Cryptanalysis of Some Attacks Applied on a Blockchain System

Lina Lumburovska, Vesna Dimitrova
Ss. Cyril and Methodius University of Skopje, Faculty of Computer Science and Engineering, Skopje, North Macedonia
lina.lumburovska@students.finki.ukim.mk
vesna.dimitrova@finki.ukim.mk

***Abstract* — The need for blockchain technology is growing constantly and its advantages are applied in different fields. One of the most sophisticated applications is the electronic voting system. Building an e-voting system based on blockchain gives a decentralized system. This research gives an overview which attacks can happen to such a system which is based on blockchain, keeping in mind that this system exists. It appears that attacks based on mathematical fields can happen such as: correlational, algebraic and birthday attacks. We analyse their performances and give an overview how each attack can happen and how to avoid it from happening. When choosing the parameters for the electronic voting system we must be careful of the presence of various attacks. This analysis showed that all processed attacks that have a mathematical basis can occur under different circumstances. They may violate voter anonymity, as in the case of correlation and birthday attacks, or they may misuse voting on behalf of others, as in the case of algebraic attacks. Fast responses, observation, correct cryptanalysis and selection of well- examined parameters are the main factors that can contribute to reliable communication.**

***Keywords* — blockchain technology, cryptanalysis, birthday attack, correlational attack, algebraic attack.**

I. INTRODUCTION

The term cryptology can be explained as a scientific field for the hidden word. As part of it, the term cryptography is defined a scientific field that aims to find and create methods for encrypting data and the rest part of the cryptology is the field of cryptanalysis, which means that it is defined a scientific field which deals with decryption of encrypted data and analysis of information systems in order to understand the hidden meanings of messages. Cryptography and cryptanalysis are closely related and often combined for different purposes. Cryptanalysis is used to encrypt cryptographic security systems to gain access to the contents of encrypted messages, even when the key is not known. In this paper, we give an overview of the cryptanalysis of some attacks on a blockchain system. As one example of a blockchain system where we did the cryptanalysis of the attacks is an electronic voting system. The attacks that we included have something in common and that is that are all based on mathematics, which means that they use different mathematical methods so they can break the cipher [1].

Decentralized systems reduce the risk for corruption, fraud and manipulation. The key concept in these systems is to build one network with more elements where each element is equally treated, has the same permissions and fraud can be less caused because each element is part of every process that is being executed in the system. Additionally, each element can contribute equally to the system and its opinion is taken into consideration. All elements build a network, a chain, they are called blocks and the technology is blockchain. In our specific case we took the electronic voting system as one blockchain system and each block is one vote in the system [2].

According to the requirements, each blockchain technology should contain four elements:

- Peer-to-peer network

The structure of the blockchain technology is basically the same as the Internet. It contains a peer-to-peer network, a network of computers, known as nodes that has the same privileges and permissions. The network is accessible to anyone and everyone, it is open and can be accessed anytime. The need for this type of network is huge, because the nodes can communicate and share information with each other remotely. With the equal treating the network is more stable and the anonymity is risen to a higher level since every element provides the same contribution to the network [3].

- Cryptography

The part of the blockchain's structure where there is a need to try to reach the highest security level that is possible is the cryptography part. This is the second element of the technology and must provide the technology with secure communication, verification of the messages, authorisation and authentication. Additional research should be done which cryptographic algorithm is the best for a certain situation. There is no right or wrong cryptographic algorithm, but it depends on the given circumstances. The main goal is to catch the "bad actors" in the chain since everyone has the access and reduce the possible damage in the system. In this step, the cryptographic algorithm uses pair of keys - public and private [3].

- Consensus Algorithm

The addition of a new block in the system must be done according to a predefined rule. This specific rule must be accepted by everyone, otherwise it will not be counted as valid. There are many type of consensus rules, but the most common are: proof of importance, proof of work, proof of activity and proof of stake. Blockchain uses the principe proof of work which means that the new node has to do some work in order to prove that it is ready and suitable for the chain. As soon as the node does the proof of work correctly, it can be added to the system because it is counted as a suitable node that can do its work. In blockchain each new node has to find a solution to a mathematical problem which requires computational power to solve [3].

- Punishment and Reward

After the node gets the rule and it provides an answer to the problem, it gets either punishment or reward. The reward is a token, a coin which is awarded every time a solution has been reached and a new block is being added to the network. On the other hand, the bad actors who try to manipulate the system, will loose their money, their coins will be taken and they will get a punishment. This is a simple decision and if the node successfully finished the proof of work, it can be added and accepted to the system [3].

Blockchain applications and systems can issue attacks. The idea is to find which types of attack are most likely to happen, so it will be easily to prevent them in time. Having in mind the attacks based on mathematical basis, and a build blockchain system, we are reaching and to trying to find if it is possible that such attacks can happen to the system and how will they have an impact to it. All attacks use different mathematical methods and each of them can occur in different circumstances. No matter what the attack is, there is one general rule that means that participants in communication must use secure and credible resources in order to protect themselves from the outside world as much as possible. In case the same cannot be done, it is necessary to respond quickly to the existing problem.

The structure of the paper is organized as it starts with a brief introduction to the topic, then the second sections is an explanation how the attacks that we take in the research work. The third section is our research question, where we explain how those attacks work on a blockchain system. Our methodology process was separated into three phases, where in the first phase we analyzed the attacks, how they work, and what are their performances and vulnerability. In the second phase we took one blockchain system (which we had built in a separate research) and we analyzed how these attacks have impact on our system. In the last phase we did the analysis of the whole work and came up with conclusions. The paper finishes with a conclusion of the whole research work.

II. ATTACKS BASED ON MATHEMATICS

A. Correlational attacks

Correlation attacks are attacks that occur in symmetric cryptography. They can only happen on stream ciphers, as one part of the symmetric cryptography and are caused by combining the output of several register shifts from the so-called LFSR (linear feedback shift register) registers using a Boolean function. An LFSR is a register whose input bit is linearly dependent on the previous bit. These registers function by means of a polynomial called characteristic polynomial or feedback polynomial. A polynomial can be defined as a designated predefined set that is represented as a polynomial by module 2. The weakness of the code stems from the poor choice of the Boolean function, and correlation attacks use exactly that weakness in order to break the code. The advantage of the code before this attack can be made when selecting the repetition period of the key, i.e if the period is longer and there are fewer repetitions, the code will be more secure before this type of attack. Another advantage is that the user defined the Boolean function and having in mind these rules, it can easily be defined so it will reduce the possibility for such attack [4].

Correlation attacks are most likely to occur when there is a correlation between the output of one register and the output of the Boolean function that combines the output of all registers. In this way a partial knowledge of the main course occurs and allows the attacker to find the key using the brute-force attack. by combining all possible options. In this paper, these types of attacks are examined with which register there is the greatest correlation when using different Boolean functions, because the choice of function is the main key factor that can violate the security of the code [4].

B. Algebraic attacks

Algebraic attacks are the next mathematically based attacks that occur on stream ciphers in symmetric cryptography. The similarity of both mentioned attacks is that both can happen in symmetric cryptography and on stream ciphers, not on block ciphers. The target they have is the same as for correlation attacks, but the way how they work is completely different. The idea behind these types of attacks is to find the key by solving a system of nonlinear equations. This type of problem is an NP difficult problem even when all equations have degree 2. The first way to solve this type of equation is with the help of linearization and it can be performed in polynomial time only if there are enough linearly independent equations. Another way to solve a system of nonlinear equations is to use Gröbner databases. Solving this type of equation is very difficult, however if successful it reveals the key and the password is no longer secure. A big role in the discovery of the key is the time performance, i.e the solution of the system depends on the available resources [5].

In recent years, algebraic attack has been enhanced by the introduction of the so-called fast algebraic attack, which involves an additional step in solving the system of nonlinear equations. This step is called the pre-calculation step and in it the degree of the system of nonlinear equations should be reduced as much as possible. Research shows that the sum of the time before the calculation and the time from the calculation itself, in most cases is less than when there is no step before the calculation. In the step before calculation, the coefficients for reducing the maximum degrees in the system of nonlinear equations should be found. Another advantage of using a pre-computation step is that less data is stored in the resource memory due to running in two independent parts. Correlation attacks stem from the weakness of Boolean function, and that is why that connection was examined. In contrast, algebraic attacks originate from solving the system of nonlinear attacks, ie whether it is possible to solve and if there are sufficient resources to perform it in a possible time [5].

C. *Birthday attacks*

The last attack in this paper that is based on mathematics is a birthday attack. This attack comes from the equally named - birthday problem in mathematics. The birthday problem in mathematics is in a room with *n* people, what is the probability of finding two people who will have a birthday on the same day. The problem is that if there are more people in the room, such an example is more likely to be found. It comes directly from the number of days in the year and if there are more than 365 people in the room, the probability of finding such a couple with a birthday on the same day will be 100%. On the other hand, the minimum value is if there is only one person in the room, i.e. the probability of finding a person with a birthday on the same day will be 0%. The existence of a leap year is not included in this paper. There are several methods to calculate the probability of having a person who would have a birthday on the same day. This attack functions similarly to the attack of brute force, i.e. tests all possible options for finding two identical matches. In this paper, an analysis is made of the probability that two different people in the room will find two identical birthdays. This analysis uses Poisson distribution as a method for calculating probability [6, 7, 8].

III. RESEARH QUESTION AND CRYPTANALYSIS OF ATTACKS BASED ON MATHEMATICS ON A BLOCKCHAIN SYSTEM

A. *Correlational attacks*

As an example of a blockchain system in our research, we chose an electronic voting system. The correlation attack uses the LFSR registers using the Boolean function where each input bit is linearly dependent on the previous one. This architecture is very similar to blockchain technology itself, where we have connected blocks and each block is dependent on the previous one. In the case of electronic voting, the blocks are votes in the system and in case of successful voting a new block is added which depends on the previous one. Each block has its own randomly generated hash and hash to the previous block. If something changes in the current block, both the previous one and the one before it must be changed and so on until the first block. The hash value of the block can be equated with the value of the register to draw a parallel for the presence of a correlation attack. A correlation attack occurs when there is a correlation between the output of one register and the output of the Boolean function that combines the output of all registers. If the blocks were combined with a Boolean function, the same behavior could occur as with registers and there could be a correlation attack. It follows that the use of multiple operators (example combination of AND, OR and XOR) provides greater security than the use of fewer operators with multiple iterations (example: 2 AND operations and 1 XOR operation). Also the presence of negation can increase the possibility of correlational attacks.

B. *Algebraic attacks*

Algebraic attacks result from finding the key by solving a system of nonlinear equations. Since all signing and verification is based on mathematical equations, the possibility of an algebraic attack is definitely present. Mathematical operations are used in the signature and verification process: degree, degree per module, multiplication by modulus, addition of points on a curve, multiplication of a scalar by a point, inverse calculation, finding prime numbers, and so on. The advantage is the use of very large numbers when choosing a secret key and all the values that are selected or obtained by some calculation are very large. Therefore, solving such a system of nonlinear equations requires extremely large amounts of performance, which to some extent protects the key. In this case, finding the private key would mean abusing the user.

C. *Birthday attacks*

In the process of signing the vote, a hash function is used for the vote in order to increase the anonymity of the voters. If a duplicate of that hash function is found with the birthday attack, the voter loses anonymity, the hash function is no longer unique and one-way. The current implementation of the system uses interactive zero-knowledge proof to prove that the encrypted text is correct, but possible improvements indicate the use of non-interactive zero-knowledge proof. In this case we do not have proof between two people, but on the other side there is a hash function. In this case there is also a danger of a birthday attack. If the hash function is detected, zero-knowledge proof will not be correct and voting may be allowed when it should not be. We can increase the security even more by using different hash functions in both cases in order to protect ourselves from the attacker to a greater extent.

IV. Conclusion

Establishing a secure system is the key to communication. Cryptanalysis of possible attacks is a key factor in informing and likelihood of attacks occurring. All attacks use different mathematical methods and each of them can occur in different circumstances. No matter what the attack is, there is one general rule that means that participants in communication must use secure and credible resources in order to protect themselves from the outside world as much as possible. In case the same cannot be done, it is necessary to respond quickly to the existing problem.

Acknowledgment

I want to thank the Faculty of Computer Science and Engineering in Skopje for giving me the opportunity to take part in this conference.

References

[1] Schaefer, Edward. "An introduction to cryptography and Cryptanalysis." (2009).

[2] Yaga, Dylan, et al. "Blockchain technology overview." arXiv preprint arXiv:1906.11078 (2019).

[3] https://99bitcoins.com/what-is-blockchain. Accessed: 23.05.2022

[4] Meier, Willi, and Othmar Staffelbach. "Fast correlation attacks on certain stream ciphers." Journal of cryptology 1.3 (1989): 159-176.

[5] Armknecht, Frederik. (2004). Algebraic attacks on stream ciphers. European Congress on Computational Methods in Applied Sciences and Engineering ECCOMAS. 24-28.

[6] Wagner, David. "A generalized birthday problem." Annual International Cryptology Conference. Springer, Berlin, Heidelberg, 2002.M. Young, *The Technical Writer's Handbook*. Mill Valley, CA: University Science, 1989

[7] Kirchner, Paul. "Improved Generalized Birthday Attack." IACR Cryptol. ePrint Arch. 2011 (2011): 377.

[8] Gupta, Ganesh. "What is Birthday attack??." (2015)..