

The Bucharest University of Economic Studies

The Faculty of Cybernetics, Statistics and Economic Informatics

Economic Informatics Specialization

Secure Blockchain-Based Application for Electronic Health Records

Bachelor Thesis

Advisor:

Prof. univ. dr. Răzvan Daniel Zota

Graduate:

Lupșan Sabrina

Bucharest 2021

Contents

[1. Introduction - 2 -](#_Toc53266857)

[2. Blockchain - 2 -](#_Toc53266858)

[a. What is Blockchain - 2 -](#_Toc53266859)

[b. Types of Blockchain - 2 -](#_Toc53266860)

[c. Why private blockchain - 2 -](#_Toc53266861)

[3. Cybersecurity - 2 -](#_Toc53266862)

[a. The rise of threaths - 2 -](#_Toc53266863)

[b. Integrity with hashing - 2 -](#_Toc53266864)

[c. Safe authentication - 2 -](#_Toc53266865)

[d. Encryption - 2 -](#_Toc53266866)

[e. SQL Injection - 2 -](#_Toc53266867)

[4. Windows Application - 2 -](#_Toc53266868)

[a. .NET Framework and Windows Forms platform - 2 -](#_Toc53266869)

[b. C# Language - 2 -](#_Toc53266870)

[c. DB Storage chapter - 2 -](#_Toc53266871)

[5. Using Blockchain to create, view and manage EHRs - 2 -](#_Toc53266872)

[6. Conclusions and future work - 2 -](#_Toc53266873)

[a. Conclusions - 2 -](#_Toc53266874)

[b. Future work - 2 -](#_Toc53266875)

[7. Bibliography - 2 -](#_Toc53266876)

[8. Annexes - 2 -](#_Toc53266877)

## Introduction

## Blockchain

## What is Blockchain

Blockchain is a technology.

## Types of Blockchain: a comparison

There are two types of Blockchain: public and private. Both types serve different needs and bring different measures of security. There is a big debate on which blockchain it is safer and, after analyzing the characteristics of both, I draw some conclusions.

Firstly, the problem of decentralization of data: in a public blockchain, there data is not stored on a server, in a single place, but rather everywhere: anyone that follows the rules imposed for the blockchain they want to contribute to can step in and verify, add and read the data on the blockchain. This is the case for Bitcoin. On the other hand, in a private blockchain, the decentralization cannot be kept. However, only trusted entities can participate in the manipulation of the blockchain. This capability of the private blockchain increases access control, because not anyone can participate in the actions specific to a blockchain, but introduces the vulnerability of being hacked.

Secondly, a public blockchain’s security increases as the number of peers that take part in it increases. This is a problem of scalability, since the transactions can only occur in a slow pace. This happens because every time an entity wants to contribute with a block, it must be approved by everyone in the blockchain. This can severely impact performance, which is not the case for a private blockchain. In the permissioned one, the number of authorized nodes is much smaller and therefore the data is processed much faster. (Sharma)

Lastly, a private blockchain is more prone to hacking. If a hacker would get access to a trusted entity’s credentials, the vulnerability created would be critical. This raises many security problems that are not to be found in the unpermissioned version. The latter one can only be attacked if over 51% of the participants in the transactions would have the same ill-intention, all blocks would be tampered with and the proof of work would be completely redone for the entire blockchain. (Blockchain Tutorial for Beginners: Learn Blockchain Technology)

## Why private blockchain

After considering the advantages and disadvantages of both permissioned and unpermissioned blockchains, I have decided that the private blockchain would be more fit for the application presented in this paper.

One capability I considered is the invitation-only aspect of the private variant, which allows the trusted entities to keep their data private in only one institution. This type of technology permits an organization to only include its employees, suppliers and clients and to keep the data off of the internet. Otherwise, in a public environment, the transactions would not be kept private. Acknowledging this aspect, a private blockchain is a secure tool to manage sensitive data in an enterprise.

There is also the possibility of implementing different levels of access and a different set of transactions for entrants. For example, in a company, an employee would have different rights and needs than a client. (Heath, 2018) This can be achieved by creating different profiles and assigning them correctly to participants.

Another decisional factor in differentiating the two blockchains is the cost of the technology. In an unpermissioned implementation, validation and proof of work are essential. Validating transactions is time-costly because every entrant is considered untrusted. In the private blockchain this is exactly the opposite, as the contributors to transactions are already trusted nodes, thus speeding up the process of accessing the blocks.

The biggest draw-back of the private blockchain is the fact that it is credentials-based, which means that anyone that has access to credentials of a trusted person (that can be obtained through cyber security attacks like social engineering, phishing and others) or gains the trust of the already-existing participants can read and add to the blocks.

Taking into accound the risks and the benefits of this type of private blockchains, many companies have adopted this technology. Walmart, Spotify, DHL are just a few examples of the giants that acquired a permissioned blockchain. (Euromoney)

## Cybersecurity

## The rise of threaths

## The CIA and AAA models

After analyzing the continuous rise of threats and the various types of complex attacks that have been developed, cyber security provides two models to guide professionals. The CIA model represents the principle that an IT component should provide the following characteristics: Confidentiality, Integrity and Availability. The AAA model provides the means to achieve the goals of cyber security described in the CIA mode. The triple-A concept refers to Authentication, Authorization and Accounting. (Nweke, 2017)

**CIA Model**

The first goal in the CIA model expresses the need for privacy of the data. Protecting data over the internet is crucial and the first step in a hacker’s malicious plans includes gathering sensitive information about their victim. Therefore, an attack can be stopped from the reconnaissance stage if an attacker cannot gain confidential information. This element of the model can be assured using data encryption and access control, among others.

Integrity assures that the original data is not tampered with by a third party. This plays the part in not allowing someone unauthorized to alter or delete information. Integrity can be achieved through hashing, which is the process of introducing data in a hash function that produces a unique output for every input.

The last principle in the CIA model is Availability. This component assumes that the owner of their data can access it and the needed resources whenever they want. DoS or DDoS attacks are built to bring down a system and block the users from accessing the needed resources. Availability can be provided when a system maintains reduncancy, fault tolerance, access lists etc. (CYBER EDU)

**AAA Model**

## Integrity with hashing

A cryptographic hash function is a mathematical algorithm which takes a variable-sized input and generated a fixed-size output which represents the hash of that data. There are a few characteristics that enable this function to be used for integrity checks.

This function is irreversible, which means that, only knowing the output of the operation, the input cannot be discovered. Another characteristic is that even if only one bit from the original input is changed, the entire result is significantly modified. Using this property it is very easy to detect any alteration in the existing data. Furthermore, the hash function will always generate the same result for the same input (that means that the hash value will not change unless the data is changed). This mathematical algorithm is also collision ressistant, meaning that no two values lead to the same hash value. (Synopsys Editorial Team, 2015)

Knowing these properties, integrity over the internet can be achieved using a cryptographic hash function. This algorithm is only vulnerable to brute-force attacks. This type of attack represents repeated attemps to guess the input by comparing its hash value to the hash value of every attempted guess. A brute-force attack required big computing power and a lot of time. The cryptographic hash algorithms have improved over time in order to make it computationally infesible to try to guess passwords.

Moreover, an enhancement named “salting” can be used to further increase the time needed to guess a value. The process of salting involves combining another string, like the username or the email of an account, in the initial input of the function. This is effective because it prevents the hacking of the most common passwords, like “123456” or „password”. (Picheta, 2019) If the user “John” would log in with the password “123”, the input data for the hash function could be “Jo123hn”. This is significantly harder to guess, thus increasing the protection from brute-force attacks. (M., 2016)

## Safe authentication

## Encryption

## SQL Injection

SQL Injection is a common web hacking technique that exploits input requests by writing malicious SQL code with the purpose to unknowingly run SQL statements on a database. Such attempts to inject code can modify or delete database data, can read sensitive information or even shutdown a DBMS. The ill-intended code can be designed to change the purpose of the affected SQL querry, cause an error or delay a command.

A more “traditional” way of dealing with this type of vulnerability is placing validation policies on the input received from users. This can mean not accepting meta characters, limiting the amount of data accepted or limiting the user to a set of allowed values. However, at some point intruders will find a way around that, since there are many possible ways to do it. A more versatile and secure way of protecting your application from SQL Injection is the use of Prepared Statements (or parameterized statements). A prepared statement is a stored procedure that doesn’t concatenate the query string and compiles it, but rather keeps the command compiled and executes the statement every time. (SQL Injection, n.d.)This is also efficient because the statement is only compiled once, so the overhead decreases. (Prepared statements and stored procedures ¶, n.d.) It is also more resilient to SQL injection because it uses placeholders in the query string and every parameter is checked if it is correct and if its type corresponds to the database column type first. (SQL Injection, n.d.)

## Proof of work and DOS

## Windows Application

## .NET Framework and Windows Forms platform

.NET is a platform built for developing different types of applications. It is open-source, cross-platform, free and can be used with different editors in different languages. It was released by Microsoft in 2002 and has reached over 3700 companies and 60.000 developers. .NET applications can be written in C#, F# or Visual Basic. Compiled code is stored in files which are called assemblies and are files with .dll or .exe extensions. A widely used tool for developing .NET applications is Visual Studio.

.NET Framework is used for building and running applications on Windows. Is is a part of the .NET platform and is its original implementation. Besides desktop apps, it supports other services and websites. There are two components that make up the .NET Framework which are: the Common Language Runtime (CLR) and the Class Library.

The CLR is the engine that executes and handles the running applications. It delivers many useful services, including: exception handling, garbage collection, thread management and others. The Class Library is a set of APIs for writing and reading files, drawing, connectiong to databases etc. (.NET Framework documentation, n.d.)

Windows Forms is a User Interface framework designed to build Windows desktop applications. With this technology, the graphical construction of the application is easy to deploy and update, can be worked on while online or offline and brings many functionality like drag and drops and print previews. Windows Forms carries many different controls, from the most used ones like buttons, textboxes and date pickers to drop-down boxes, contextual menus and error providers. The flexibility of this platform also allows developers to create and design their own controls as well as drawing pie charts, histograms, etc. (Desktop Guide (Windows Forms .NET), 2020)

## C# Programming Language

C# is an object-oriented programming language rooted in the C family of languages. It is a modern programming language that provides garbage collection, exception handling, lambda expressions, asynchronous operations and many more. (A tour of the C# language, 2020)

## Microsoft Azure with SQl

## Using Blockchain to create, view and manage EHRs

## Conclusions and future work

## Conclusions

## Future work

## Bibliography

## Annexes

# References

*.NET Framework documentation*. (n.d.). Retrieved from Microsoft: https://docs.microsoft.com/en-gb/dotnet/framework/

*A tour of the C# language*. (2020, 08 06). Retrieved from Microsoft: https://docs.microsoft.com/en-us/dotnet/csharp/tour-of-csharp/

*Blockchain Tutorial for Beginners: Learn Blockchain Technology.* (n.d.). Retrieved from Guru99: https://www.guru99.com/blockchain-tutorial.html

CYBER EDU. (n.d.). *What is the CIA Triad?* Retrieved from Forcepoint: https://www.forcepoint.com/cyber-edu/cia-triad

*Desktop Guide (Windows Forms .NET)*. (2020, 10 26). Retrieved from Microsoft: https://docs.microsoft.com/en-us/dotnet/desktop/winforms/overview/?view=netdesktop-5.0

Euromoney. (n.d.). *The rise of private blockchains.* Retrieved from Euromoney: https://www.euromoney.com/learning/blockchain-explained/the-rise-of-private-blockchains

Heath, N. (2018, September 05). *What’s the difference between a private and public blockchain?* Retrieved from In The Black: https://www.intheblack.com/articles/2018/09/05/difference-between-private-public-blockchain

M., I. (2016, October 27). *Three random words or #thinkrandom.* Retrieved from NCSC UK: https://www.ncsc.gov.uk/blog-post/three-random-words-or-thinkrandom-0

Nweke, L. O. (2017). Using the CIA and AAA Models to Explain Cybersecurity Activities. *PM World Journal*, 3.

Picheta, R. (2019, April 23). *How hackable is your password?* Retrieved from CNN: https://edition.cnn.com/2019/04/22/uk/most-common-passwords-scli-gbr-intl/index.html

*Prepared statements and stored procedures ¶*. (n.d.). Retrieved from php.net: https://www.php.net/manual/en/pdo.prepared-statements.php

Sharma, T. K. (n.d.). *PUBLIC VS. PRIVATE BLOCKCHAIN : A COMPREHENSIVE COMPARISON.* Retrieved from Blockchain Council: https://www.blockchain-council.org/blockchain/public-vs-private-blockchain-a-comprehensive-comparison/

*SQL Injection*. (n.d.). Retrieved from Owasp: https://owasp.org/www-community/attacks/SQL\_Injection

*SQL Injection*. (n.d.). Retrieved from W3schools.com: https://www.w3schools.com/sql/sql\_injection.asp

Synopsys Editorial Team. (2015, December 10). *What are cryptographic hash functions?* Retrieved from Synopsys: https://www.synopsys.com/blogs/software-security/cryptographic-hash-functions/