Data Protection Standards in the Business Environment

Miloš Milašinović[1,*], Nemanja Veselinović[1], Miloš Jovanović[1], Aca Aleksić[2] and Chad Ehrlich[3]

[1]Belgrade Metropolitan University, Faculty of Information Technology, Tadeuša Košćuška 63, Belgrade, Serbia
[2]Academy of Technical and Art Applied Studies, School of Electrical and Computer Engineering, Belgrade, Serbia
[3]Whitman School of Management, Syracuse University, Syracuse, NY, United States of America

Abstract

This paper considers the strategy of protecting important data from attacks caused by the external or internal environment, as well as attacks carried out using social engineering methods. Encrypting data ensures a high level of protection. When developing an encryption strategy, it is necessary to know data flows at the application level and to have access controls to keys and places to store them. The influence of cryptographic algorithms on the operation of the database is also analyzed. The applied solutions show the strategy of raising protection in databases in the business world, which today has become unavoidable in all environments. The paper presents frameworks with an emphasis on legal and technical aspects that are necessary for the company's compliance process with the GDPR. With regard to the necessary implementation of technological solutions, the paper presents suggestions for the development of application support that will be applicable during the company's compliance with the GDPR.

Keywords

Data protection, Cryptography, Encrypting, Key management, GDPR

1. Introduction

The development of information technologies today, in addition to a number of positive effects, has also caused an incredible growth of data collections on almost all individuals around the world. Billions of data are being collected in the virtual world at any moment, sold as a commodity, misused in many ways. Every individual is faced with the fact that various state institutions, private companies and individuals collect, process and use their personal data. Due to the incredible ease and speed with which data can be collected and misused, and the large number of individuals who can be affected by it, one of the most important issues of the 21st century is the issue of data protection. In professional information systems, data protection plays an important role. The goal of every organization is to protect sensitive data in databases.

Network databases are the heart of any organization. They contain business information, transaction information, financial data and customer information. The above types of data and information are often the target of attackers. A successful attack can cause huge financial losses to a company and damage its corporate reputation. More frequent targets of attacks are transactions based on the web environment, which threaten companies' credibility and customer relations. The only way to protect the company's assets is to implement certain measures and regulations for protection. Security measures include encryption of data exchanged over the network and data stored on storage devices. Another reliable method of protection is based on access control that protects data and information from insiders in the environment. Data protection strategy is important and complex. As an additional guarantee and prerequisite for realistic assessments of the level of security and the use of reliable supporting software components whose source code is available. One of the intermediate conditions for the success of the solution is its ease of use.

The ideal, in concrete solutions, is that end users feel the presence of cryptographic solutions as little as possible, i.e. that their work is not complicated and that time resources are not significantly changed when executing business processes.

Implementing protection over the database itself provides an excellent method for protecting sensitive data, but on the other hand, it leads to a decrease in performance and complicates use. This means that only some important data can be encrypted, such as e.g. credit card numbers, customer information, etc.

This paper deals with the implementation of encryption within the database, as well as the implementation of the encryption process on the application server side. These solutions offer the availability of source code, primarily the extension of available encryption algorithms with their own algorithm. Through the analysis of the performance of the mentioned solutions, the complexity of the implementation and the load on the processor are analyzed in detail.

BISEC'22: 13th International Conference on Business Information Security, December 03, 2022, Belgrade, Serbia
*Corresponding author.
milos.milasinovic@metropolitan.ac.rs (M. Milašinović);
milos.milasinovic@metropolitan.ac.rs (N. Veselinović);
milos.jovanovic@metropolitan.ac.rs (M. Jovanović);
milos.milasinovic@metropolitan.ac.rs (A. Aleksić);
cdehrlic@gmail.com (C. Ehrlich)

2. Data protection by Encrypting

The goal of any business environment is to protect the data of its clients. Due to the large number of places where sensitive data is located within a business environment, it is necessary to protect archived data, which are located on employees' computers or on some long media.

When developing a strategy, in order to be sure that the data will be protected, we must fulfill two basic conditions. The first is the use of reliable cryptographic mechanisms, and the second is the precisely defined management of access to cryptographic keys. Only by fulfilling these conditions, the protection of confidential data can be secure [1].

The next step is to create a strategy for the encryption process, and for its implementation it is important to do the following:

- decide whether the encryption process is performed inside or outside the database
- precisely optimize the system in order to reduce the number of users who have access to cryptographic keys
- determine a safe place to store cryptographic keys
- separate the ciphers from the cryptographic keys in (in which case the ciphers in the hands of the attacker are completely useless),
- strike a balance between performance and application

A precise policy of access control and management of cryptographic keys builds a high level of trust and control over the information infrastructure [1].

Access to decrypted data should be strictly limited by access control and users who have access to such data must be monitored by a system for monitoring and analyzing log files.

2.1. Planning Data Encryption Strategy

Before starting to design an encryption strategy in databases it is necessary to understand:

- how cryptographic mechanisms work
- what the data flows look like in the application i
- how database protection fits into the organization's security policy

So there are two strategies:

- use of functional encryption mechanisms within the database (DBMS)
- use of functional application-side encryption mechanisms

Each of these approaches has its advantages and disadvantages. It is important to make a good overview of both, and depending on the organizational infrastructure, choose the one that best fits the functioning of the organization [2].

2.2. Social Engineering as an Essential Factor in Strategy Development

creating a database protection strategy. Without strong identity verification, application-side technologies open the way to access to data decryption mechanisms.

A big problem is database administrators and business application developers who know how to access that information, as well as disgruntled employees who find a way through social engineering to access passwords. Many employees have compromised their companies by revealing monthly earnings, social security numbers, cell phone numbers, information about external partners, and more [2].

Database protection is not only the protection of important data, but also the protection of employees from this type of exploitation. You should think about developing a security strategy in time, primarily the protection of important data, such as encryption, access control, event monitoring and log monitoring [3].

2.3. Impact of Encrypted Data on the Database

Encryption provides a high level of protection and is accepted as a best practice in restrictions for protecting important information, but it has an impact on data and the database.

Encryption increases the amount of data and reduces the performance of the application. When planning the development of the application, it is important to anticipate possible impacts caused by the use of encryption systems.

By knowing which data should be protected, we get more flexibility and better performance. When processing a credit card, we can encrypt the entire number or just the last four digits [4].

Based on the knowledge of the amount of data that needs to be protected, the application is designed. Encryption affects data size. The algorithm can often change the size of the fixed data from the database, so that the data is not saved. Especially the encryption of some small data can greatly increase its size, and this leads to an increase in the column sizes of the relational database [4].

Encryption is a process in which we transform characters into binary meaningless strings that can affect the size if we transform the encrypted data into characters.

By using the BASE64 algorithm, we can convert the encrypted data into a string of characters, but this would

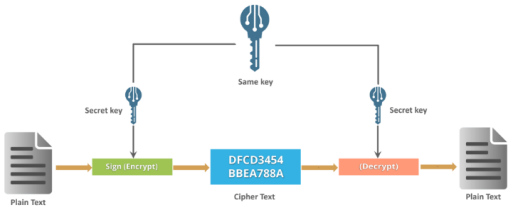

Figure 1: Cryptography.

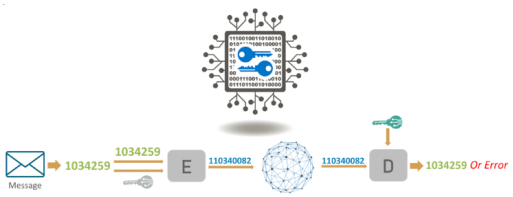

Figure 2: The most popular symmetric-key system - DES.

increase the size of the data by one third. When selecting the fields in the database tables, which will be encrypted, it is important to take care of the search process.

If every search, in large databases, is followed by a process of encryption and decryption, it can degrade the performance of applications, as well as their functioning, encryption can complicate the initial process and cause frustration for the user at the beginning, and this requires hard new considerations and planning around the selection encryption fields [5].

3. Cryptography and Cryptographic Algorithms

Data, in order to reach the highest level of security, must be stored in an encrypted form. The goal of encryption is to make them unreadable by unauthorized readers and protect them from decryption in the event of an attack.

The encryption operation is performed with randomly generated cryptographic keys. Which make the cipher secure and difficult to attack or decipher. Keys are usually stored in encrypted form [6].

Not all encryption averages are the same. The security of encrypted data depends on various factors such as the length of the cryptological key, and the algorithm and the way it is implemented in the system used.

Many databases use AES and DES algorithm to protect fields with important data, but DES has long been considered insecure for protection. When choosing a cryptographic algorithm, it is important to find a reliable supplier of commercial algorithms or decide on your own solution [6].

Apart from the quality of the cryptographic algorithm, a quality strategy for managing cryptographic keys is also very important. In such systems there is a constant tension in the spheres of protection between control and access. In any case, the system must have access to the keys to decrypt something, and their distribution is complicated and leads to a decrease in performance and sometimes security [7].

Administrators and developers are at a loss if they implement such security measures.

3.1. Cryptographic Key Management Strategy

Cryptography is based on cryptographic keys, which depend on the storage of encrypted data in databases. They have secret places where they are archived and who can access them. Cryptographic key management control is generally envisaged when developing a strategy for database protection [8].

The most important roles in planning a key control strategy:

- the number of keys required for encryption
- the way they are used
- the place where they are stored
- protected and limited access to them
- change of keys

4. Required Number of Keys and Modes of Management

The problem of cryptographic key management is difficult to solve. Using the same key in an application is easy to implement and maintain, but if that key becomes compromised, the data will become vulnerable.

If there are a large number of users, requiring access to different applications, the scenario becomes more complex because encrypted data can only be decrypted with the appropriate key, and each system or application must have access to the appropriate key [8].

The more applications that have access to the keys, the greater the risk of attack, so they should be kept in a secure management system. A good practice is to use a smaller number of encryption keys, it is a simpler key management solution, but because of this the critical point becomes the key [9].

4.1. Safe Place to Keep Keys

When developing a key control strategy, consider a safe place to store the keys. One solution is to store it in a separate database or in a file with strictly limited access. With such solutions, the problem is administrators who have authorized access to the keys and can decrypt the

data and then cover their tracks. In this case, the security of the database does not depend on best practices, but on the honesty and honor of the employees.

To reduce risk and have better control, it is best if only one person has access to the keys. The next possibility to consider is to separate the keys from the database and store them in an encrypted form on some hardware devices. It is important that the key cannot leave the hardware device, that access is controlled, and that neither the administrator nor the attacker can get to it and take it [9].

The security of keys depends on good management and a place to store them. According to a 2005 survey, one-third of organizations store their sensitive data with keys.

Without a strategy for separating database keys and storing them in secure places with limited access, the entire system is at great risk to many users, developers, and database administrators.

Systems that do not have this strategy survive due to the quality of the applied authentication methods for identity verification and access control [10].

5. Data protection Law

Along with the development of theories about the understanding of privacy and the protection of the right to privacy before the courts in Europe, from the seventies of the twentieth century until today, legal regulations were developed that regulate the issue of data protection and create a complex system of data protection.

In the beginning, these were the legal regulations of certain European countries, and then the main role in the creation of the data protection system was taken by the European Union [11].

The legal system of data protection in Europe, based on the Data Protection Directive from 1995, was not sufficient to adequately respond to the accelerated development of technology and numerous ways of inadequate collection and processing of personal data, and for this reason, in January 2012, a legislative procedure for the adoption of a regulation that would comprehensively regulate issues of the right to data protection.

After a long public discussion and harmonization of numerous proposals and consideration of over 4,000 amendments, the text of the General Data Protection Regulation was agreed and adopted in April 2016. This regulation came into force on May 25, 2018 [12]. The GDPR regulation provides individuals with greater control over personal data and imposes many obligations on companies that collect and analyze personal data. The regulation under private data includes all those data from which the identity of a person can be determined, then data on political, sexual orientation, race, property status, but also data such as search history, meta data on files created by a natural person, which they can reveal identity, information about health, movement, habits [12].

Personal data means any combination of personal facts that accurately determines an individual, i.e., among others, first and last name, personal identification number, location data, physical, physiological, genetic, mental, economic, social, cultural or any other factors.

The GDPR regulation on the management of personal data is based on the following principles:

- Legal and transparent processing of personal data of European Union residents is required.
- Collection of personal data for specific, precisely defined purposes. The company must not process data in a way that is inconsistent with the stated purpose. Storing only minimally necessary data.
- The company does not have the authority to ask the individual for additional information that goes beyond the defined purpose of data collection and processing.
- The company is not authorized to keep data longer than defined.
- It is necessary for the company to implement certain security measures, • Personal data must be accurate and updated.
- In an adequate way, it is necessary to ensure the corresponding level of security of personal data, including protection against unauthorized processing. It is necessary to ensure data confidentiality.

The GDPR regulation requires a review of business models in companies that offer their goods and services to natural persons resident in the European Union [13].

A company that collects personal data before processing must obtain the consent of the individual for their use, and it is necessary that the natural person who gives consent be informed about the purpose of collecting personal data. One of the simpler ways of adapting a company to the GDPR regulation is by engaging third parties - agencies that have legal and technical knowledge for security and data protection.

They will be able to adequately and professionally fulfill their obligations to companies in terms of efficient privacy protection [13].

6. Application of Application Support for GDPR Implementation

In order to improve the protection of personal data, companies will improve their IT infrastructure, implementing

the necessary protective measures for the process of physical control of access to data, storage and the process of data archiving.

The regulation mentions the need for archiving and data access records. This will lead companies to focus on data protection processes. The necessary alignment of domestic companies with the GDPR regulation will increase market demands for specialized application support [14]

The Regulation does not explicitly state which technology must be used for data protection activities. The Regulation mentions the mandatory continuous monitoring of personal data and the necessary activities of search, management, protection, monitoring of personal data and reporting activities.

The technology that would help companies in the GDPR compliance process can be arbitrary. The first and basic phase of the implementation of the GDPR regulation is the phase of searching existing personal data.

It is necessary to enable the search of the competent scope of personal data. In addition, it is necessary to determine identifiers that will clearly define personal data, such as: first name, last name, email, address, telephone, JMBG, etc. After determining these identifiers of personal data, the data to be searched will be divided into structured and unstructured . Structured data is found in tables in which names, surnames, JMBG numbers, etc. are structurally listed [14].

It is necessary to search complete databases in order to separate those files that contain structured data with defined identifiers of personal data. Application support will help companies to find as many personal data as possible based on certain search rules. Once the data is determined, then it needs to be classified.

Personal data can be marked by type (structured/unstructured), by secrecy (public data, secret data). Every user of the personal data base should get access to this tool. He will manually or mechanically classify data (files, office documents, photos, video, audio, email).

The classification of data will make it possible to determine the types of data processed by business processes within the company: whether the data is personal data at all, whether it needs encryption, a digital signature, and the like. In order to reveal the largest percentage of personal data that the company has in its databases, it is necessary to define - what all needs to be searched [15].

The search tool should be able to search photos, as an important identifier of personal data, as well as data from social networks, various online identifiers: such as user logins, MAC data, and GPS data about the user's geolocations. In order for the data search system to be able to find a larger percentage of personal data, it is necessary to define a larger spectrum of identifiers that the GDPR regulation recognizes as personal data. It is necessary to define the method of data search. It is possible to search structured data in which it is easier to recognize identifiers such as: first name, last name, phone number, JMBG, etc.

These data are found in databases such as accounting ERP systems or customer relationship management systems CRM. Within the organization, it is necessary to define that both IP and MAC addresses are also personal data. In order for the search system to be able to work with such data, it is necessary to note that these data have control digits and from them it can be concluded that a series of numbers (JMBG, MAC address, etc.) belongs to personal data. It is suggested to use machine learning systems, because they will independently recognize within similar documents whether something is specific to the organization [15].

Thus, the optimal application support system will be able to recognize what is personal data from a large number of documents and contracts. It is necessary to use application support with search mechanisms in order for the company to determine where personal data is located.

It is necessary to search all hardware data storage units, as it is possible that copies of personal data are located on multiple storage units. Structured data will be easier to find, but for unstructured data it is necessary to develop new search algorithms within e-mails or the cloud [15].

7. Conclusion

Attacks on databases, from year to year, are on the rise, which increases the risk of compromise. Today, financial service providers and healthcare organizations must strictly comply with data privacy laws.

Users of these services due to concerns about data disclosure or misuse will inevitably expand the responsibilities of each organization to provide certain services.

The negative effects, caused as a result of the attack, would be reflected in the form of legal liability, negative publicity, lost trust in the public, and loss of money and productivity. In order to avoid the mentioned negative effects in our environment, it is inevitable to plan an encryption strategy in databases to protect important data and a strategy against attacks and other abuses.

Implementers of application support will have procedures for compliance with the GDPR regulation. However, users will be required to clearly define changes in business processes. Companies must be aware that IT firms will not take the risk and responsibility on themselves. IT companies will proceed from the assumptions that they avoid taking risks due to large fines and due to often unclear points within the GDPR regulation.

IT companies will implement application support according to the requests received from clients and thus avoid the risk related to the GDPR regulation. Therefore, it is important that company managers undergo

the necessary training in order to be able to implement compliance with the GDPR regulation. An IT company that develops application support must be aware that its clients are subject to the GDPR and that the very fact that companies use their application support to enter personal data creates an obligation for them to comply with the GDPR regulation.

Some of the principles that the application support must adhere to are: it is necessary to implement the necessary settings for the protection of personal data in the application support, as well as the possibility of competent deletion of personal data. It is necessary that all interfaces to third-party systems comply with the rules of personal data protection.

It is necessary to simplify the process of changing customers' personal data, in such a way that from one place the change is reflected in all places in the company's databases. It is necessary to offer a printout of the Consent to be signed by the respondents in the application. The process of user logging into the system is also changing. It is important that each user accesses the system with his or her own unique user login that only he or she uses.

References

[1] M. E. Smid, D. K. Branstad, Data encryption standard: past and future, Proceedings of the IEEE 76 (1988) 550–559.

[2] R. Davis, The data encryption standard in perspective, IEEE Communications Society Magazine 16 (1978) 5–9.

[3] D. Coppersmith, The data encryption standard (des) and its strength against attacks, IBM journal of research and development 38 (1994) 243–250.

[4] D. E. Standard, et al., Data encryption standard, Federal Information Processing Standards Publication 112 (1999).

[5] E. F. Schaefer, A simplified data encryption standard algorithm, Cryptologia 20 (1996) 77–84.

[6] G. C. Kessler, An overview of cryptography, 2003.

[7] D. Dolev, C. Dwork, M. Naor, Non-malleable cryptography, in: Proceedings of the twenty-third annual ACM symposium on Theory of computing, 1991, pp. 542–552.

[8] W. Diffie, M. E. Hellman, New directions in cryptography, in: Democratizing Cryptography: The Work of Whitfield Diffie and Martin Hellman, 2022, pp. 365–390.

[9] K.-H. Lee, P.-L. Chiu, An extended visual cryptography algorithm for general access structures, IEEE transactions on information forensics and security 7 (2011) 219–229.

[10] O. G. Abood, S. K. Guirguis, A survey on cryptography algorithms, International Journal of Scientific and Research Publications 8 (2018) 495–516.

[11] K. Hjerppe, J. Ruohonen, V. Leppänen, The general data protection regulation: Requirements, architectures, and constraints, in: 2019 IEEE 27th International Requirements Engineering Conference (RE), IEEE, 2019, pp. 265–275.

[12] M. Goddard, The eu general data protection regulation (gdpr): European regulation that has a global impact, International Journal of Market Research 59 (2017) 703–705.

[13] C. Tankard, What the gdpr means for businesses, Network Security 2016 (2016) 5–8.

[14] J. P. Albrecht, How the gdpr will change the world, Eur. Data Prot. L. Rev. 2 (2016) 287.

[15] W. B. Tesfay, P. Hofmann, T. Nakamura, S. Kiyomoto, J. Serna, Privacyguide: towards an implementation of the eu gdpr on internet privacy policy evaluation, in: Proceedings of the Fourth ACM International Workshop on Security and Privacy Analytics, 2018, pp. 15–21.