A Survey Research of Satisfaction Levels on Preventing Data Loss and Preserving Privacy

[*]H. Lakshmi, [#]Dr. K. Nageswara Rao

[*]Research Scholar, Rayalaseema University, Kurnool, Andhra Pradesh, India.

itslakshmi.h@gmail.com

[#]Principal, Potti Sriramulu Chalavadi Mallikarjuna Rao College of Engineering & Technology, Andhra Pradesh, India.

Abstract **– At present in our digital world, data comes and leaves cyberspace at huge rates. A representative organization transfers millions of email messages and downloads, stores, and transmits millions of data sets via various channels on a regular basis. Companies always hold private data of customers, stake holders, industry partners, regulators and they expect them to protect. Unfortunately, today's industries constantly fall victim to massive data loss, and high-profile data leakages involving sensitive personal and corporate data continue to appear (http://opensecurityfoundation.org). Loss of data could significantly damage a company's goodwill and reputation and could also invite legal issues or regulatory consequences for negligent security. That's why, organizations should take measures to manage the sensitive data they carried out, how it's restricted, and how to prevent the loss from being leaked or compromised. In this respect, over the years the database security community has developed a number of different techniques and approaches to assure data confidentiality, integrity, and availability[14]. Thus data loss prevention and in particular protection of data from unauthorized accesses remain important goal of any data management system. Multi Category Security labeling from a user and system administrator standpoint is straightforward. It consists of configuring a set of categories, which are simply text labels, such as "Company_Confidential" or "Medical_Records", and then assigning users to those categories. The system administrator first configures the categories, then assigns users to them as required. The users can then use the labels as they see fit. A system in a home environment may have only one category of "Private", and be configured so that only trusted local users are assigned to this category. In this paper, we first survey the most relevant concepts underlying the notion of database security, types of losses and summarize the menaces to databases and different categories of vulnerabilities in database. This paper focused on Virtual private database, stops various sensitive data from leaving the corporation's private confines. This paper illustrates and demonstrates how to enable mutli-level access restrictions which ensures accuracy and security.**

Index Terms –Software vulnerabilities, Data Secrecy, Data Privacy, Fine-grained access level.

I. INTRODUCTION

As companies increase their adoption of database systems as the key database management technology for day-to-day operations, knowledge sharing and decision making, the data security managed by these systems becomes crucial[1]. Damage and misuse of data affect not only a single user or application, but may have disastrous consequences on the entire organization.

The recent rapid proliferation of Web based applications and information systems have further increased the risk exposure of databases and, thus, data protection is today more crucial than ever. It is also important to appreciate that data needs to be protected not only from external threats, but also from insider threats.

Database security issues are typically categorized as unauthorized data inspection, data unavailability and incorrect data manipulation. Unauthorized data surveillance results in the disclosure of private data to users not permitted to gain access to such information. All organizations, ranging from commercial private organizations to government sectors, in a variety of domains such as medical, military and public protection, may suffer heavy losses from both economic and human points of view as a consequence of unauthorized data inspection. When data is not available in

 DOI : 10.35291/2454-9150.2020.0001

time most of the organizations will suffer with project deliver commitment issues. This may lose their customers as well as price of the company shares. Incorrect data manipulation involves either planned or unplanned which may result in an erroneous database situation. Any use of incorrect data may result in heavy losses for the companies. This may leads to company's bad reputation towards the customers.

A. Organization of the Paper

The remainder of the paper is organized as follows: Section 2 discusses types of losses. Section 3 presents Data privacy and Section 4 presents software vulnerabilities. Section 5 specifies an overview of Virtual Private database concepts and Section 6 outlines the benefits of multi level access restrictions. Finally, Section 7 discuss about the limitation and conclusion.

II. FACEBOOK USER'S DATA SECURITY AND AWARENESS: A LITERATURE REVIEW

In 2005, a year after Facebook was launched; Govani and Pashley (2005) studied how student shares their information despite the many privacy issues in the new system. They found out that about 84% of Facebook users were aware of the privacy settings but only half of them used it. These numbers are almost similar as reported by Jones and Soltren (2005). It was also found out that adult and young Facebook users have similar behaviour when it comes to privacy settings[15]. Five years after, the picture is still the same significantly smaller proportion from older and younger users of Facebook actually used privacy settings (Christofides et al., 2010). There are however reports showing different numbers.

It might be attributed to other factors such as knowing how the information generated from Facebook accounts can be used against the user.

Year	Privacy Glitches
2006	News Feed: The new feature allowed every post of the user to appear to friends Facebook wall. Privacy controls were introduced after about 1 million users protested. The feature later becomes one of the major parts of its success.
2007	Advertisement: The feature allowed the company to track purchases of by Facebook users and notify their friends what was bought even without user's consent.
2013	Bug exposes private contact information: About 6 million users were revealed to anyone who had some connection to the person or if they have at least one contact information.
2014	Mood-manipulation experiment: It involved more than half a million randomly selected Facebook user for the experiment. The result of the experiment was published and later removed due to ethical issues.
2015	Cuts off apps from taking any data from Facebook: An app downloaded by user A can allow extracting user's A friend's data. Even after the stoppage, third-party users were known to have still used the previously collected data.
2018	Privacy Bug: About 14 million users were affected. They may have unknowingly posted private information to the public. 2018 87 million user's data: A researcher had sold Facebook data collected via a personal quiz was revealed.

Sources: Newcomb, 2018; Rodriguez et al., 2018; Lee, 2018)

A study in The Netherlands reveals that almost 6 in every 10 of their Facebook users are not comfortable with the level of exposure (Yazici, 2017), but they remain to use it for the benefit they get from it[15]. Facebook users have been giving away personal for more than a decade. Those who have exhibited a higher level of privacy awareness are the ones who have negative social network site experiences (ibid). The number of aware about privacy settings since 2005 to date is ever changing.

III. CATEGORIES OF LOSS

We can divide data loss into two sometimes overlapping categories:

• Leakage, in which sensitive data is no longer under the organization's control. This type of data loss is often due to hacked customer confidential databases, making its most common consequence identity theft. Example for this type of Data loss: Hackers stole 130 million credit-card records from one of the US's largest payment processors (datalossdb.org)[2]. Another involved 94 million customer records held at a major retailer.

• Loss, data sets are lost from the company's database and no longer available. An example occurred in 2009, when a major cell phone service provider suffered widespread loss of customer data that was supposed to be housed by a third-party cloud-based storage service[2]. In normal operation, the smart phone would automatically sync its data at power-off with the central server, which stores it for use when the phone is on again.

• Misuse, hackers will misuse the data sets, include some malicious code in it and upload in the public domains so that clients of specific organizations will suffer.

IV. DATA PRIVACY

The center of attention of any privacy law is on personal information. Generally, this includes any private data of an individual. It can be as little as a name, e-mail address, or phone number, pictures or it can include much more extensive data such as an individual's economic or medical issues.

Privacy laws regulate various aspects of the collection, use, processing, storage, and disclosure of all such personal information. US federal privacy law imposes comprehensive regulations in the financial and healthcare sectors, but there are few privacy rules outside those sectors[3]. In the European Union and several other countries, however, all personal information is subject to comprehensive regulation in all sectors.

Most countries, including the US, often apply special rules to the privacy of more sensitive types of personal information, regardless of sector. In the US, depending on the jurisdiction, such rules might apply, for example, to Social Security numbers, drivers' license numbers, information regarding

 DOI : 10.35291/2454-9150.2020.0001

medical or health conditions, credit or debit card numbers, and financial account information.

In the EU, data protection laws apply special rules to sensitive personal information such as racial or ethnic origin, political opinions, religious or philosophical beliefs, union membership, or information specifying an individual's sexual orientation.

V. SOFTWARE VULNERABILITIES

Various software threats are stated every year, some common and others rare. A frequently mentioned security breach in any given year might subside in subsequent years because of prevention and detection methods used by developers[4]. Every time security subsystems are getting modified based on the vulnerabilities. But, when a security threat is addressed and resolved, intruders are trying to find out new security breaches and issues and uploading them to penetrate the organization's software. Security management systems are working towards every new attack. Security professionals then develop more countermeasures to defend systems. In other words, both attacker and defender are always are working towards their duties like harming the data and protecting the data.

A. Tracing Vulnerabilities

Vulnerability repositories and databases can be traced to study trends and find severe vulnerabilities. The three most popular repositories are the Open Source Vulnerability Database (OSVDB; blog.osvdb.org), the Exploit database (www.exploit-db.com), and the National Vulnerability Database (NVD; nvd.nist.gov)[4]. OSVDB covers only web application vulnerabilities. The Exploit database archives exploits and vulnerable software; as such, it's not a suitable source for the discovery of vulnerabilities. According to its website, "NVD is the US government repository of standards based vulnerability management data represented using the Security Content Automation Protocol."

NVD, which contains common vulnerability exposures (CVE) vulnerabilities, US Computer Emergency Readiness Team (US-CERT) alerts, US-CERT vulnerability notes, and Open Vulnerability and Assessment Language (OVAL) queries, is one of the most complete repositories of reported vulnerabilities. For example, NVD reported 6,514 vulnerabilities in 2007, but only 2,779 of them have been classified. NVD uses a subset of Common Weakness Enumeration (CWE; cwe.mitre.org) graph construction to categorize vulnerabilities. In the revision applied by NVD in 2014, this classification encompassed 35 categories. However, the earlier version had only 23 categories.

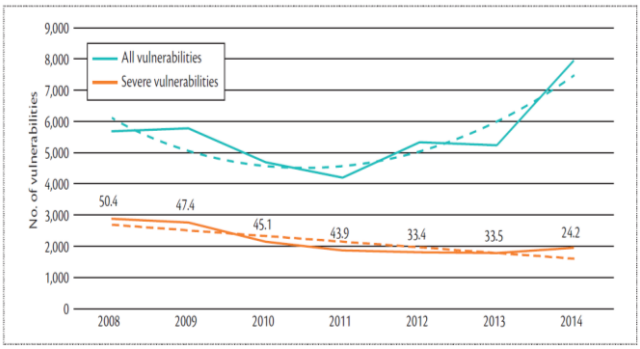

Figure 1: Blue and orange lines represent all vulnerabilities and severe vulnerabilities, respectively, registered on the National Vulnerability Database. Dashed lines represent trends extracted from solid lines.

B. VIRTUAL PRIVATE DATABASE

Virtual Private Database is also known as Fine-Grained Access Control (FGAC). It allows defining, which rows users may access. Modern database applications with large numbers of users require fine-grained access control (FGAC) mechanisms at the level of individual tuples, not just entire relations/views, to control which parts of the data can be accessed by each user[5].

Consider the following scenario: In a commercial organization's human resources database, the human resources manager should have access to all the personal details of employees. At the same time, individual employees should only be able to see their particulars, not other employees' information. In the above case, authorization is required at a very fine-grained level, such as at the level of individual tuples. Similar scenarios exist in many environments, including finance, law, government, and military applications. Consumer privacy requirements are yet another emerging driver for finer control of data.

Currently, general data authorization mechanisms in relational databases permit access control at the level of complete tables or columns, or on views. There is no direct way to specify fine-grained authorization to control, which tuples can be accessed by users. In theory, FGAC, at the level of individual tuples, can be achieved by creating an access control list for each tuple.

C. Benefits of FGAC

Today's network applications require much more secure data storages than ever before. With millions of anonymous users using same networking applications, security of data behind the applications have become a major concern of database developers and security experts. In most security incidents, the databases attached to the applications are targeted, and attacks have been made. Most of these applications require allowing data manipulation at several granular levels to the users accessing the applications—not just table and view level, but tuple level. A database that supports fine-grained access control restricts the rows a user sees, based on his/her credentials.

 DOI : 10.35291/2454-9150.2020.0001

Generally, this restriction is enforced by a query modification mechanism automatically done at the database. This feature enables per-user data access within a single database, with the assurance of physical data separation. It is enabled by associating one or more security policies with tables, views, table columns, and table rows. Such a model is ideal for minimizing the complexity of the security enforcements in databases based on network applications. With fine-grained access controls, one can create fast, scalable, and secure network applications.

Each application can be written to find the correct balance between performance and security, so that each data transaction is performed as quickly and safely as possible. Today, the database vendors like Oracle 10g, and IBM DB2 provides commercial implementations of fine-grained access control methods, such as filtering rows, masking columns selectively based on the policy, and applying the policy only when certain columns are accessed.

The behavior of the fine-grained access control model can also be increased through the use of multiple types of policies based on the nature of the application, making the feature applicable to multiple situations. Meanwhile, Microsoft SQL Server2005 has also come up with emerging features to control the access to databases using fine-grained access controls. Fine-grained access control does not cover all the security issues related to Internet databases, but when implemented, it supports building secure databases rapidly and bringing down the complexity of security management issues.

VI. MULTI LEVEL ACCESS RESTRICTIONS

A complete solution to data protection must meet three key requirements: (i) Confidentiality — it refers to the protection of data against unauthorized access; (ii) integrity — it refers to the prevention of improper data modifications; and (iii) Availability — it refers to the prevention and recovery from hardware and software errors and from malicious data access denials making the database system unavailable. These three requirements arise practically in all applications.

Consider a database storing medical information about patients of a hospital. It is important that patient records not be released to unauthorized subjects, that records be modified only by the subjects who are properly authorized and their accuracy be assured, and that patient records be readily available to doctors in charge especially in emergency situations. Securing data is a challenging task. It is ensured collectively by various components of a database management system (DBMS) and may also require components external to the DBMS, such as secure co-processors. A key component for assuring data protection is represented by the access control mechanism.

When a user attempts to access some data, the access control mechanism checks whether or not the user has the authorization to perform the action on the data. Authorizations are granted to users according to the access control policies of the organization.

A. Multi-Level Security (MLS)

Protecting sensitive or confidential data is paramount in many businesses. In the event such information is made public, businesses may face legal or financial ramifications. At the very least, they will suffer a loss of customer trust. In most cases, however, they can recover from these financial and other losses with appropriate investment or compensation.

The same cannot be said of the defense and related communities, which includes military services, intelligence organizations and some areas of police service. These organizations cannot easily recover should sensitive information be leaked, and may not recover at all. These communities require higher levels of security than those employed by businesses and other organizations.

Having information of different security levels on the same computer systems poses a real threat. It is not a straight-forward matter to isolate different information security levels, even though different users log in using different accounts, with different permissions and different access controls.

Some organizations go as far as to purchase dedicated systems for each security level. This is often prohibitively expensive, however. A mechanism is required to enable users at different security levels to access systems simultaneously, without fear of information contamination.

B. Need of Multi-Level

The term multi-level arises from the defense community's security classifications: Confidential, Secret, and Top Secret.

Individuals must be granted appropriate clearances before they can see classified information. Those with Confidential clearance are only authorized to view confidential documents; they are not trusted to look at Secret or Top Secret information[6]. The rules that apply to data flow operate from lower levels to higher levels, and never the reverse. This is illustrated below.

Fig 2: Information Security Levels

 DOI : 10.35291/2454-9150.2020.0001

VII. CONCLUSION AND LIMITATIONS

Database security refers to the collective measures used to protect and secure a database or database management software from illegitimate use and malicious threats and attacks. Database security is generally planned, implemented and maintained by a database administrator and or other information security professional.

Some of the ways database security is analyzed and implemented include: Restricting unauthorized access, Load/stress testing and capacity testing of a database, Physical security of the database server and reviewing existing system for any known or unknown vulnerabilities.

Security situations arise in many everyday activities, although sometimes it can be difficult to distinguish between a security attack and an ordinary human or technological breakdown. Alas, clever attackers realize this confusion, so they may make their attack seem like a simple, random failure.

A threat is an incident that could cause harm. Vulnerability is a weakness through which harm could occur. These two problems combine: Either without the other causes no harm, but a threat exercising vulnerabilities means damage. To control such a situation, we can either block or diminish the threat, or close the vulnerability (or both). Countermeasures and controls can be applied to the data, the programs, the system, the physical devices, the communications links, the environment, and the personnel. Sometimes several controls are needed to cover a single vulnerability, but sometimes one control addresses many problems at once.

In theory, FGAC, at the level of individual tuples, can be achieved by creating an access control list for each tuple. However, this approach is not scalable (Jain, 2004) and would be totally impractical in systems with millions of tuples and thousands or millions of users, since it would require millions of access control specifications to be provided (manually) by the administrator[5]. It is also possible to create views for specific users, which allow those users access to only selected tuples of a table, but again, this approach is not scalable with large numbers of users.

REFERENCES

[1] Elisa Bertino, Fellow, IEEE, and Ravi Sandhu, Fellow, IEEE, "Database Security—Concepts, Approaches, and Challenges", IEEE transactions on dependable and secure computing, vol. 2, no. 1, January- March 2005,

[2] Simon Liu and Rick Kuhn, "Data Loss Prevention", Published by the IEEE Computer Society ©2010 IEEE

Brian M. Gaff, Thomas J. Smedinghoff, and Socheth Sor

[3] Edwards Wildman Palmer LLP "Privacy and Data security", Published by the IEEE Computer Society © 2012 IEEE.

[4] Hossein Homaei and Hamid Reza Shahriari, Amirkabir University of Technology, "Seven Years of Software Vulnerabilities: e Ebb and Flow",Copublished by the IEEE Computer and Reliability Societies, January/February 2017, © 2012 IEEE.

[5] http://www.cgisecurity.com/database/oracle/pdf/VPD9ir2twp.pdf

[6] https://web.mit.edu/rhel-doc/5/RHEL-5-manual/Deployment_Guide-en-US/sec-mls-ov.html

[7] Sohail IMRAN, "Security Issues in Databases", 2009 Second International Conference on Future Information Technology and Management Engineering

[8] Xiaolei Qian, Computer Science Laboratory, SRI International "View - Bases Access Control with High Assurance."

[9] Ravi S. Sandhu and Sushil Jajodia, " Data and Database Security and Controls", Handbook of Information Security Management, Auerbach Publishers, 1993

[10] The virtual private database in oracle9ir2: An oracle technical white paper.

[11] E.Bertino, L.M. Haas, B.G.Lindsay, View Management In Distributed Data Base Systems

[12] Surajit Chaudhuri, Raghav Kaushik, Ravi Ramamurthy, Microsoft Research, "Database Access Control & Privacy: Is There a Common Ground?"

[13] *Meg Coffin Murray*, "Database Security: What Students Need to Know", Journal of Information Technology Education Volume 9, 2010.

[14] Lakshmi, B., et al. "Data Confidentiality and Loss Prevention using Virtual Private Database." *International Journal on Computer Science and Engineering* 5.3 (2013): 143.

[15] https://www.researchgate.net/publication/331408062_Facebook_User's_Data_Security_and_Awareness_A_Literature_Review

[16] Yadav, Rajesh and Sharma, Anand, A Critical Review of Data Security in Cloud Computing Infrastructure (January 14, 2019). International Journal of Advanced Studies of Scientific Research, Volume 3, Issue 9, 2018. Available at SSRN: https://ssrn.com/abstract=3315422

[17] Sudhakar, Kumar, S. An emerging threat Fileless malware: a survey and research challenges. *Cybersecur* **3,** 1 (2020) doi:10.1186/s42400-019-0043-x

 DOI : 10.35291/2454-9150.2020.0001

[18] Tung Bui, Eric Clemons, Introduction to Information Security and Privacy in Business and Society Minitrack, Proceedings of the 52nd Hawaii International Conference on System Sciences | 2019,

[19] Pawitar Dulari, Brijender Bhushan, A Novel Approach for Cloud Data SecurityEnhancement through Cryptography and Biometricin the Government Cloud Environment,International Journal of Computer Science and Mobile Computing, Vol.8 Issue.12, December- 2019, pg. 59-63.

[20] Anil Lamba, Satinderjeet Singh, Balvinder Singh, Sivakumar Sai Rela Muni, A STUDY PAPER ON SECURITY RELATED ISSUE BEFORE ADOPTINGCLOUD COMPUTING SERVICE MODEL, International Journal For Technological Research In Engineering Volume 3, Issue 4, December-2015ISSN (Online): 2347 -4718.

AUTHORS' BIOGRAPHY

Mrs. B.Lakshmi is currently working as an Asst. Professor, Department of Computer Applications, VRSEC (Autonomous), Vijayawada, Andhra Pradesh. She has 13 years of teaching experience. Her areas of interest include Database Security, Database Management Systems, Data Warehousing and Data Mining. She had received M.C.A from Acharya Nagarjuna University and M.Tech in CSE from JNTUK, Kakinada. She has ratified under both Acharya Nagarjuna University and JNTUK, Kakinada. She had completed the OCA certification. She is a Member of CSI.

Dr. K. Nageswara Rao Garu is currently working as Principal, Potti Sriramulu Chalavadi Mallikharjuna Rao College of Engineering & Technology, Vijayawada-7. He had completed his Ph.D. in computer science and system engineering and former professor of Andhra University, Vishakhapatnam. He has an excellent academic and research experience. He has contributed a number of research papers in various National and International Journals. His area of interest includes Robotics, Data warehousing and Data Mining and Database Security.

 DOI : 10.35291/2454-9150.2020.0001