research is designed as a non experimental exploratory research which utilizes both quantitative and qualitative data analysis approaches.

Through the preliminary literature review it was realized that in the research domain, Information security of financial institutions literature is mainly focused on outside attacks. Literature available on insider attacks are comparatively rare and infrequent.

This research acknowledges this knowledge gap and tries to reduce this gap by focusing on the information security issues concerning insiders. Intention of the research is to capture the standard everyday information security environment prevailing inside a financial institution

based on the data available on that system. Any manipulations in data would create a false picture of the security environment inside the financial institution. As a result, no data manipulations can be performed on the original data during the research. This condition makes the research fall under the category of nonexperimental research. In nonexperimental research, the independent variables involved can not be manipulated by the researchers but should be analysed as it is available[38].

3.3.1. Research Purpose

This research examines the information security issues financial institutions face due to parties inside financial institutions commonly referred to as insider attacks. Whole research process involves an element of discovery, which includes the discovery of potential information security issues and vulnerabilities existing in the system, patterns of data present in the back end of the system, back end procedures, and relationships between data structures that may lead to insider threats and vulnerabilities. Therefore due to this exploratory nature of the research, it is evident that the purpose of the research is exploratory.

3.3.2. Research Approach

The case study of the research is done on a bank database containing a large amount of transactional data. In the research it is required to analyse the data and identify the threats and vulnerabilities present in the system that can be exploited by an insider. This involves quantitative analysis techniques such as recognising data patterns, trends, forecasting,etc. Further it is also required to perform qualitative analysis techniques in order to examine the structural and qualitative characteristics of the data. Furthermore research also involves interviews carried out with the employees of the bank which basically involves qualitative analysis. Therefore in this research both qualitative as well as quantitative research analysis techniques are being used. A research approach used by researchers that involves more than one type of research method is known as mixed method research. It may include a mix of both quantitative and qualitative methods, a mix of quantitative methods or a mix of qualitative methods.[29]. Therefore this research follows a mixed method research approach.

3.4. Data Collection

This research study looks into the information security issues financial institutions are facing and the underlying root causes for the existence of those information security issues by analysing the data available in a database of a Sri Lankan bank. Therefore mainly the type of data used in the research is secondary. Secondary data is defined as data that has been already collected and tabulated by other sources[34]. In this study use of secondary data is appropriate and justifiable as it provided researchers with a large quantity of data spanning a longer period (18 years) which would have been impractical and hugely time consuming if primary data collection methods were used in this research[34]. Finding a bank willing to cooperate with the researchers to carry out the research was the main concern that had to be addressed when it comes to data collection.

Through an agreement with a reputed Sri Lanakan bank(name not disclosed to preserve privacy), a suitable database was finalized for the research. The database was provided to the researchers after sanitizing it first. Sanitization is a process of hiding sensitive data in a data set in order to preserve the confidentiality of sensitive data in a data set. Further more data sanitization process promotes the sharing of transactional databases among organizations by minimizing the impact on data utility of non sensitive data in a data set whilst protecting the confidentiality of the sensitive data in it[35].

The database obtained is a Microsoft SQL Server 2000 (MSSQL 2000) database dump, including both LDF(log data file) and MDF(main data file). Database consists of a total of 440 tables of which 120 tables contained data covering the main operations including deposits on the bank, withdrawals from the bank, interests, payments, loans, savings accounts, etc. It also contained data on customers, system users.

Initially the database dump was attached to a MSSQL 2000 local server running on a Windows XP operating system, as MSSQL 2000 was compatible only with windows XP operating system. Functionalities and the graphical support provided by the MS SQL Server 2000 was minimal therefore the database was upgraded to MS SQL Server 2008 R2. This upgrade enabled access to improved functions and more user friendly interfaces and compatibility with modern operating systems(Windows 10) with better performance.

3.5. Data Analysis

Data analysis refers to the process of examining data in order to come up with conclusions about the information contained in the data[30]. The database contained both qualitative and quantitative data. Therefore both quantitative and qualitative analysis techniques were used to analyse the database data.

The main intention of analysing database data is to:

- Identify relationships between tables based on table interactions
- Perform timeline analysis on business processes, tables, transactions based on the log history of the database.
- Identify significant data attributes in terms of information security
- Identify anomalies and patterns visible in the data
- Identify potential security issues existing in the system

The 2 main tools that were used in this research to analyse data are:

1. Microsoft SQL Server Management Studio version 18.0 (MSSMS)
2. Database Analyser tool specifically built for the research

3.5.1. Database Analyser Tool

To Identify foreign key relationships among tables in the database, Database Analyzer tool was implemented using JavaFX[31] and Java libraries; JDBC[32] and GSON[33]. When 2 columns of 2 corresponding tables are selected, database analyzer tool checks for the similarity of the content between the 2 columns and predicts whether a relationship is possible or not between those 2 tables.

JavaFX was used to build Database Analyser since Java was already familiar with the research team and had the functional requirements necessary to develop the tool. Tool was designed as a standalone application in order to avoid unnecessary exposure of data to the internet, which was a condition the researchers had to obliged as per the agreement with the bank. Further using JavaFX simple user interfaces can be easily developed. JDBC library was used to connect the SQL server database as it was the most used and standard database library used for Java standalone applications. GSON is a java library provided by Google for converting java objects into JSONrepresentation. GSON was used because of its stability and robustness.

The high level architecture of the tool is shown in Fig. 3.1.

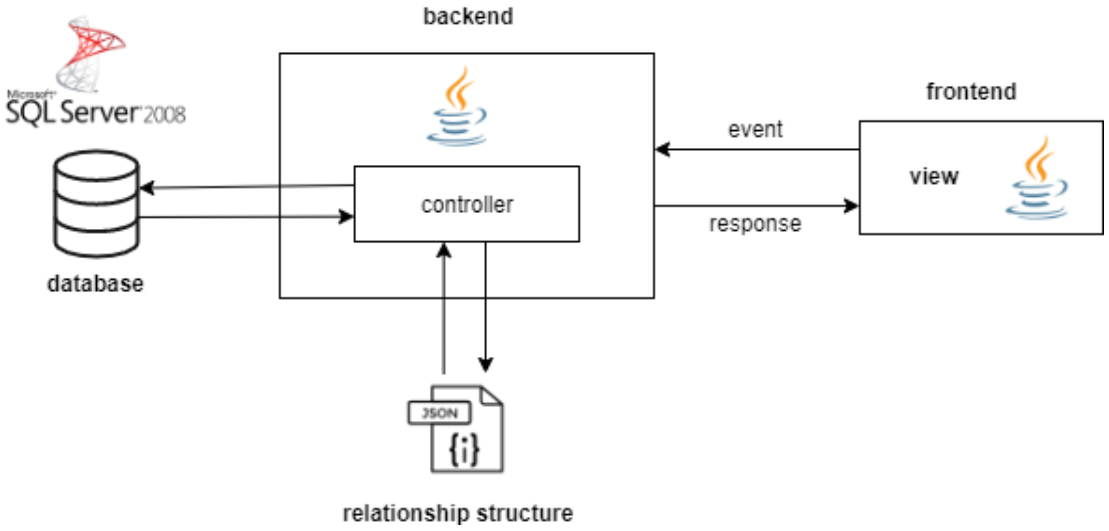

Fig. 3.1. High level architecture of the Database Analyser

Following functional capabilities were added to the Database Analyser to support the data analysis process:

- Find relationships between tables
- Filter and extract data by columns
- Perform table joints and extract data
- Download data in JSON format
- Profile and visualize customer data

3.5.2. Table Relationships

Tool utilizes an algorithm which checks for similar contents in 2 table columns for a threshold value of 80% to determine whether there is a relationship between the two tables or not. If the similarity between the contents exceeds 80%, the tool determines that a relationship is present between the 2 tables and is stored in a JSON file separately. In this file, there's a JSON array that contains objects for every table in the database holding another relation tables object array within each object. Using Java as the core programming language, the algorithm was implemented to create table relationships. This way all the table relationships identified through Database Analyser were saved in the JSON file. Information in this JSON file is used by the Database Analyser tool as source data when implementing the other functionalities of the tool.

```
{
  "structure": [
    {
      "table": "table1_name",
      "relations": [
        {
          "table":"relation1_name",
          "key1":"table1_key",
          "key2":"relation1_key"
        }
      ]
    }
  ]
}
```

Fig. 3.2. JSON relationship structure format

As illustrated in Fig. 3.2, "structure" is a JSON array that contains an array of JSON Objects divided among the tables in the database. Each table object has a key named as relations and it holds all the tables and the corresponding columns (keys) by which the foreign key relationships are established with that table. Following Fig. 3.3 shows the steps involved in the process of generating table relationships.

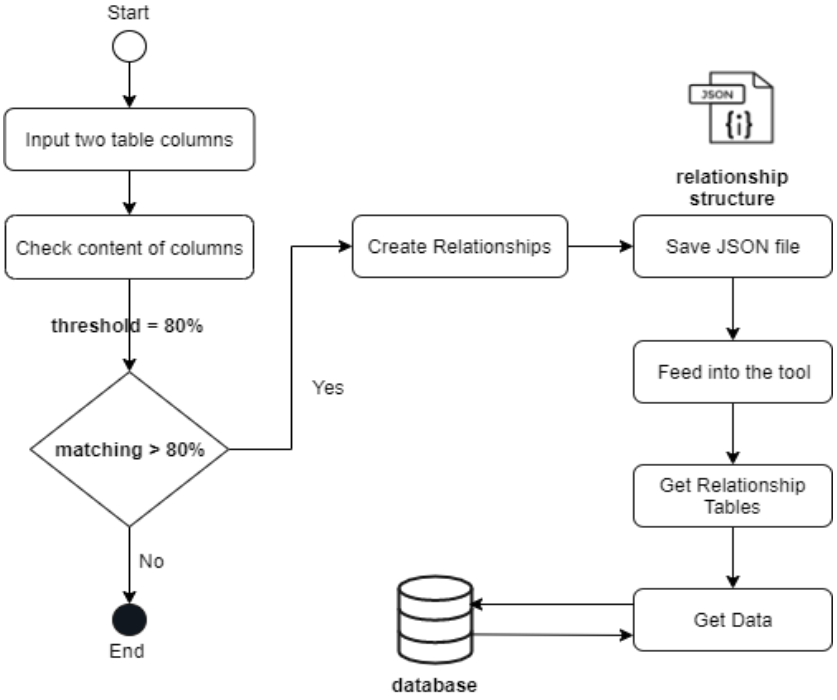

Fig. 3.3. Relationship generation and data retrieval process

Through the interface shown in Fig. 3.4. when 2 tables are selected all possible relationships that can exist between the 2 tables are generated. The list of generated relationships may contain irrational relationships. Those irrational relationships were removed when creating the JSON file which was used as the source data file for other functionalities.

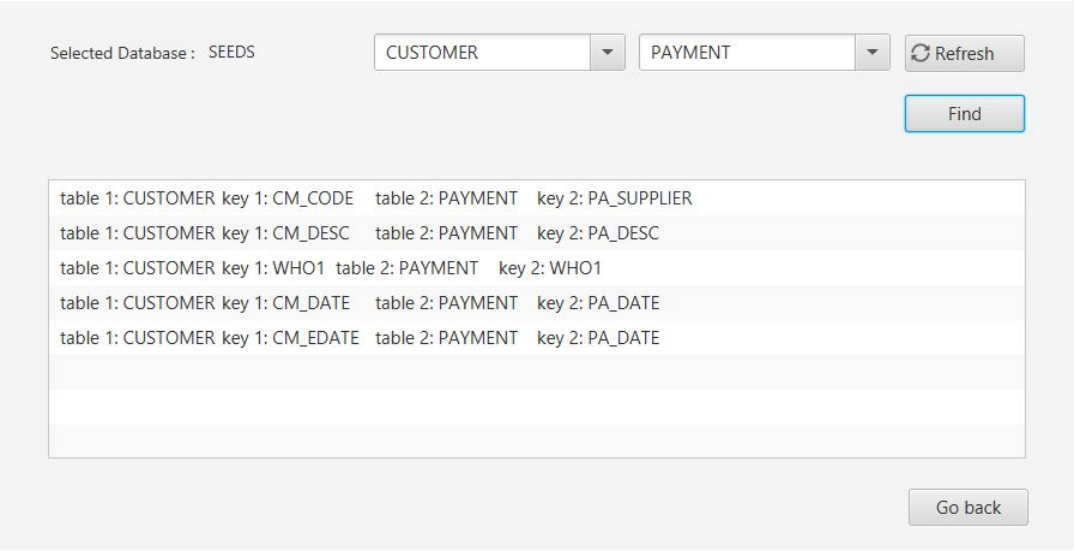

Fig. 3.4. Interface of database tool for checking related columns

Through the interface shown in Fig 3.5. it was checked whether any relationship was present between 2 specific columns. Here again only the rational relationships were used to create the JSON file which was used as the source data file for other functionalities. Format of the relationship object that was saved into the "structure" array, is shown in the interface below.

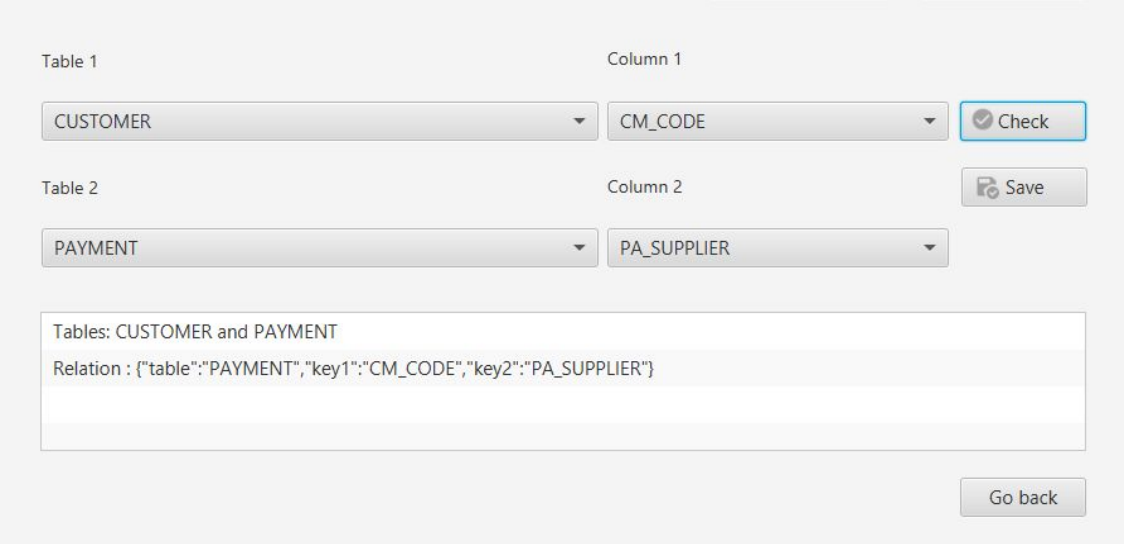

Fig. 3.5. Interface of database tool for relationship generation process

3.5.3. Data Visualization

Functionality to visualize data on to graphs was implemented using JavaFX chart libraries. 3 types of charts were used to graphically represent data; line charts, bar charts and stacked bar charts.

- Line chart - To visualize the flow of money; how the receipts and payments change with time
- Stacked bar chart - To visualize the frequency of transactions performed change with time
- Bar chart - To visualize the change of total of receipts and payments with time
- Bar chart - To visualize the change in bank receipts and payments with time

The time framed considered in the analysis was changed between weekly, monthly and yearly as requirements dictated.

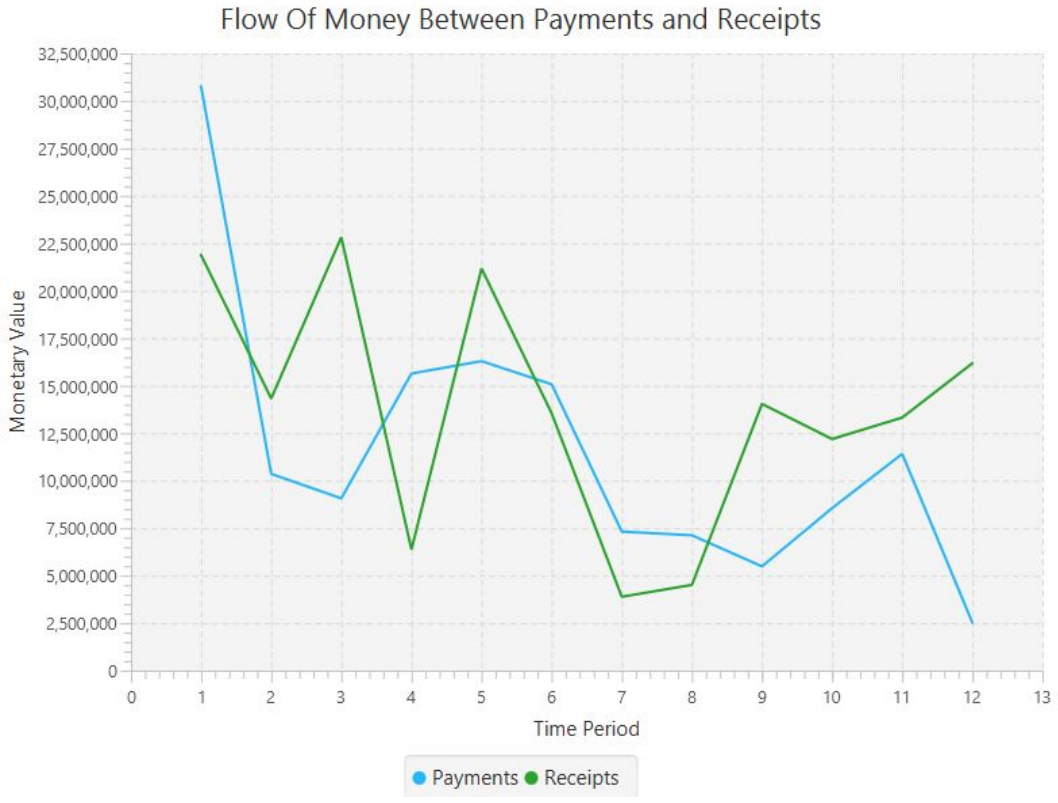

Fig. 3.6. Sample chart created using Database Analyser

Based on the relationships identified through the Database Analyser tool, using MS SQL Server Management Studio 18 different sql queries were executed to view data stored in tables and then identify anomalies as a part of the qualitative analysis. Parallely, Database Analyser

was also used to extract data when it was needed. Since some of the data in the database was in sinhala language, data was exported to csv format, translated and read manually as required during the analysis process.

3.6. Results Evaluation

3.6.1. Query Parser

The precision of the relationships generated using the Database analyser tool was evaluated by comparing them with the foreign key relationships derived using the sql queries that had been used to build the views of the database. Sql queries that had been used to create database views contained join conditions. Join conditions are usually used to match a primary key of a table to one of its foreign keys [50]. Based on this logic, by identifying the columns and the corresponding tables used in a join condition, foreign key relationships between tables were derived. These derived foreign key relationships were compared with the database relationships obtained through the Database analyser tool.

Sql queries that have been used to create views were used to compare and evaluate the table relationships obtained through the Database Analyser because it was the only place in the database that contained, a justifiable source (join queries) that could be used to recreate relationships that could have existed in the database.

This process of comparing and evaluating the precision of the table relationships generated using Database analyser tool was automated using the java program, “Query Parser” developed by the research team. Query parser captures sql queries stored in view_definition column in the INFORMATION_SCHEMA.VIEWS systems table of the database. Next, all SQL keywords and special characters included in each query were removed except JOIN keyword. Then all statements are iteratively split based on the logic shown in Fig. 3.7 and the relevant table relationships are obtained in the format show in Fig. 3.8.

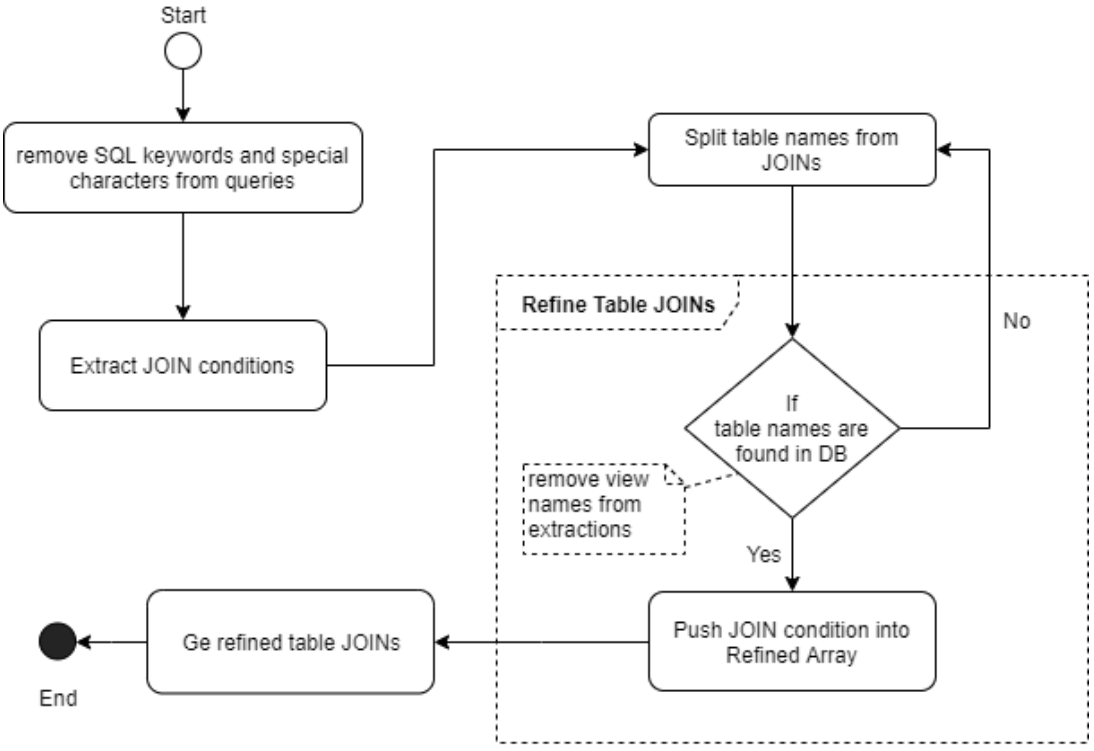

Fig. 3.7. Query parsing process used in the Query parser tool

```
[
        table1.column1 = table2.column2 ,
        table3.column1 = table4.column2
]
```

Fig. 3.8. Format of the relationship results obtained from query parser

Then the results obtained from the Query parser were compared with the relationships generated through the Database Analyser tool.

3.6.2. Interviews

After analysing the database dump, based on the information collected from the analysis an interview script was created and Interviews were conducted with employees of the bank. The intention was to identify whether the employees were aware of the anomalies or the possible information security issues present in the bank database we detected and if they have discovered

them evasive action taken. It was also intended to identify the control and monitoring mechanisms, security frameworks deployed, security principles and the security culture present in the institute.

Interview findings were used as a medium to evaluate the validity of the issues identified in the database analysis and identify the root causes behind those issues and recommend solutions to mitigate the risks involved with those vulnerabilities in terms of frameworks, best practices, rules or guidelines. Information obtained through the interviews were mainly qualitative data. Therefore the evaluation of the validity of the findings of the analysis phase was performed as a qualitative evaluation.

Chapter 4 - Results and Evaluation

This chapter describes the results including the evaluation results of this research. The aim of the research was to identify the information security issues financial institutions are facing at present and understand what are the actual reasons behind the existence of those security issues. Research took a database approach to answer the questions. A case study was carried out using a database dump of a Sri Lankan bank, database was analysed and potential information security issues existing in the system were identified. This chapter describes the results obtained through the analysis and the evaluation. Results are ordered as Direct and Indirect security issues.

4.1. Results

The results of the database analysis phase are a set of anomalies that were identified, that could potentially be classified as information security issues prevailing in the bank's information system. Here, the word anomalies is used to describe the deviations of the regulations and techniques adopted by the bank from standard practice or the guidelines set out by the CBSL, where CIA (Confidentiality, integrity, availability) of information is breached. Also it refers to the patterns of data that could be derived from the dataset that could potentially compromise CIA in information security.

4.1.1. Direct Information Security Issues

4.1.1.1. Users have NULL values as their password

Through the analysis it was identified, PASS table is used to store the usernames, passwords and privileges and other login information of each system user. 49 user records are available in the 'pass' table. Out of these 49 user records 23 of the records does not contain any value for the password (password column is null). This indicates that 46.9% of the user accounts does not require a password to login to their relevant accounts.

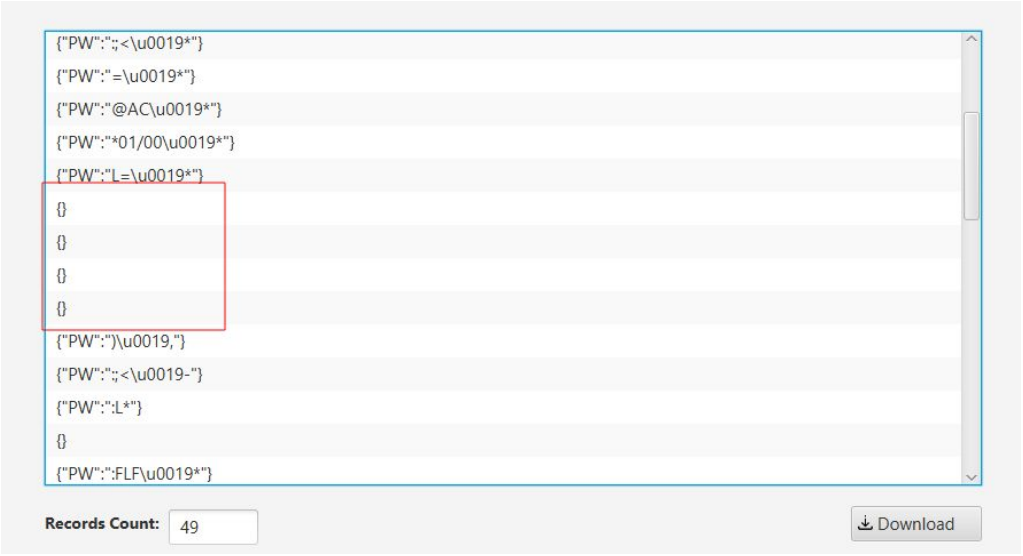

Fig. 4.1. Accounts with NULL passwords from PASS table

4.1.1.2. Session logout times and the authority level used are not recorded in certain user login sessions.

According to our analysis, the system stores user login sessions in a table called In_OUT table. In this table the time of login, time of logout, username, privilege level used during the login (ex: super user) are recorded. In the analysis it was identified that out of the total 7071 records available in the table, 2451 of the records did not contain a logout time (logout time is NULL). This amounts to 34.66% of the records.

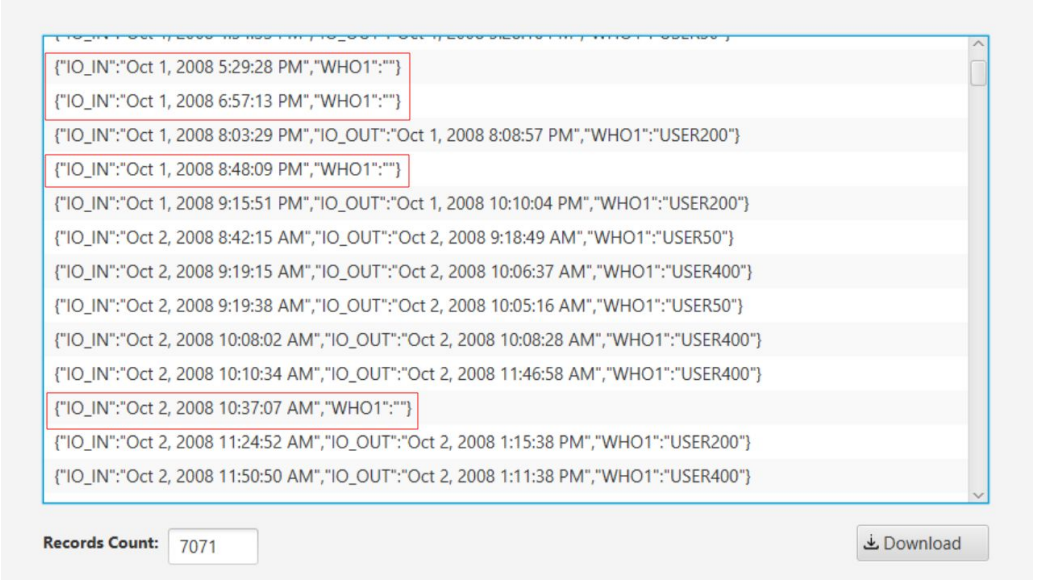

Fig. 4.2. Empty IO_OUT values

Interestingly it was observed that all records in the table have a login time entry, whether it had a log out time or not. Based on this observation it can be assumed that the culture in the bank does not stress on the importance of logging out of the system. It can also imply that the system used by the bank does not have the functionality to record a logout when the user is inactive for a stipulated time period or in an instance where the system shut down takes place.

In the table which records login sessions, logged in users are recorded twice: once the username is stored, then the privilege level of the user is also stored (super, user1 etc). In some records the value representing this privilege level of the logged in user is missing. This indicates a flaw in the logic behind the recording of login sessions that can be used to evade the session login stamp of a user.

Here again the hypothesis is made based on the observations made at the database level and no analysis was carried at the system level. In order to confirm the hypothesis a comprehensive system analysis must be carried out.

4.1.1.3. Relationships between tables can be recreated

Initial analysis showed that the database does not contain explicitly defined foreign key relationships. This can be attributed as a positive security measure[43]. Although foreign keys were not explicitly defined, researchers tried to identify the foreign key relationships between the tables.

As mentioned previously in the methodology, relationship finding functionality of the Database Analyser tool was used for this. Database Analyser uses an algorithm which checks the similarity between the contents of 2 columns of 2 corresponding tables based on a 80% rule and if the 80% threshold was obtained, tool suggests it as a possible relationship. Rationality of these suggested relationships were examined by the researchers and the most probable foreign key relationships were selected. Based on the results of this method, it was realised that even without a predefined set of foreign key relationships among tables, using a simple mechanism (that was used in the Database Analyser tool), relationships between tables could be identified.

As the researchers were successful in recreating relationships in the database using the logic implemented using the Database Analyser tool, complex and advanced methods to discover foreign key relationships such as machine learning methods[44] were not required to be used in this research.

Table relationships generated from database tool

Total of 33 tables were used for the process of recreating foreign key relationships between tables. The 2 main reasons for taking only a sample of 33 tables for recreating table relationships were: 1) the intention was to determine whether even if explicit foreign relationships are not defined in the database, can the implicit relationship be identified and it was proven that these implicit relationships can be identified. 2) because there was a large number of tables available in the the database, researchers did not focus on all the tables. Instead focused more on the tables that contained data that would have a direct impact on the customers and their transactions. A total of 30 unique relationships between these 33 tables were identified using the Database Analyser tool. Table 4.1 lists the 33 database tables used to recreate table relationships and Fig. 4.3 shows the list of table relations obtained (recreated).

FFROM_TTO	PA_EXPENCE	FIRST_LETTER
CREDIT	PA_DTAIL	DEPOSIT
INCOME_EXPENCE_LEDGER	RE_INCOME	REALIZE
CUST_CATEGORY	INCOME_EXPENCE	RE_DTAIL
CUSTOMER	LOAN_CARD_TYPE	CASH_WITHDRAW
INCOME_EXPENCE_SUB	SLEDGER	ACCOUNT
TR_DTAIL	CUSTOMER_SEX	INCOME_EXPENCE_LEDGER
LCODE	TR_TYPE	TAX
CR_DTAIL	RECEIPT	CUSTOMER_STATUS
CREDIT	DIVISION	CUST_CATEGORY
FT_DTAIL	PA_DTAIL PA_EXPENCE	PAYMENT

Table 4.1. List of database tables used to recreate table relationships

```
ACCOUNT.AC_CODE = CASH_WITHDRAW.CW_ACCOUNT
CUSTOMER.CM_CODE = FFROM_TTO.FT_FROMMEMBER
CUSTOMER.CM_CODE = FT_DTAIL.FD_TOMEMBER
CUSTOMER.CM_CODE = PAYMENT.PA_SUPPLIER
CUSTOMER.CM_CODE = RECEIPT.RE_CUSTOMER
CUSTOMER.CM_CODE = CR_DTAIL.CD_MEMBER
CUSTOMER.CM_CODE = INCOME_EXPENCE_SUB.IE_MEMBER
CUSTOMER.CM_CATEGORY = CUST_CATEGORY.CC_CODE
CUSTOMER.CM_STAT = CUSTOMER_STATUS.ID
CUSTOMER.CM_SEX = CUSTOMER_SEX.ID
CUST_CATEGORY_MAIN.CC_CODE = CUST_CATEGORY.CC_CODE
DEPOSIT.DP_CHQNO = RE_DTAIL.RD_CHEQUE2
DEPOSIT.DP_CHQNO = RE_DTAIL.RD_CHEQUE
DEPOSIT.DP_CHQNO = RE_INCOME.RI_CHEQUE
DEPOSIT.DP_CHQNO = REALIZE.RZ_CHQNO
RE_DTAIL.RD_CODE = DEPOSIT.DP_DSLIP
RE_DTAIL.RD_CODE = RECEIPT.RE_CODE
RE_INCOME.RI_CODE = RECEIPT.RE_CODE
INCOME_EXPENCE.IE_CODE = INCOME_EXPENCE_LEDGER.IE_CODE
TR_DTAIL.TD_LCODE = LCODE.LC_CODE
TR_DTAIL.TT_CODE = TR_TYPE.TD_CODE
SLEDGER.SG_LCODE = LCODE.LC_CODE
LCODE.LC_DIVISION = DIVISION.DI_CODE
INCOME_EXPENCE_SUB.IE_CODE = FIRST_LETTER.FL_LOAN_CODE
INCOME_EXPENCE_SUB.IE_CODE = TAX.TA_CODE
INCOME_EXPENCE_SUB.IE_CODE = PA_EXPENCE.PE_EXPENCE
INCOME_EXPENCE_SUB.IE_CODE = RE_INCOME.RI_INCOME
CREDIT.CR_CODE = CR_DTAIL.CD_CODE
PA_DTAIL.PD_CHEQUE = PA_EXPENCE.PE_CHEQUE
PAYMENT.PA_SUPPLIER = PA_EXPENCE.PE_CODE
```

Fig. 4.3. List of unique table relationships obtained from Database Analyser

The ability to identify table relationships allowed researchers to map between corresponding data stored across different tables and find out the different aggregate data patterns observable when data stored in different tables were combined. Information that was derived based on the table relationships identified are described below.

4.1.1.3.1. Fluctuation patterns of cash and money deposits of the bank can be obtained

Through the database analysis it was identified that information on all money receipts by the bank is recorded in a single table called RECEIPT table. By analysing this table, basic estimations on the amount of money received by the bank on a single day, week of the month, quarter, etc. can be made. Furthermore future money receipts can also be predicted through basic statistical analysis methods.

Same way it was identified that the PAYMENT table is used to record information on all payments made by the bank. Therefore total of all payment values represent the total payments made by the bank or in other words the total money outflow from the bank.

Combining the aggregate of all PAYMENT table values and the aggregate of all RECEIPT table values an estimation on the total money available in the bank (deposits) at a given day, week of the month, month or quarter can be obtained. These aggregate values were graphically represented as bar graphs using the visualization functionality available in Database Analyser tool. Visual representations provided better insights on the fluctuation patterns of the total deposits in the bank. Further through statistical models future predictions on the money available with the bank on a given future instance can also be calculated. This effectively provides an attacker with sufficient insights on the optimal days to carry out a digital or a physical fraud on the banks financial assets. (ex: days, quarter, or the week where there is a greater amount of liquidity/cash in hand)

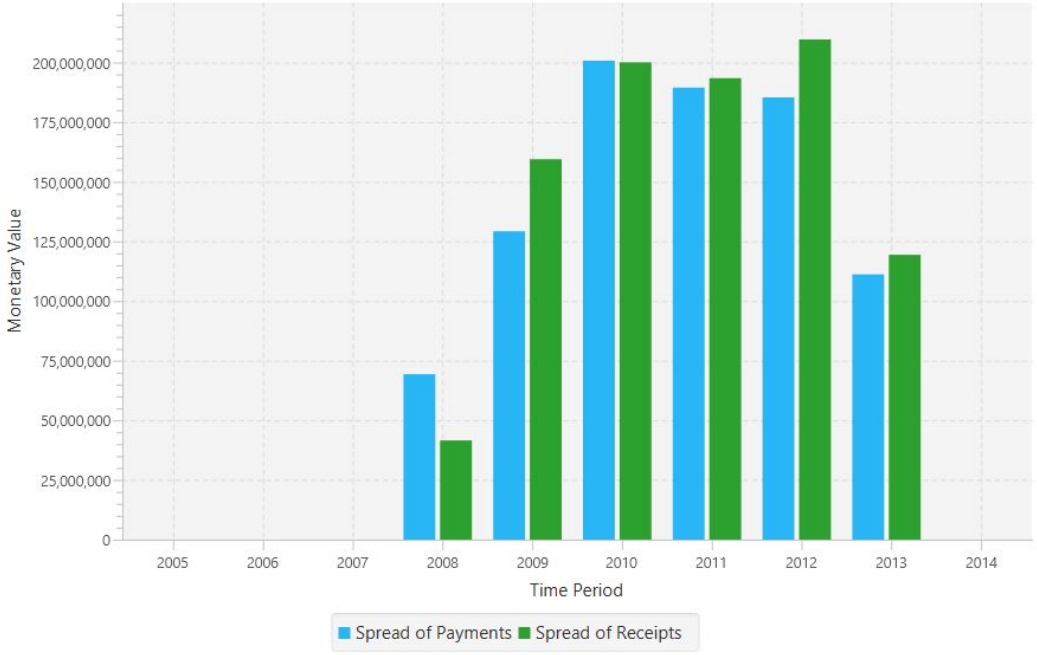

Fig. 4.4. Total receipts and payments by each year in bank

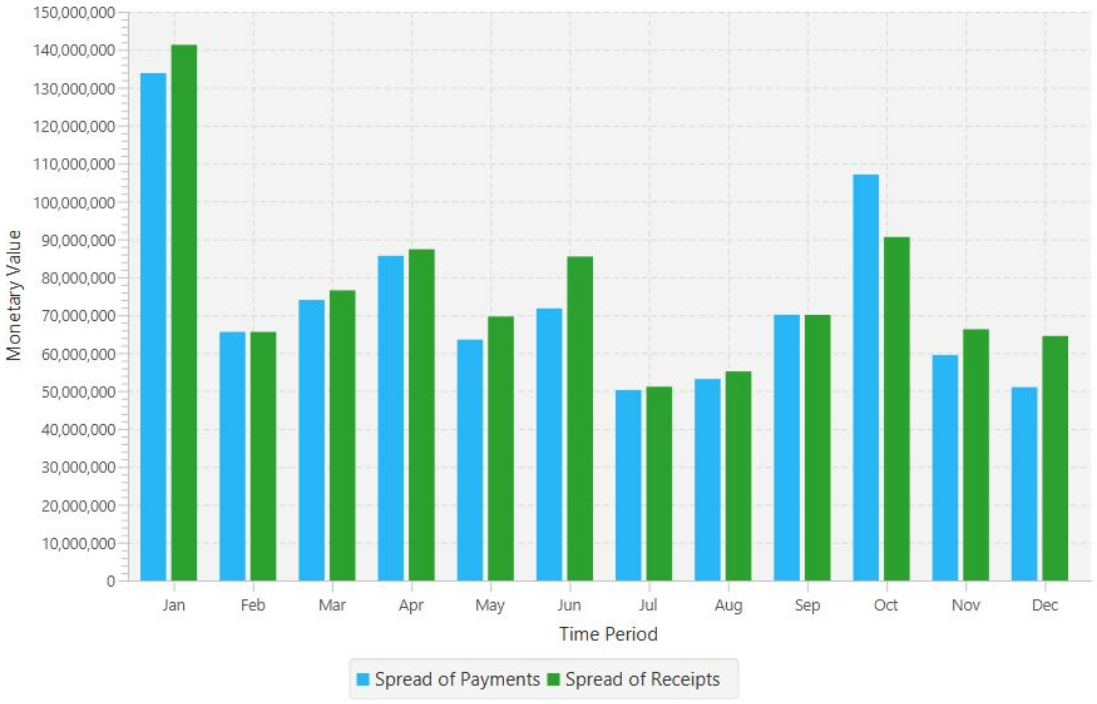

Fig. 4.5. Total receipts and payments by each month in bank

For example, Fig. 4.4 shows the total payments and receipts of the bank by year. It can be determined that the bank was not doing well in terms of liquidity in their financial assets based on the fact that total payments are higher than the receipts. Further it can also be said that liquidity wise bank had gone through tighter periods during the years 2010 and 2011.

Fig. 4.5. represents the total receipts and total payments of the bank for the year 2013 arranged by month. Analysing the graph it can be determined that January is the month the bank has handled the highest amount of money. It is also identifiable that generally the month following a peak in the amount of money handled there is a drastic reduction in the amount of money handled. Months January and June are the months with the highest surplus of money in the bank. Therefore an attack on the bank may yield the highest reward for an attacker in such a month.

4.1.1.3.2. Transaction and behavioural profile of a bank account holder can be created

In the database analysis it was found out that a single transaction made by a customer might be stored in several tables. This is because the transaction made by the customer may have been done to compensate for an earlier transaction done by the customer and also because of the dual nature of a transaction (every transaction results in debit as well as credit). For example a

customer may deposit money in the bank as a partial payment payment of a loan. This transaction will be recorded in the receipts table, income expense table, income table, and loan payment account as well. These would result in an information security loophole that can be used to map out all the transactions corresponding to a particular CUSTOMER_ID. This effectively enables the creation of a transactional history profile for a selected customer. This can be used as a method to target a potential victim for an attack.

The relationship between different entries corresponding to the same transactions are generally made using the CUSTOMER_ID and a unique TRANSACTION_ID. In the database, CUSTOMER table is used to store personal details of customers of the bank. This information is stored in plain text. So using the database analyzer tool, it is possible to track the trail of every transaction back to a single account holder using the CUSTOMER_ID. By utilizing these capabilities, high profile customers with large funds in their accounts as well as customers who transact with the bank in large volumes can be profiled as their personal details are also plainly available in the CUSTOMER table.

Any dormant accounts with no transaction history can be identified through the database analyzer findings, as a result any insider with adequate access and permissions could change the status of these dormant accounts to active and carry out illicit transactions such as money laundering and siphoning available funds to external accounts outside the country to go undetected [46].

The customer related to Fig. 4.6, Fig. 4.7, Fig. 4.8, Fig. 4.9 figures, has done 161 payments and 176 receipts within the period of 2008-2013. All these transactions have been plotted in the following charts by each month of the year. 2013, 2012, 2011 and 2010 years are taken for chart visualizations. When monetary values are considered, individual transactions of this specific customer exceed Rs. 50,000 on average. According to the following illustration, in 2013, this specific customer has not conducted that many transactions with the bank, in fact his deposits to the bank have come to an all time low. But in 2012, this customer has received around Rs. 2.5million into the account altogether.

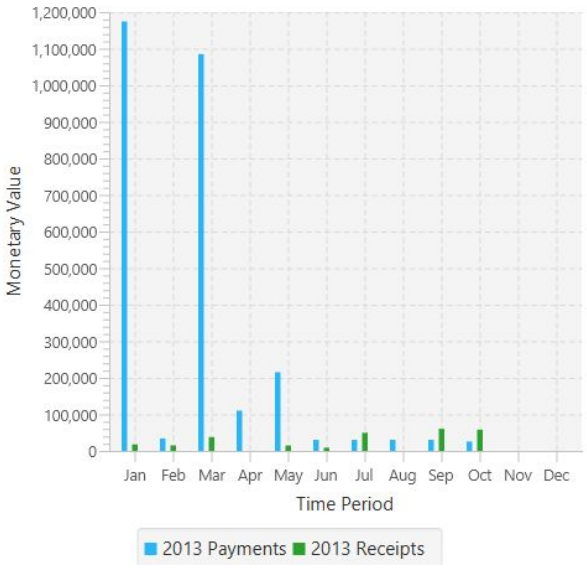

Fig. 4.6. 2013 receipts & payments

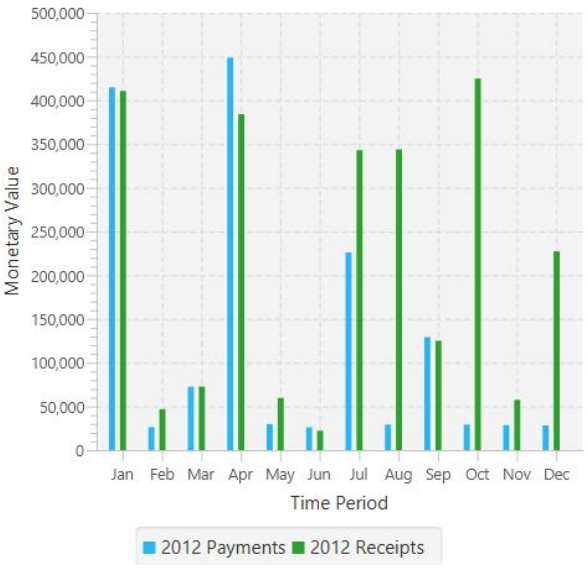

Fig. 4.7. 2012 receipts & payments

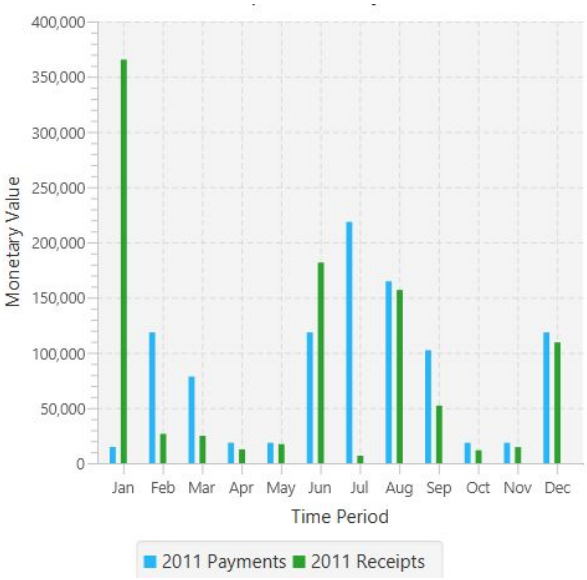

Fig. 4.8. 2011 receipts & payments

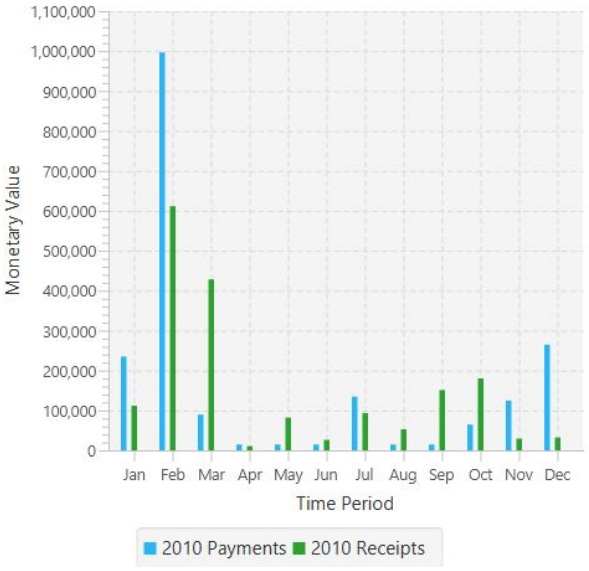

Fig. 4.9. 2010 receipts & payments

4.1.2. Indirect Information Security Issues

Under indirect information security issues, anomalies which were identified from the analysis, but do not pose a direct threat to the CIA triad in information security have been described.

4.1.2.1. Essential data columns contain empty values

Out of the total of 2432 customers in the database 924 of the customer records does not contain any NIC number value. This amount refers to 38% of all customer records. NIC is generally a mandatory information that must be collected from customers (account holders) by a financial institution[49]. NIC number is generally used by banks as proof of an individual's identity. But in this bank for 38% of the customer records, this vital information is missing.

It was also identified that the NIC number is not stored in the database in a standard format. Most of the time the final letter 'v' was stored in simple letters while in a few other records the final letter 'v' was in capital letters. Further although most of the time 'v' was stored adjacent to the rest of the numbers, in records were found where there was a 'space' separating v from the rest of the numbers as shown in Fig.4.10.

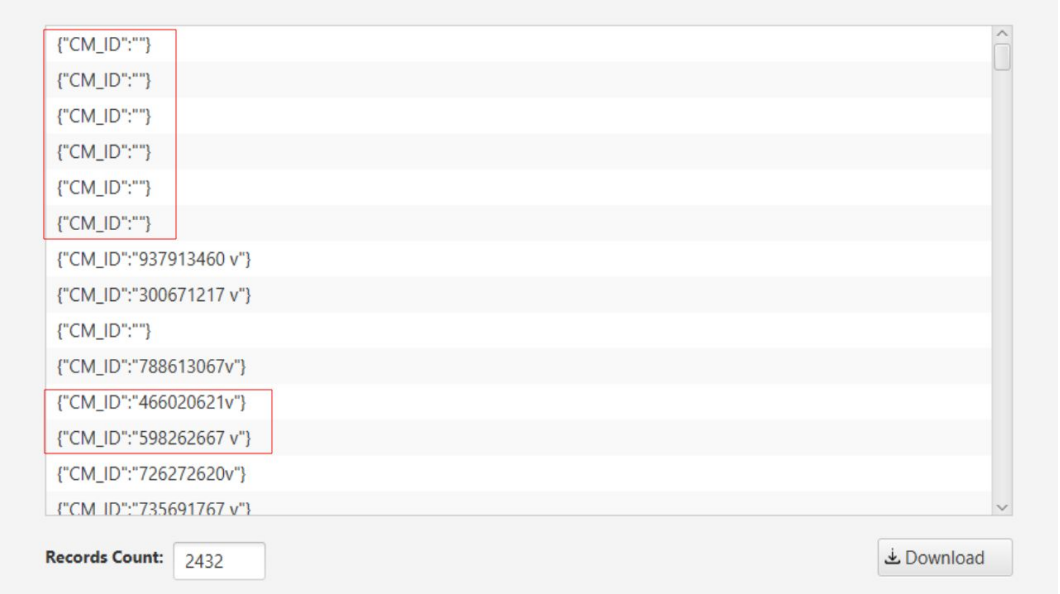

Fig. 4.10. Empty NIC records and NIC stored in different formats in CUSTOMER table

Researchers found out that in the CUSTOMER table the primary key is the account number. Therefore if a customer has 2 accounts with the bank, details of the customer is saved under both account numbers.

Further, when the date of birth of the customers were analysed, it was identified that in some records where the same NIC number values were present, different date of birth values were present. Since NIC numbers are unique value for everyone, there can not be duplicate values on date of birth for any NIC. Fig. 4.11. shows several instances where the date of birth representation in the NIC number does not match with the recorded date of birth value in the customer table.

This information represents that in the information system used by the bank, there are loopholes in the validation process of customers. It also indicates that internal employees of the bank are not adequately familiar with the information system used and lack the knowledge on the standard conventions that must be followed in day to day operations (ex: correct format to enter NIC number). These vulnerabilities in the system can create grey areas that could be exploited to propagate insider attacks on the system, which constitutes the greatest amount of attacks on financial institutions[28].

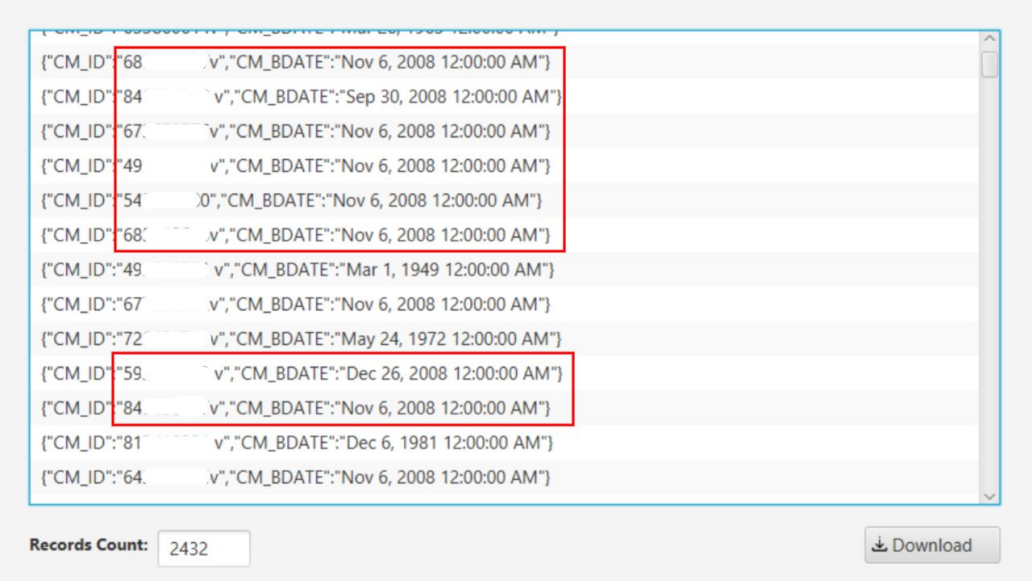

Fig. 4.11. Records with mismatching NIC numbers and date of birth values

4.1.2.2. Primary keys contain two formats

Through the analysis it was identified that in most of the tables that record transactional information, the format of the primary key has undergone a sudden change, starting from a specific time period (dates not disclosed due to privacy concerns). Based on the observation that the change is visible across the whole database starting from a given time period, the change is most likely due to a change in the whole system or a major upgrade in the system. Bulk insertion

operations that have been carried out during this period which provide evidence to confirm this assumption. Fig. 4.13. shows part of the bulk insertion operation carried out on 2008-09-30. Before upgrading a new system, necessary precautionary measures must be taken when converting old data to the new formats required by the new system. Fig.4.12 shows the PAYMENT table which has the primary key - PA_CODE in two formats.

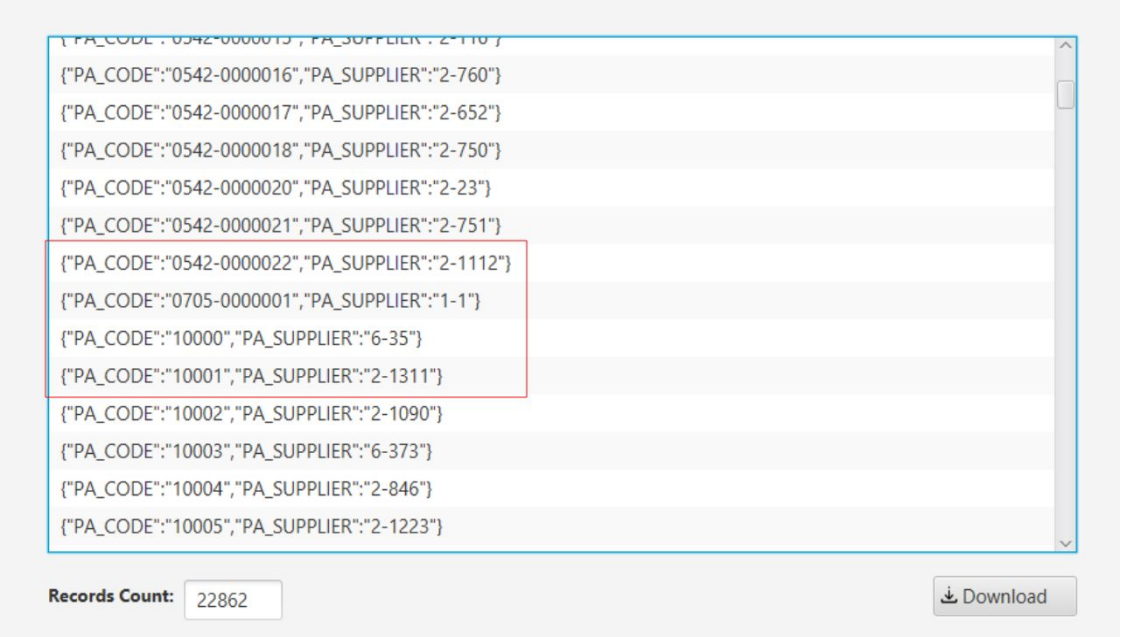

Fig.4.12. The two formats of the primary key present in the PAYMENT table

Results | Messages

	PA_CODE	PA_SUPPLIER	PA_DATE	WHO1	WHEN1	PA_CASH	PA_CHE
1	0521-0000015	2-74	2008-09-12 00:00:00	USER50	2008-09-30 07:14:00	0.00	0.00
2	0521-0000034	2-341	2008-09-03 00:00:00	USER50	2008-09-30 07:28:00	0.00	0.00
3	0521-0000041	2-175	2008-09-30 00:00:00	USER50	2008-09-30 07:29:00	0.00	0.00
4	0521-0000046	2-264	2008-09-29 00:00:00	USER50	2008-09-30 07:30:00	0.00	0.00
5	0521-0000047	2-94	2008-08-27 00:00:00	USER50	2008-09-30 07:31:00	0.00	0.00
6	0521-0000052	2-98	2008-08-06 00:00:00	USER50	2008-09-30 07:32:00	0.00	0.00
7	0521-0000054	2-100	2008-09-30 00:00:00	USER50	2008-09-30 07:33:00	0.00	0.00
8	0521-0000056	2-251	2008-08-08 00:00:00	USER50	2008-09-30 07:34:00	0.00	0.00
9	0521-0000059	2-120	2008-03-27 00:00:00	USER50	2008-09-30 07:35:00	0.00	0.00
10	0521-0000062	2-145	2007-07-28 00:00:00	USER50	2008-09-30 07:38:00	0.00	0.00
11	0521-0000075	2-277	2008-09-02 00:00:00	USER50	2008-09-30 07:43:00	0.00	0.00
12	0521-0000080	2-291	2008-09-30 00:00:00	USER50	2008-09-30 07:45:00	0.00	0.00
13	0521-0000082	2-292	2008-09-19 00:00:00	USER50	2008-09-30 07:45:00	0.00	0.00
14	0521-0000085	2-244	2008-09-10 00:00:00	USER50	2008-09-30 07:46:00	0.00	0.00
15	0521-0000100	2-359	2008-09-29 00:00:00	USER50	2008-09-30 07:47:00	0.00	0.00
16	0521-0000101	2-287	2008-09-29 00:00:00	USER50	2008-09-30 07:48:00	0.00	0.00
17	0521-0000110	2-419	2008-09-25 00:00:00	USER50	2008-09-30 07:49:00	0.00	0.00
18	0521-0000014	2-414	2008-09-10 00:00:00	USER50	2008-09-30 07:51:00	0.00	0.00
19	0521-0000186	2-597	2008-06-19 00:00:00	USER50	2008-09-30 07:54:00	0.00	0.00
20	0521-0000205	2-649	2008-09-25 00:00:00	USER50	2008-09-30 07:55:00	0.00	0.00
21	0521-0000225	2-266	2008-06-06 00:00:00	USER50	2008-09-30 07:56:00	0.00	0.00
22	0521-0000238	2-512	2008-08-08 00:00:00	USER50	2008-09-30 07:56:00	0.00	0.00
23	0521-0000343	2-887	2008-09-27 00:00:00	USER50	2008-09-30 07:58:00	0.00	0.00
24	0521-0000482	2-569	2008-09-08 00:00:00	USER50	2008-09-30 07:59:00	0.00	0.00
25	0520-0000486	2-881	2008-09-22 00:00:00	USER50	2008-09-30 08:01:00	0.00	0.00
26	0520-0000492	2-1125	2008-05-29 00:00:00	USER50	2008-09-30 08:02:00	0.00	0.00
27	0520-0000507	2-1198	2008-08-25 00:00:00	USER50	2008-09-30 08:03:00	0.00	0.00
28	0520-0000515	2-1170	2008-09-21 00:00:00	USER50	2008-09-30 08:04:00	0.00	0.00
29	0513-0000002	2-3	2008-08-11 00:00:00	USER50	2008-09-30 08:33:00	0.00	0.00
30	0513-0000003	2-23	2008-09-30 00:00:00	USER50	2008-09-30 08:34:00	0.00	0.00
31	0513-0000005	2-13	2008-09-30 00:00:00	USER50	2008-09-30 08:36:00	0.00	0.00
32	0513-0000009	2-175	2008-09-30 00:00:00	USER50	2008-09-30 08:37:00	0.00	0.00
33	0513-0000010	2-87	2008-09-30 00:00:00	USER50	2008-09-30 08:38:00	0.00	0.00

Fig. 4.13. Bulk insertion operations carried on 2008-09-30

4.1.2.3. Two date attributes for the same record

Each table in the database that records transactions has two columns for storing ‘date’ values. But only one of these columns stores time along with the date, while the other column is storing only date values. In most of the tables the column that stored both date and time was WHEN1. This results in data redundancy. It is most likely that the system transition may have caused this issue to exist. It is mostly likely that the older software system used by the bank has only captured date value and the new system has introduced WHEN1 column to capture both time and date values.

For transactions that have taken place before the system transition, date value stored in WHEN1 column and the other date column is different. Further analysis showed that this has happened because during the bulk transfers of data that have been carried out during the system

transition, the WHEN1 column has been used to store the ‘date and time’ the data transfer from old system to current system has taken place. The other date column has been used to store the original date the transaction has occurred. During the time after the system transfer both date columns are used to record the same date: WHEN1 including time and the other without time.

Because for the transactions that have been carried before the system transition WHEN1 and the other date column records 2 different date values, unless the system has been configured correctly to fetch the correct date column for records before and after the system transition it can result in ambiguity when it comes to querying for historical transactions. Fig. 4.14. shows a list of records in the PAYMENT table with the 2 columns used to represent date: WHEN1 and PA_DATE.

If something fraudulent happens, date is important since the database actions are monitored through system LOGS arranged by dates. As per the findings of the research, WHEN1 is the column to store date with time and the other is only storing date. If dates are not correct and accurately saved, the possibility of tracking the damage is reduced. Insiders who do the damage can be escaped from the easiest way of tracking them. There are sets of records which have the same date without time. There can be fraudulent transactions present during the time of transition of this system because it appears that bulk insertions of data have been carried out during the transition period. Date columns may have been replaced during this process.

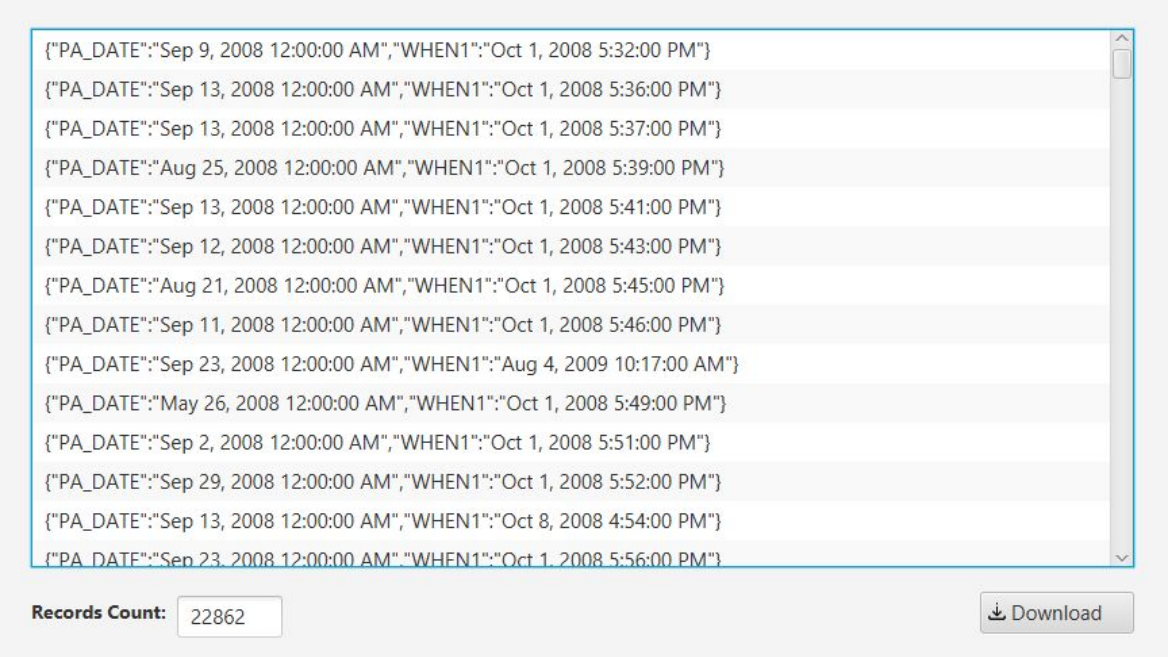

Fig. 4.14. Two date columns with time and without time in PAYMENT table

4.2. Evaluation

4.2.1. Evaluating the precision of table relationships generated using the Database Analyser tool

All the relationships generated using the Database Analyser tool (using 80%-20% rule) were evaluated for precision using the sql queries that had been used to build the views of the database (sql queries of views). These sql queries had been stored in the view_definition column in the INFORMATION_SCHEMA.VIEWS systems table of the database. Sql join conditions are generally used to match a primary key of a table to one of its foreign keys[50]. Based on this logic a list of foreign key relationships were derived. Relationships generated using the database analyser was then compared with the derived list of foreign key relationships.

The process of deriving foreign key relationships from sql queries of views and comparing them with the relationships generated using the Database analyser was automated using the tool Query Parser as it was mentioned in the Methodology chapter(sub topic 3.6.1. Query parser). The following result was obtained when the evaluation was done using the query parser.

When the 30, table relationships generated using the Database analyser were compared with the relationships derived from the sql queries of views, 24 matches and 6 miss matches were obtained. This results indicate that at a precision level of 80%, researchers were able to generate table relationships for this database using the Database Analyser tool.

4.2.2. Interview Findings

Interviews were conducted with the employees of the Bank, whose database on which the case study of this research was carried out. Interviews were conducted as structured, face to face interviews.

The intention was to identify whether the employees were aware of the anomalies or the possible information security issues present in the bank database we detected and if they have discovered them evasive action taken. It was also intended to identify the control and monitoring mechanisms, security frameworks deployed, security principles and the security culture present in the institute. Interview findings were also intended to use as a medium to evaluate the validity of the issues identified in the database analysis and identify the root causes behind those issues.

Through the interviews mainly qualitative information was collected. Due to the privacy concerns and the policies of the bank, certain information was withheld from sharing with the researchers. Below table lists down the findings from the interviews with the information security issue or the anomaly they were related to.

Issue	Findings from the interview related to the issue
Users have NULL values as their password	• Public key/private key encryption is present • MD5 is used for SLIPS • No password sharing is encouraged and therefore no password sharing • Password validation is done:the 3 most recent passwords used can not be used as the new • Regex check is present:a criteria specified by the bank has to be used when creating the password (ex: at least one capital letter, simple letter, special character, and a number must be used when creating the password)
Log out times and the authority level used in certain user login sessions are not recorded in the database	• A log system is available to monitor and keep track of user activities • Hash lock system is used to control transaction communications with the database.
Relationships between tables can be recreated - Fluctuation patterns of cash and money deposits of the bank can be obtained - Transaction and behavioural profile of a bank account holder can be created	• No information could be collected (Bank refused share information due to bank policies)

Important and essential data columns contain empty values	• No information could be collected (Bank refused to share information due to bank policies)
Primary keys contain two formats	• There has been a major system transition in the bank • Bulk importing of data to the new database using sql queries have been carried out during the transition from the old system to the current system • There had often been mismatches between the new format and the old format when transferring data from the old system to the new system • Some issues had been found which could not be found reconciled in the new system, those data were entered again into the older system in the format compatible with the new system and those data had been transferred to the newly added system.
Two date attributes for the same record	

Table 4.2: Interview results

Chapter 5 - Discussion

5.1. Overview of issues identified

Insiders in banks have the potential to be involved with an attack easily since they know the background and environment well. The research team considered a bank as the case study to find the vulnerabilities existing in the backend systems and processes which can be exploited by authorized insiders. It is evident that the identified issues during the database analysis depict that there's a considerable threat to the sensitive information stored in the bank database answering the first research question. The study showed that there are both major and minor issues leading to breach of information security.

During the analysis of the database, the research team introduced a novel approach to establish relationships between table columns by checking contents for similarity. It has not been used for SQL server databases before this attempt, in this research domain. In addition to that, newly introduced Query Parser can be used on SQL Server databases to derive foreign key relationships from the join conditions from a database which consists of views which are already created. SQL database structure was written into a JSON script that can hold the relationships at a single place. That was the source data to implement the Database Analyser tool.

Through extensive analysis the researchers presented the issues that were discovered from the database analysis tool in the results section. Initial focus is on the issues which pose a direct threat to information security of the bank's information system. From the list of issues that were listed out, it was clearly evident that every issue directly breaches the CIA principles of information security, and the vulnerabilities these issues give rise to, have far greater impact on the information security of the bank and to its financial assets at large compared to the issues which are identified later on as indirect issues. The impact and the extent of the damage that can be caused by these direct issues are discussed below.

One of the issues identified was, 'missing values (NULL) in the password attribute' in the database. Passwords are one of the primary forms of protection to a system and its resources from unauthorized access. Constructing secure passwords and ensuring proper password management are essential. Poor password management and construction can allow both the dissemination of information to undesirable parties and unauthorized access to a system's resources [51]. Even an unprofessional attacker can easily compromise poorly chosen passwords.

So in the case of this particular scenario, when there are empty values (NULL) for the password attribute, it is possible to assume that certain users are created in the system without valid passwords, which implies for the same users, a successful login can be initiated by just entering the username. So, if an insider who has necessary permissions and access to the backend of the system (i.e. the database), the usernames can easily be viewed and the fact that corresponding password values are NULL can also be discovered. So the said insider can easily log into the banking system as a system user without entering a password. This amounts to breaching the integrity of the system thereby breaching the CIA triad of information security. However, the above hypothesis is based on the observations made at the database level and not on the system level, so in order to confirm the hypothesis, a comprehensive system analysis must be carried out.

With the help of the tool, as explained in the results section, the researchers were able to recreate the relationships between tables. As a result, the transaction history and pattern of bank's customers can be easily discovered. Also it was identified that bank's customers' personal details are contained in the CUSTOMER table including their full name, address and NIC numbers. These are considered to be very confidential and sensitive information of individuals. It is quite evident and obvious that no means of encryption were used to mask these sensitive attributes. This fact alone indicates the extent to which the confidentiality of the CIA triad has been breached as any one with access to the backend of the system or even a database dump, can easily take note of the sensitive information of customers that is readily available to view in the database. The types of attacks an insider could manoeuvre from those highly sensitive information are vast in number and greater in impact.

Similarly, the analyzer tool was able to visualize the transaction patterns of customers in terms of their deposits and withdrawals with the bank. This would shed some light into the amount of wealth a specific customer has saved in the bank as well as the volume of the transactions of the said customer. So if an insider can get these insights about the bank's customers, the customer will be vulnerable as their confidential details are exposed without their consent.

Since both personally identifiable information as well as confidential information (such as approximate account balance and the transaction pattern) of the customer is exposed, an expert can put together this information to carefully profile each and every customer down to their residential address. This combined information about the bank's customers can be used to tailor make social engineering attacks of fraudulent nature to exploit the vulnerabilities of

individual customers and maybe even used for identity theft [52]. Furthermore, this profiled information about bank's customers can be sold to marketing agencies, insurance agencies, other banks and financial institutions, etc. for substantial monetary rewards. This is the ultimate breach of confidentiality as per the CIA triad of information security

Similar to visualizing customer transaction patterns, the bank's total receipts and payments can also be populated into a bar graph. By visualizing such information, an insider can predict the days, weeks or months in which the bank handles the most number of deposits. This information can be further taken up to plan a coordinated attack on the bank on a day, a week or a month in which the bank has the highest liquidity. This would be a potential vulnerability that could be exploited by an attacker.

5.2. Challenges faced during the research

During the research, there were various challenges researchers had to face: for some the researchers were able to find answers and for few problems, researchers failed to come up with an answer. All the difficulties faced are discussed below.

5.1.1. Challenges faced in reading LDF files

LDFs of the database were initially tried to read using the 3rd party tool called SQL Log Rescue by RedGate. But the log records which were in 'blob' format could not be read using this tool and the Log Rescue did not provide assistance in converting logs into a more usable format such as csv. Therefore two MS SQL Server built in functions namely; fn_dblog and fn_dump_dblog, and some third party proprietary tools like ApexSQL were tried in order to read the log contents. But it was unsuccessful due to 4 reasons: 1) inability to convert the log data into a desired format and couldn't read blob type records, 2) incompatibility of the LDF of the bank database with MS SQL server versions that provided built in functions to read LDF files , 3) lack of support given to older versions of SQL server databases 3rd parties, 4) stored procedures used by 3rd parties that had been used read the log files were in '.dll' format and could not be decoded. Therefore although it was initially intended to use LDF content as well for the analysis, it had to be achieved.

5.1.2. Challenges faced in relationship creating process implementation

There were 320 empty tables out of all the tables. Those empty tables were rejected in this process since we can not check the contents of columns in empty tables. But there can be a relationship between two tables even if one of them is empty. The below mentioned relationship is an instance where empty tables are present.

E.g.: ACCOUNT and CASH_WITHDRAW tables

Here, CASH_WITHDRAW table was empty, but it's connected with ACCOUNT table

There was less content in some tables which did not satisfy the 80% threshold when its columns were compared with another table. But there was a relationship between the two tables.

E.g. : ACCOUNT table has 1 record

BANK table has 6 records

Tool recorded as there's no relationship, but an actual connection was there between these two tables.

5.1.3. Challenges faced in implementing the query parser

Database query parser was not 100% perfect. So, in 146 results came out from parser, there were a set of wrongly parsed joins due to the naming conventions and high degree of dynamic nature in queries. They were removed after the refining process shown in Fig. 3.7.

In database views creation queries, there were JOI conditions in a format which were not the same when compared to the relationship JOIN included in JSON file.

E.g.

1. Json - CASH_WITHDRAW.CW_ACCOUNT = ACCOUNT.AC_CODE
2. Query - ACCOUNT.AC_CODE = CASH_WITHDRAW.CW_ACCOUNT

Sometimes, when query is parsed, only the format in (2) is found within query and not the format (1). But the relationship is existing both ways according to the SQL join conditions. When extracted join is compared with tool joins, it's not counted as a matched record since it's not in format of (1). But if one join is found in format (1), it should be in format (2) also due to the existence of a relationship.

5.3. Reasons for the identified issues

In discussing cause and effect of the discovered vulnerabilities, the researchers were able to pinpoint the reasons for the existence of such loopholes and vulnerabilities directly by assessing the database and the results of the database analyzer tool.

One main reason identified for the existence of security issues in the studied bank is, Organizational culture in the bank is not stressing on the importance of information security. Researchers identified, this is partly responsible for the presence of null values for passwords and for having user sessions with no log out times. Researchers assumed that in an organizational culture which promotes information security positively employees would be extra cautious when it comes to information security and whether password policies are present or not would always use a strong password for their account. It was also assumed that employees would always be mindful of the risk of unauthorised access, if log out procedures are not followed as prescribed and would always stick to those log out procedures minimizing the number of unintentional log outs missed.

Validations methods implemented by the bank do not meet the basic security standards is the second reason identified for the existence of security issues in the studied bank. This was also identified, as partly responsible for the presence of null values as passwords and for having user sessions with no log out times. Validations are a must for any given system and validations are implemented to satisfy 2 main objectives: prohibit activities intentionally performed by users, outside the desired procedures and to prohibit unintentional mistakes made by the users. When validations implemented are not upto the standard, the system may not notify the users that they have not logged out or, may not notify the users when data inserted is in the wrong format, that may lead to intentional as well as unintentional insider attacks.

But the information gathered from the interviews did not support the root causes researchers identified based on the evidence collected at the database analysis. In fact, for some security issues, such as the presence of null password values, the information gathered from the interviews contradict the observations made at database analysis. For almost all the other security issues as well, the information received from the interviews were inadequate to make a reasonable decision on the underlying root causes for the existence of those information security issues.

Therefore researchers suggest that a database level study does not provide the required level of in depth analysis to understand the real reasons behind the existence of these loopholes

within the system and a complete system wide study must be carried out to fully understand the real reasons behind the existence of these loopholes within the information system of the bank.

5.4. Recommended mitigation techniques for information security issues

5.4.1. Application Validation

Whole banking system should be built well with clearly defined validations for the application side. It reduces the inconsistencies available in the database such as NULL values in passwords and format errors in NIC. Any insider can not inset malicious values into the system if there's a proper validation in the front-end and back-end of the system. The easiest way to do validation is input validation. Input should be tested for missing data, field length, class, range and invalid values. In addition to that, performing cross reference checks, comparing values with stored data before saving and setting up self validating codes(digits check) are some advanced validations[42].

5.4.2. Use strong encryption mechanisms for passwords

In the company system SLIPS, MD5 is used as the encryption algorithm as it was revealed in the interview. Nowadays it is considered as a less secure encryption mechanism since brute force attacks are fast due to the advancement in computer technologies[47]. Lookup tables for MD5 hashes have been rapidly grown allowing users to find common passwords easily. So, encryption policies should be updated with strong encryption algorithms. As an example, instead of using even SHA256 directly, using a salt with SHA256 makes it more difficult to be broken by an attacker[48].

5.4.3. Oracle Database Vault

This is a better alternative to be used when information security is vital to an organization. It enables separations of duties and supports for privileged database access user controls protecting database, data and applications. Oracle Database Vault[27] can restrict access to specific areas in an Oracle database including users who have administrative access. For example, we can restrict administrative access to employee salaries, customer medical records, or other sensitive information. Oracle confirms that using Oracle Database Vault, data can be protected against insider threats, meet regulatory compliance requirements and enforce

separation of duty. Oracle vault includes realms, factors, rules and command rules and using them Oracle ensures that data can be protected against insider threats[27]. Combining rules and factors, conditions can be established which commands in the database are allowed to execute by users. To adopt this for the bank, they have to move RDBMS from SQL Server to Oracle.

5.5. Conclusion

With the limited data resources that were available through the database dump and the limited knowledge on the processes and the systems of the bank, the research team was able to derive the relationships among the tables and there by infiltrating a considerable amount of information about the bank, its customers and transactions. Considering the intimate knowledge an insider of the bank possesses about the banks internal process and systems, as well as with the higher level of access to the system, the extent to which the information system can be compromised is far more serious and imminent compared to the extent the research team was able to compromise. Since this research was done within the Sri Lankan context and the bank we gathered data from isnt a well established bank within Sri Lanka, hence the security controls are not up to the required level and many vulnerabilities exist in the financial information system of the bank. The potential for an insider to breach the information security of the bank via the database is high. Therefore the bank should look into the root causes behind those vulnerabilities and implement adequate controls to mitigate the risks involved with those vulnerabilities.

In the process of answering the research question, “Are there threats and vulnerabilities exposed to insider attacks, in systems and processes of Financial Institutions in Sri Lanka?”, this research could derive information security issues already existing in the database taken for the case study. When the second research question was considered, there was a lack of information obtained from the interviews to provide exact reasons for the issues. But after the database analysis, a set of main reasons were identified which have been mentioned under the topic 5.3 in this chapter itself. So, this work will open doors to critically analyze and find the reasons for information security loopholes in the back-end aspects like databases in information systems in banks.

The Database Analyser tool created by the research team can contribute to the community as a tool to create relationships between 2 tables where foreign key constraints have been not established. JSON based relationships structure definition was also a novel approach to be used along with the Database Analyser tool to store relationships. The research team believes

that this work will motivate the research community to design and implement more tools on analyzing SQL format databases on various conditions.

5.6. Future Work

This research is based on one database obtained from a reputed bank in Sri Lanka. This can be extended more to obtain data from several banks and perform database analysis. Then more issues can be identified and find the reasons according to multiple banking environments. A generalized solution can be suggested for risk mitigation, as a framework or model. With the limited time frame that was not a goal for the research team. More interview sessions can be conducted and gain the real understanding on banking policies focused on information security.

As a future work, the Database Analyser tool can be improved in several ways. Currently it's running on a single threaded environment. It can be extended to use a multithreading environment with Java to obtain higher efficiency. After finding the patterns in the data, if the tool can give hints or suggestions about the vulnerabilities existing, it will be a real use for this kind of scenario.

It is possible to improve the accuracy of the query parser since it's not 100% perfect due to the high degree of dynamic nature in SQL queries used for creating views. But still it extracts joins from complex queries to some extent correctly. Some queries in this database included multiple nested queries making the query complex. Parser implemented by the research team could cater to the majority of the queries for extracting joins successfully. To convert this query parser as a universal tool, more parsing mechanisms can be added.

For the anomalies mentioned under indirect information security issues in the result section researchers were unsuccessful in finding direct vulnerabilities or, threats posed by these anomalies. But further research using advanced analysis techniques such as machine learning, may reveal security implications caused by these anomalies as well.

In this research, researchers were unable to utilize the log files of the database during the analysis phase due to the technical challenges the team had to face. This research can be extended to include the log files as well in the analysis. Log files provide data on the time wise changes in the database providing with the capability to perform time series analysis on the changes in the database.

References

[1] M. Ula, Z. Ismail and Z. Sidek, "A Framework for the Governance of Information Security in Banking System", Journal of Information Assurance & Cybersecurity, pp. 1-12, 2011.

[2] "Banking Sector | Central Bank of Sri Lanka", Cbsl.gov.lk, 2019. [Online]. Available: https://www.cbsl.gov.lk/en/financial-system/financial-system-stability/banking-sector. [Accessed: 25- Nov- 2019].

[3] Symantec.com, 2019. [Online]. Available: https://www.symantec.com/content/en/us/enterprise/other_resources/b_Financial_Attacks_Exec_Report.pdf. [Accessed: 27- Nov- 2019].

[4] "Data Breach at Desjardins Bank Caused by Malicious Insider | CyberArk", *CyberArk*, 2020. [Online]. Available: https://www.cyberark.com/blog/data-breach-at-desjardins-bank-caused-by-malicious-insider/. [Accessed: 05- Jan- 2020].

[5] "Bangladesh probes 2013 hack for links to Swift-linked central bank heist", CNBC, 2019. [Online]. Available: https://www.cnbc.com/2016/05/25/bangladesh-probes-2013-hack-for-links-to-swift-linked-central-bank-heist.html. [Accessed: 28- Nov- 2019].

[6] "Massive Fraud At Sampath Bank: Millions Of Rupees Siphoned Off By Manager And Partners In Crime: Media Hushed Using Advertising Power", Colombo Telegraph, 2019. [Online]. Available: https://www.colombotelegraph.com/index.php/massive-fraud-at-sampath-bank-millions-of-rupees-siphoned-off-by-manager-and-partners-in-crime-media-hushed-using-advertising-power.[Accessed: 28- Nov- 2019].

[7] Bankinfosecurity.com. (2019). Commercial Bank of Ceylon Hacked?. [online] Available at: https://www.bankinfosecurity.com/commercial-bank-ceylon-apparently-hacked-a-9103. [Accessed: 28- Nov- 2019].

[8] "Banks issue warning on ATM skimming activities", Dailymirror.lk, 2019. [Online]. Available: http://www.dailymirror.lk/article/Banks-issue-warning-on-ATM-skimming activities -161961.html. [Accessed: 23- Dec- 2019].

[9] "What is Information Security (InfoSec)? -- Definition from WhatIs.com", SearchSecurity, 2019. [Online]. Available: https://searchsecurity.techtarget.com/definition/information-security -infosec. [Accessed: 23- Dec- 2019].

[10] L. Matsakis et al., "Everything We Know About Facebook's Massive Security Breach", WIRED, 2019. [Online]. Available: https://www.wired.com/story/facebook-security-breach-50-million-accounts. [Accessed: 28- Dec- 2019].

[11] N. Alam, "The great Bangladesh cyber heist shows truth is stranger than fiction", Dhaka Tribune, 2019. [Online]. Available: https://www.dhakatribune.com/uncategorized/2016/03/12/the-great-bangladesh-cyber-heist-shows-truth-is-stranger-than-fiction.[Accessed: 31- Dec- 2019].

[12] M. Stamps, “The Evolution of the Hacker,” The Evolution of the Hacker. [Online]. Available: https://blog.techguard.com/the-evolution-of-the-hacker. [Accessed: 31- Dec- 2019].

[13] "Baseline Security Standard for Information Security Management", Cbsl.gov.lk, 2019. [Online]. Available: https://www.cbsl.gov.lk/sites/default/files/cbslweb_documents/laws/cdg/Attachement_4_BaselineSecurityStandard.pdf. [Accessed: 02- Jan- 2020].

[14] "Introduction to FIS - Financial Services", Financial Services, 2020. [Online]. Available: https://finance.utoronto.ca/policies/gtfm/financial-information-system-fis/introduction-to-fis/. [Accessed: 02- Jan- 2020].

[15] G. Kul and S. Upadhyaya, "Towards a Cyber Ontology for Insider Threats in the Financial Sector", vol. 6, pp. 64-85, 2015. Available: https://www.researchgate.net/publication/290448680.

[16] T. Damenu and C. Beaumont, "Analysing information security in a bank using soft systems methodology", Information and Computer Security, vol. 25, no. 3, pp. 240-258, 2017.

[17] M. Ula, Z. Ismail and Z. Sidek, "A Framework for the Governance of Information Security in Banking System", Journal of Information Assurance & Cybersecurity, pp. 1-12, 2011.

[18] A. Da Veiga and J. Eloff, "A framework and assessment instrument for information security culture", Computers & Security, vol. 29, no. 2, pp. 196-207, 2010.

[19] M. Kooper, R. Maes and E. Lindgreen, "On the governance of information: Introducing a new concept of governance to support the management of information", International Journal of Information Management, vol. 31, no. 3, pp. 195-200, 2011.

[20] B. Bojinov, "Information Security in Banking Conceptual Issues", SSRN Electronic Journal, 2016. Available: 10.2139/ssrn.2905801.

[21] L. T. Khrais, "Highlighting the Vulnerabilities of Online Banking System", The Journal of Internet Banking and Commerce, 14-Sep-2015. [Online]. Available: http://www.icommercecentral.com/open-access/highlighting-the-vulnerabilities-of-online-banking-system.php?aid=61518. [Accessed: 09-Sep-2019].

[22] A. Malpani, and R. Singh, “An Investigation of the Factors Affecting the Security of Management Information System in Financial Institutions”, The IUP Journal of Systems Management, vol. IX, no. 2, pp. 7-27, 2011.

[23] M. Fazlida and J. Said, "Information Security: Risk, Governance and Implementation Setback", Procedia Economics and Finance, vol. 28, pp. 243-248, 2015.

[24] BUBU, Ioan Alexandru. "Banking Software Applications Security. Journal of Mobile", Embedded and Distributed Systems, [S.l.], vol. 7, no. 1, pp. 35-40, mar. 2015. ISSN 2067-4074.

Available:http://www.jmeds.eu/index.php/jmeds/article/view/Banking_Software_Applications_Security.

[25] Drabikova, K. Matiasko and A. Lieskovsky, "Database Object Dependency tool", 2016 7th IEEE International Conference on Software Engineering and Service Science (ICSESS), 2016. Available: 10.1109/icsess.2016.7883240.

[26] J. Wagner, A. Rasin and J. Grier, "Database forensic analysis through internal structure carving", Digital Investigation, vol. 14, pp. S106-S115, 2015. Available: 10.1016/j.diin.2015.05.013.

[27] "Introducing Oracle Database Vault", Docs.oracle.com, 2020. [Online]. Available: https://docs.oracle.com/cd/B28359_01/server.111/b31222/dvintro.htm#DVADM700. [Accessed: 05- Jan- 2020].

[28] Media.scmagazine.com, 2020. [Online]. Available: https://media.scmagazine.com/documents/296/2017_ibm_x-force-_security_tre_73846.pdf. [Accessed: 06- Jan- 2020].

[29] "Mixed Methods Research: A discussion paper", *Citeseerx.ist.psu.edu*, 2020. [Online]. Available: http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.468.360. [Accessed: 17- Sep- 2019].

[30] "Social Research Methods", *SAGE Publications Inc*, 2012. [Online]. Available: https://us.sagepub.com/en-us/nam/social-research-methods/book233436. [Accessed: 17- Apr- 2019].

[31] "JavaFX", Openjfx.io, 2020. [Online]. Available: https://openjfx.io/. [Accessed: 17- Dec- 2019].

[32] "Lesson: JDBC Basics (The Java™ Tutorials > JDBC(TM) Database Access)", Docs.oracle.com, 2020. [Online]. Available: https://docs.oracle.com/javase/tutorial/jdbc/basics/index.html. [Accessed: 17- Dec- 2019].

[33] "Gson User Guide - gson", Sites.google.com, 2020. [Online]. Available: https://sites.google.com/site/gson/gson-user-guide. [Accessed: 17- Dec- 2019].

[34] A. Bhattacherjee, Social science research: Principles, methods, and practices. 2012, p.39

[35] A. Amiri, "Dare to share: Protecting sensitive knowledge with data sanitization", Decision Support Systems, vol. 43, no. 1, pp. 181-191, 2007. Available:10.1016/j.dss.2006.08.007 [Accessed 6- Dec- 2019].

[36] T. Mbelli and B. Dwolatzky, "Cyber Security, a Threat to Cyber Banking in South Africa: An Approach to Network and Application Security", 2016 IEEE 3rd International Conference on Cyber Security and Cloud Computing (CSCloud), 2016. Available: 10.1109/cscloud.2016.18 [Accessed 2 May 2019].

[37] P. Subsorn and S. Limwiriyakul, "A comparative analysis of the security of internet banking in australia: A customer perspective", 2011.

[38] "Overview of Nonexperimental Research – Research Methods in Psychology", Opentextbc.ca, 2019. [Online]. Available: https://opentextbc.ca/researchmethods/chapter/overview-of-nonexperimental-research/. [Accessed: 06- Jul- 2019]

[39] T. Brar, D. Sharma, and S. S. Khurmi,"Vulnerabilities in e-banking: A study of various security aspects in e-banking,"International Journal of Computing & Business Research//T.Brar, D. Sharma, S. Khurmi, no. 6,pp. 127–132, 2012.

[40] Ibrahim and Y. Fadlalla, "A review on online-banking security models, successes, and failures," Feb. 2018.

[41] J. Gates, "What's in a Password?", EDPACS: The EDP Audit, Control, and Security Newsletter, 19(8), pp.5-11,1992.

[42]"Ensuring Data Quality Through Input Validation", W3computing.com, 2020. [Online]. Available: https://www.w3computing.com/systemsanalysis/ensuring-data-quality-input-validation/. [Accessed: 18- Feb- 2020]

[43] D. Trivedi, P. Zavarsky and S. Butakov, "Enhancing Relational Database Security by Metadata Segregation", Procedia Computer Science, vol. 94, pp. 453-458, 2016. Available: 10.1016/j.procs.2016.08.070.

[44] R. Alexandra, et al. ,"A machine learning approach to foreign key discovery", *WebDB*. 2009.

[45] A. Ilyas, et al. "Extracting syntactical patterns from databases", In 2018 IEEE 34th International Conference on Data Engineering (ICDE), pp. 41-52, 2018.

[46] M. Lopez, "Why Dormant Financial Account Fraud Goes Undetected", AppGate | Total Fraud Protection Blog, 2015. [Online]. Available: https://blog.easysol.net/dormant-financial-account-fraud/. [Accessed: 18- Oct- 2019].

[47] "3 Reasons why MD5 is not Secure | MD5Online", Md5online.org, 2016. [Online]. Available: https://www.md5online.org/blog/why-md5-is-not-safe/. [Accessed: 14- Dec-2019].

[48] "SHA-256 is not a secure password hashing algorithm - Dusted Codes", Dusted.codes, 2016. [Online]. Available: https://dusted.codes/sha-256-is-not-a-secure-password-hashing-algorithm. [Accessed: 14 - Dec- 2019].

[49] Central Bank of Sri Lanka, "Directions, Rules, Determinations, Notices, And Guidelines Applicable To Licensed Finance Companies And Specialised Leasing Companies", Department of Supervision of Non-Bank Financial Institutions, Colombo 01,pp. 56, 2013. [Online]. Available: https://www.cbsl.gov.lk/sites/default/files/cbslweb_documents/laws/cdg/snbfi_LFCsDirectionBook_Nov2013.compressed_0.pdf. [Accessed: 20 - Dec- 2019].

[50] "SQL Join - Introduction to SQL Join Syntax and Concepts - Essential SQL", Essential SQL, 2020. [Online]. Available: https://www.essentialsql.com/introduction-sql-joins/. [Accessed: 12- Oct- 2019].

[51] "Password Policy | Policies, Standards, and Guidelines | Policies and Regulations | Office of Information Security | Access and Security | EITS", Eits.uga.edu, 2020. [Online]. Available: https://eits.uga.edu/access_and_security/infosec/pols_regs/policies/passwords/. [Accessed: 19- Oct- 2019].

[52]C. Osborne, "UniCredit reveals data breach exposing 3 million customer records | ZDNet", ZDNet, 2020. [Online]. Available:

https://www.zdnet.com/article/unicredit-reveals-data-breach-exposing-3-million-customer-records/. [Accessed: 04- Jan- 2020].

Appendix A

1. Code for relationship generator

Link to source code: https://github.com/SalithaUCSC/DB-Tool

Relationship generation process was implemented to check the existence of a relationship between two tables. Technique used here is matching the content similarity of selected 2 columns of 2 tables.

```
void genBtn(ActionEvent event) throws SQLException, IOException {
    if (t1Drop.getValue() != null && t2Drop.getValue() != null && c1Drop.getValue() != null && c2Drop.getValue() != null) {
        ArrayList<String> list1 = db.getTableColumnData(
            db.getConnectionToDb(selectedDb.getText()),t1Drop.getValue(),c1Drop.getValue());
        ArrayList<String> list2 = db.getTableColumnData(
            db.getConnectionToDb(selectedDb.getText()),t2Drop.getValue(),c2Drop.getValue());
        int count = 0;
        for (int i=0; i<list1.size();i++) {
            for (int j=0; j<list2.size();j++) {
                if (list1.get(i).equals(list2.get(j))) {
                    count++;
                }
            }
        }
        double res1 = ((double) count)/list1.size();
        double res2 = ((double) count)/list2.size();
        double threshold = 80;
        if (ob.has("structure")){
            JsonArray objArray = ob.get("structure").getAsJsonArray();
            if (res1*100>threshold || res2*100>threshold) {
                jsonObject = getKeyRelation(objArray);
                relListView.getItems().clear();
                relListView.getItems().add("Tables: "+t1Drop.getValue()+" and "+t2Drop.getValue());
                JsonArray jArr = jsonObject.get("relations").getAsJsonArray();
                relListView.getItems().add("Relation : "+jArr.get(jArr.size()-1));
                saveBtn.setDisable(false);
            }
            else {
                Alert alert = new Alert(Alert.AlertType.WARNING);
                alert.setTitle("Warning Message");
                alert.setHeaderText(null);
                alert.setContentText("There's no relation between the entities!");
                alert.showAndWait();
            }
            saveBtn.setDisable(false);
        }
    }
    else {
        Alert alert = new Alert(Alert.AlertType.ERROR);
        alert.setTitle("Error Message");
        alert.setContentText("Please select two tables!");
        alert.showAndWait();
    }
}
```

```
public JsonObject getKeyRelation(JsonArray objArray){
    JsonObject objRel = new JsonObject();
    for(JsonElement x: objArray) {
        if (x.getAsJsonObject().get("table").getAsString().equals(t1Drop.getValue())) {
            objRel = x.getAsJsonObject();
            JsonArray relations = x.getAsJsonObject().get("relations").getAsJsonArray();
            JsonObject jobj = new JsonObject();
            jobj.addProperty("table", t2Drop.getValue());
            jobj.addProperty("key1", c1Drop.getValue());
            jobj.addProperty("key2", c2Drop.getValue());
            relations.add(jobj);
            objRel.add("relations",relations);
            relListView.getItems().addAll(objRel);
        }
    }
    System.out.println(objRel);
    return objRel;
}
```

2. Code for query parser

Query parser was implemented to extract JOIN statements from the CREATE VIEW sql queries stored in the database. It was written in Java with the use of JDBC and GSON libraries.

2.1. Parsing method

```
public Set<String> getDbViewJoins() throws SQLException {
    String splitString = "JOIN";
    Pattern pattern = Pattern.compile(splitString);
    ArrayList<String> splitWords = new ArrayList<>();
    ArrayList<String> splitNoConflicts = new ArrayList<>();
    Set<String> set = new HashSet<>();
    Set<String> joinsArray = new HashSet<>();
    DBHandler dbHandler = new DBHandler();
    ArrayList<JsonObject> views = dbHandler.getViews("SEEDS");
    for (int n = 0; n < views.size(); n++) {
        String[] split = pattern.split(String.valueOf(views.get(n).getAsJsonObject().get("code")));
        for(int i=1; i < split.length; i++){
            split[i] = split[i].trim();
            split[i] = split[i].replaceAll("dbo.", "");
            split[i] = split[i].replaceAll(",", "");
            split[i] = split[i].replaceAll("'", "");
            split[i] = split[i].replaceAll("INNER", "");
            split[i] = split[i].replaceAll("OUTER", "");
            split[i] = split[i].replaceAll("CROSS", "");
            split[i] = split[i].replaceAll("\\)", "");
            split[i] = split[i].replaceAll("\\(", "");
            split[i] = split[i].replaceAll("FFROM_TTO", "FFRO_TTO");
            split[i] = split[i].replaceAll("LEFT", "");
            split[i] = split[i].replaceAll("\"",  Matcher.quoteReplacement(" "));

            if(split[i].contains("SELECT")) {
                split[i] = split[i].trim();
                int selectIndex = split[i].indexOf("SELECT");
                int fromIndex1 = split[i].indexOf("FROM");
                StringBuilder str1 = new StringBuilder(split[i]);
                split[i] = String.valueOf(str1.delete(selectIndex, fromIndex1));
                if(split[i].contains("FROM")) {
                    int fromIndex2 = split[i].indexOf("FROM");
                    StringBuilder str2 = new StringBuilder(split[i]);
                    split[i] = String.valueOf(str2.delete(fromIndex2, str2.length()));
                }
            }

            if(split[i].contains("GROUP BY")) {
                int groupIndex = split[i].indexOf("GROUP BY");
                StringBuilder str3 = new StringBuilder(split[i]);
                split[i] = String.valueOf(str3.delete(groupIndex, str3.length()));
            }
```

```
            if(split[i].contains("ORDER BY")) {
                int byIndex = split[i].indexOf("ORDER BY");
                StringBuilder str4 = new StringBuilder(split[i]);
                split[i] = String.valueOf(str4.delete(byIndex, str4.length()));
            }

            if(split[i].contains("WHERE")) {
                int whereIndex = split[i].indexOf("WHERE");
                StringBuilder str5 = new StringBuilder(split[i]);
                split[i] = String.valueOf(str5.delete(whereIndex, str5.length()));
            }

            String[] keyWords = {"UNION ALL", "WHERE", "CONVERT", "DERIVEDTBL", "SUBSTRING"};

            for (int j = 0; j < keyWords.length; j++) {
                if(split[i].contains(keyWords[j])) {
                    split[i] = split[i].replaceAll(keyWords[j], "");
                }
                else if(split[i].contains(keyWords[j].toLowerCase())) {
                    split[i] = split[i].replaceAll(keyWords[j].toLowerCase(), "");
                }
            }

            if(split[i].contains("ON")) {
                String[] noOns = split[i].split("ON");
                for (int j = 1; j < noOns.length; j++) {
                    noOns[j] = noOns[j].trim();
                    splitWords.add(noOns[j]);
                }
            }
        }

        for (int m = 0; m < splitWords.size(); m++) {
            if(splitWords.get(m).contains("FFRO_TTO")){
                String x = splitWords.get(m).replaceAll("FFRO_TTO", "FFROM_TTO");
                splitNoConflicts.add(x);
            }
            else {
                splitNoConflicts.add(splitWords.get(m));
            }
        }

        for (int k = 0; k < splitNoConflicts.size(); k++) {
            if(splitNoConflicts.get(k).contains("AND")) {
                String[] andSplits = splitNoConflicts.get(k).split("AND ");
                for (int j = 0; j < andSplits.length; j++) {
                    joinsArray.add(andSplits[j].trim());
                }
            }
            else {
                joinsArray.add(splitNoConflicts.get(k));
            }
        }
    }
    ArrayList<String> list = new ArrayList<>(joinsArray);
    set.addAll(list);
    return set;
}
```

2.2. Refining method

```
public Set<String> RefineDbJoins(Set<String> dbJoinSet) throws SQLException, IOException {
    ArrayList<String> dbTables = handler.getTables(handler.getConnectionUrl());
    FileWriter dbTablesFile = new FileWriter(System.getProperty("user.dir")+"\\src\\tool\\config\\db-tables.txt");
    for (int i = 0; i < dbTables.size(); i++) {
        dbTablesFile.append(dbTables.get(i)+"\n");
    }
    dbTablesFile.close();
    ArrayList<String> dbJoins = new ArrayList<>(dbJoinSet);
    Set<String> refined = new HashSet<>();
    System.out.println("dbJoins: "+dbJoins);
    for (int j = 0; j < dbJoins.size(); j++) {
        String[] splits = dbJoins.get(j).split(" = ");
        if(splits.length == 2 && splits[0].contains(".") && splits[1].contains(".")){
            String table1 = splits[0].split("\\.")[0];
            String table2 = splits[1].split("\\.")[0];
            String column1 = splits[0].split("\\.")[1];
            String column2 = splits[1].split("\\.")[1];
            if(dbTables.contains(table1) && dbTables.contains(table2)) {
                ArrayList<String> tab1cols = handler.getTableColumns(conn,table1);
                ArrayList<String> tab2cols = handler.getTableColumns(conn,table2);
                if(dbTables.contains(table1) && tab1cols.contains(column1) &&
                dbTables.contains(table2) && tab2cols.contains(column2)){
                    refined.add(dbJoins.get(j));
                }
            }

        }
    }
    System.out.println("after refine: "+refined);
    FileWriter writer = new FileWriter(System.getProperty("user.dir")+"\\src\\tool\\config\\refined-joins.txt");
    for(String s: refined) {
        writer.append(s + "\n");
    }
    writer.close();
    return refined;
}
```

Appendix B

1. Refined table joins obtained from query parser

Count = 57

1. RECEIPT.RE_CODE = RE_INCOME.RI_CODE
2. CR_DTAIL.CD_MEMBER = CUSTOMER.CM_CODE
3. RE_DTAIL.RD_CODE = RE_INCOME.RI_CODE
4. RE_ITEM.RI_CODE = RECEIPT.RE_CODE
5. TRANSFER.TR_CHQNO = RE_DTAIL.RD_CHEQUE
6. INCOME_EXPENCE_SUB.IE_MAIN_NUMBER_CODE = INCOME_EXPENCE_LOAN_OPPU_SUB.IS_CODE
7. INCOME_EXPENCE_SUB.IE_CODE = INCOME_EXPENCE_SUB_DTAIL.IE_CODE
8. INCOME_EXPENCE.IE_CODE = INCOME_EXPENCE_LEDGER.IE_CODE
9. TR_DTAIL.TD_LCODE = LCODE.LC_CODE
10. PA_DTAIL.PD_CODE = PA_EXPENCE.PE_CODE
11. PA_EXPENCE.PE_EXPENCE = INCOME_EXPENCE_SUB.IE_CODE
12. RE_INCOME.RI_CODE = RE_DTAIL.RD_CODE
13. DR_DTAIL.DD_MEMBER = CUSTOMER.CM_CODE
14. DEBIT.DR_CODE = DR_DTAIL.DD_CODE
15. PAYMENT.PA_CODE = PA_EXPENCE.PE_CODE
16. CUST_CATEGORY_MAIN.CC_CODE = CUST_CATEGORY.CC_MAIN
17. FFROM_TTO.FT_CODE = FT_DTAIL.FD_CODE
18. INCOME_EXPENCE_SUB.IE_MAIN = INCOME_EXPENCE_GIFT.IE_CODE
19. RE_INCOME.RI_CHEQUE = RE_DTAIL.RD_CHEQUE
20. INCOME_EXPENCE_LOAN_DEFFINE_SUB_DOC.IS_CODE = INCOME_EXPENCE_LOAN_DEFFINE.IE_CODE
21. RECEIPT.RE_CODE = RE_DTAIL.RD_CODE
22. INCOME_EXPENCE_LEDGER.IE_CODE = INCOME_EXPENCE.IE_CODE
23. INCOME_EXPENCE_SUB.IE_CODE = RE_INCOME.RI_INCOME
24. RECEIPT.RE_OUT_RECEIPT = OUT_RECEIPT.RE_CODE
25. INCOME_EXPENCE_SUB.IE_CODE = MEMBER_REMEMBER_DTAIL.MD_LOAN
26. FFROM_TTO.FT_FROMEXPENCE = INCOME_EXPENCE_SUB.IE_CODE

27. RE_CASH.RC_CODE = RECEIPT.RE_CODE
28. RE_INCOME.RI_CODE = RECEIPT.RE_CODE
29. PRINT5.PE_EXPENCE = INCOME_EXPENCE_SUB.IE_CODE
30. INCOME_EXPENCE.IE_CODE = TAX.TA_CODE
31. DEPOSIT.DP_CHQNO = REALIZE.RZ_CHQNO
32. FFROM_TTO.FT_FROMMEMBER = CUSTOMER.CM_CODE
33. TR_DTAIL.TD_CODE = TR_TYPE.TT_CODE
34. PA_DTAIL.PD_CHEQUE = PA_EXPENCE.PE_CHEQUE
35. DEPOSIT.DP_CHQNO = RE_DTAIL.RD_CHEQUE
36. FT_DTAIL.FD_TOMEMBER = CUSTOMER.CM_CODE
37. TR_TYPE.TT_CODE = TR_DTAIL.TD_CODE
38. CREDIT.CR_CODE = CR_DTAIL.CD_CODE
39. AV_DTAIL.AD_MEMBER = CUSTOMER.CM_CODE
40. RE_INCOME.RI_INCOME = INCOME_EXPENCE_SUB.IE_CODE
41. PA_ITEM.PI_CODE = PAYMENT.PA_CODE
42. CHQ_RETURN.RT_CHQNO = RE_DTAIL.RD_CHEQUE
43. PA_CASH.PC_CODE = PAYMENT.PA_CODE
44. INCOME_EXPENCE_SUB.IE_MEMBER = CUSTOMER.CM_CODE
45. CT_CASH.CC_CODE = CASH_TRANSFER.CT_CODE
46. PAYMENT.PA_CODE = PA_DTAIL.PD_CODE
47. CUSTOMER.CM_SEX = SEX.CODE
48. DEPOSIT.DP_CHQNO = RE_INCOME.RI_CHEQUE
49. INCOME_EXPENCE_SUB_MEMBER.IE_CODE = MEMBER_REMEMBER_DTAIL.MD_LOAN
50. INCOME_EXPENCE_SUB.IE_MAIN = INCOME_EXPENCE.IE_CODE
51. RE_DTAIL.RD_CHEQUE = RE_INCOME.RI_CHEQUE
52. PAYMENT.PA_SUPPLIER = CUSTOMER.CM_CODE
53. CASH_WITHDRAW.CW_ACCOUNT = ACCOUNT.AC_CODE
54. SLEDGER.SG_LCODE = LCODE.LC_CODE
55. CASH_DEPO.CD_ACCOUNT = ACCOUNT.AC_CODE
56. RECEIPT.RE_CUSTOMER = CUSTOMER.CM_CODE
57. CUSTOMER.CM_CATEGORY = CUST_CATEGORY.CC_CODE

2.Table joins obtained from Database Analyser

Unique joins count = 60/2 = 30

1. ACCOUNT.AC_CODE = CASH_WITHDRAW.CW_ACCOUNT
2. RECEIPT.RE_CODE = RE_INCOME.RI_CODE
3. CUSTOMER.CM_CODE = FFROM_TTO.FT_FROMMEMBER
4. CR_DTAIL.CD_MEMBER = CUSTOMER.CM_CODE
5. CUSTOMER.CM_CODE = FT_DTAIL.FD_TOMEMBER
6. CUSTOMER.CM_SEX = CUSTOMER_SEX.ID
7. CUSTOMER.CM_CODE = PAYMENT.PA_SUPPLIER
8. DEPOSIT.DP_CHQNO = RE_DTAIL.RD_CHEQUE2
9. INCOME_EXPENCE.IE_CODE = INCOME_EXPENCE_LEDGER.IE_CODE
10. TR_DTAIL.TD_LCODE = LCODE.LC_CODE
11. CR_DTAIL.CD_CODE = CREDIT.CR_CODE
12. PA_EXPENCE.PE_EXPENCE = INCOME_EXPENCE_SUB.IE_CODE
13. CUSTOMER_SEX.ID = CUSTOMER.CM_SEX
14. PAYMENT.PA_CODE = PA_EXPENCE.PE_CODE
15. TAX.TA_CODE = INCOME_EXPENCE_SUB.IE_CODE
16. INCOME_EXPENCE_SUB.IE_CODE = FIRST_LETTER.FL_LOAN_CODE
17. RE_INCOME.RI_CHEQUE = DEPOSIT.DP_CHQNO
18. RE_DTAIL.RD_CODE = RECEIPT.RE_CODE
19. PA_EXPENCE.PE_CHEQUE = PA_DTAIL.PD_CHEQUE
20. INCOME_EXPENCE_LEDGER.IE_CODE = INCOME_EXPENCE.IE_CODE
21. INCOME_EXPENCE_SUB.IE_CODE = RE_INCOME.RI_INCOME
22. INCOME_EXPENCE_SUB.IE_CODE = TAX.TA_CODE
23. INCOME_EXPENCE_SUB.IE_CODE = PA_EXPENCE.PE_EXPENCE
24. PA_EXPENCE.PE_CODE = PAYMENT.PA_CODE
25. RE_INCOME.RI_CODE = RECEIPT.RE_CODE
26. DEPOSIT.DP_CHQNO = REALIZE.RZ_CHQNO
27. FFROM_TTO.FT_FROMMEMBER = CUSTOMER.CM_CODE
28. PA_DTAIL.PD_CHEQUE = PA_EXPENCE.PE_CHEQUE
29. CUST_CATEGORY.CC_CODE = CUST_CATEGORY_MAIN.CC_CODE
30. DEPOSIT.DP_CHQNO = RE_DTAIL.RD_CHEQUE
31. RE_DTAIL.RD_CHEQUE2 = DEPOSIT.DP_CHQNO
32. CUSTOMER.CM_STAT = CUSTOMER_STATUS.ID

33. FT_DTAIL.FD_TOMEMBER = CUSTOMER.CM_CODE
34. CREDIT.CR_CODE = CR_DTAIL.CD_CODE
35. CUST_CATEGORY.CC_CODE = CUSTOMER.CM_CATEGORY
36. CUSTOMER.CM_CODE = INCOME_EXPENCE_SUB.IE_MEMBER
37. FIRST_LETTER.FL_LOAN_CODE = INCOME_EXPENCE_SUB.IE_CODE
38. LCODE.LC_CODE = SLEDGER.SG_LCODE
39. RE_DTAIL.RD_CODE = DEPOSIT.DP_DSLIP
40. RE_DTAIL.RD_CHEQUE = DEPOSIT.DP_CHQNO
41. RE_INCOME.RI_INCOME = INCOME_EXPENCE_SUB.IE_CODE
42. DEPOSIT.DP_DSLIP = RE_DTAIL.RD_CODE
43. INCOME_EXPENCE_SUB.IE_MEMBER = CUSTOMER.CM_CODE
44. TR_TYPE.TD_CODE = TR_DTAIL.TT_CODE
45. CUSTOMER.CM_CODE = RECEIPT.RE_CUSTOMER
46. LCODE.LC_DIVISION = DIVISION.DI_CODE
47. LCODE.LC_CODE = TR_DTAIL.TD_LCODE
48. REALIZE.RZ_CHQNO = DEPOSIT.DP_CHQNO
49. CUST_CATEGORY_MAIN.CC_CODE = CUST_CATEGORY.CC_CODE
50. DEPOSIT.DP_CHQNO = RE_INCOME.RI_CHEQUE
51. CUSTOMER.CM_CODE = CR_DTAIL.CD_MEMBER
52. PAYMENT.PA_SUPPLIER = CUSTOMER.CM_CODE
53. CASH_WITHDRAW.CW_ACCOUNT = ACCOUNT.AC_CODE
54. DIVISION.DI_CODE = LCODE.LC_DIVISION
55. RECEIPT.RD_CODE = RE_DTAIL.RE_CODE
56. SLEDGER.SG_LCODE = LCODE.LC_CODE
57. CUSTOMER_STATUS.ID = CUSTOMER.CM_STAT
58. TR_DTAIL.TT_CODE = TR_TYPE.TD_CODE
59. RECEIPT.RE_CUSTOMER = CUSTOMER.CM_CODE
60. CUSTOMER.CM_CATEGORY = CUST_CATEGORY.CC_CODE

Appendix C

Interview Script

1. What is your role at the bank?
2. On average how many transactions does the system handle during a day?
3. Is this system isolated or is it made available to other locations over a network/internet?
4. How is your system connected to the network/internet?
5. How do you manage passwords? What are the preconditions on setting up passwords?
6. Do you believe there's adequate training/information provided at work to the users of the system?
7. What type of documentation or manuals were provided along with the software?
8. How often did the supplier conduct maintenance activity on the software?
9. Who was in charge of overseeing the frequent issues faced by employees with regard to the software system?
10. How often were system audits conducted?
11. Was the software an off-the-shelf purchase or did the supplier gathered specifics of your requirement and developed a unique solution for the bank?
12. How often did users share passwords with each other to access the system?
13. How often were backups created?
14. How was data initially migrated to the current system for operations?
15. Have you experienced downtime in the software as a result of load and performance issues?
16. Does the system comply with the rules and guidelines laid out by the CBSL with regard to the maintenance of financial information security?
17. Do the bank's accounting systems properly manage and report bank transactions in accordance with the proper accounting standards?
18. What were the widely used user roles within the system and the bank?
19. What sort of encryption mechanisms has been used to protect confidential data of customers?
20. What sort of reports were periodically generated, and for what purpose?