EasyChair Preprint

№ 9764

A Review on the Strength of Data Security in Blockchain Through Cryptographic Techniques

Naman Grover, Chandan Kumar Sangewar, Anchita Singh and S Sandosh

February 22, 2023

A Review on the Strength of Data Security in Blockchain Through Cryptographic Techniques

Naman Grover
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
naman.grover2020@vitstudent.ac.in

Chandan Kumar Sangewar
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
chandankumar.sangewar2020@vitstudent.ac.in

Anchita Singh
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
anchita.singh2020@vitstudent.ac.in

Dr. S. Sandosh
Assistant Professor(Sr.)
School of Computer Science and Engineering
Vellore Institute of Technology, Chennai, India
sandosh.s@vit.ac.in

Abstract
Changes in technology occur every day. Technology is advanced through data security. For this, data authentication at the time of execution is required. Cryptography ensures the independency and security of data. Before sending information, it encrypts it. There are several different cryptographic algorithms in use today. Block chains are distributed, decentralised, and immutable digital ledgers that can identify and stop tampering (i.e., a bank, company, or government). Blockchains are distributed databases that enable users to record transactions in a public distributed ledger that cannot be altered after publication for added security. This article provides a technical overview of blockchain technology and cryptography. To shed light on how the combination of blockchain technology and cryptography offers unbreakable security and safeguards against data loss.
Keywords— Blockchain, Attacks, Blowfish Algorithm,

Introduction

Cryptography is a mechanism for protecting private communications. When literally translated, the Greek phrase means "secret writing." However, thanks to today's encryption, only the receiver can read the sent data, keeping personal and company secrets safe. Because of its long and illustrious history, cryptography may be seen as a tried-and-true method that is always being refined. The Egyptians used "secret" hieroglyphics approximately 2000 B.C.E., but there is also evidence of similar practises in the form of ancient Greek hidden inscriptions and the famous Roman Caesar cypher. Without realising it, billions of people across the globe use cryptography every day to keep their financial and personal data private. While cryptography has numerous useful uses, it is also considered very fragile because of the risk of a whole system compromise from even a single implementation error.

Blockchain technology's use has increased exponentially in the ICT sector during the last several years. Most people's awareness of blockchain technology and its quick development may be attributed to the precipitous surge in cryptocurrency values and the large infusion of venture funding into blockchain firms. It is predicted that blockchain's rapid expansion would continue until at least the year 2021. After Bitcoin's first release and subsequent years, over 1500 other crypto currencies have emerged. Bitcoin

pioneered the concept of a decentralised cryptocurrency. It ensures that transactions are executed decentralised, without relying on a trusted third party.

With blockchain, a middleman (or shill) isn't needed to complete a financial transaction. As a replacement, it employs validators (often miners) to independently validate all monetary transactions. This can only be achieved by the formation of a dispersed consensus among several, mutually suspicious parties. In the world of cryptocurrencies, this computational problem is analogous to the double-spending problem, which refers to the challenge of ensuring that a specific amount of a digital coin has not already been spent without the help of a trusted third party (typically a bank) that maintains a record of all transactions and user balances.

Symmetric Cipher Model

Before going on to a general review of the many substitution methods, it is important to have a firm grasp of the symmetric cypher concept. It is correct that the same key is used for both encryption and decryption in this cryptosystem. The term "cryptography" was used to describe the practise of exchanging secret information between two parties.

- Equivalent to symmetric cryptography[9]. As part of this symmetric scheme, we use the notation "plaintext" (X), "ciphertext" (Y), "key," "encryption," and "decryption," respectively) (D).
- The definitions of plaintext, ciphertext, key, encryption, and decryption used here are the same as those used in the paper's introduction. In order to use traditional encryption in a risk-free manner, two prerequisites must be met:
- The first order of business is to identify a secure means of encryption.
- The sender and the receiver are both responsible for safeguarding a private key that was obtained by them.

Attacks on Symmetric Cipher Model

In most cases, the ciphertext itself is not the objective of an attack on an encryption system; rather, the aim of such an assault is the key that may be used to decode the information. Following an examination of the two possible methods for cracking the Symmetric Cipher model, we will next make an effort to determine which of these strategies is most likely to be successful. The two alternatives left at this point are to either try to guess the password or use brute force to enter the system.

Cryptanalysis

The first kind of attack is known as a cryptanalytic assault, and it involves not just technical know-how but also maybe some knowledge of the characteristics of the plaintext, or even simply a limited number of example pairings of plaintext and ciphertext. There are a lot of different entry points that may be used for this kind of attack. An attack like this uses the algorithm's properties to understand how it deciphers plaintext or detects the active key.

Brute force attack

In the second type of assault called a "brute force attack," the attacker systematically tries every conceivable key on the ciphertext until a decipherable "plaintext" version is found. More than half of the potential key combinations have to be tried before success is guaranteed, on average.

If the key is deduced, all messages encrypted using that key, whether current or historical, are vulnerable to attack.

Classical encryption techniques

There are two primary methods of encryption:

- Substitution Ciphers
- Transposition Methods

Methods of Substitution and Transposition Cryptography: A Comparison

To safeguard the information, use a cypher or factor such the Caesar Cipher, Monoalphabetic Cipher, Polyalphabetic Cipher, Playfair Cipher, Hill Cipher, RailFence Cipher, or Columnar

Transportation Cipher. Other alternatives include the Playfair Cipher, Hill Cipher, RailFence Cipher, and Columnar Transportation Cipher. All of them were produced by individuals from various countries. Julius Caesar, who lived in the nineteenth century, is credited with the invention of the Caeser Cipher. This cypher makes use of a replacement key and has a fast encryption speed, but its decoding performance is considerably worse when compared to that of competing methods. This encryption makes less of a demand on memory and has a little less avalanche effect. Unfortunately, even this particular encryption key may be broken using a Brute Force Attack; as a result, this cypher is not foolproof.

The letter replacement technique that this system employs is where the inventiveness of this method lies. The Caesar Cipher and the Monoalphabetic Encryption Algorithm are both extremely close to one another. The Caesar Cipher makes use of the replacement key type. This particular form of encryption has a minor avalanche effect, a sluggish encryption rate, and hardly no memory use at all. Because this encryption method only uses a single letter replacement with a fixed substitution, it is vulnerable to an attack using a text-only cypher because it only uses a single letter substitution. Using a substitution key, Leone Battista Alberti developed the Polyalphabetic cypher in the year 1467. A slow pace is maintained throughout both the encryption and decryption processes. This design uses a very small amount of memory and has a moderate influence on avalanches. This design flaw, which has been around for a long time, might be easily exploited by cypher text and plaintext attacks, the knowledge of which is already widespread. This pattern is one of a kind due to the fact that it can substitute letters and also arrange them in a matrix of 2626 cells. In terms of deciphering, the Playfair cypher and the Hill cypher are not too unlike from one another. Each was discovered on its own by Charles Wheatstone in 1854 and Lester S. Hill in 1929, respectively. Both of these options are for replacement key types. In this new development, playfair cypher and hill cypher are differentiated from one another by the fact that playfair cypher delivers encryption at a breakneck speed, but hill cypher takes an eternity.

Comparable decryption rates are observed. In terms of how much memory is being used, the playfair cypher is located in the middle, while the hill cypher is located at the very top. A considerable avalanche effect is produced by both the playfair and the hill cyphers. Both of them are sensitive to the effects of a physical force. The Hill cypher is one of a kind due to the fact that it can encrypt plaintext in addition to using a substitution matrix that is based on linear algebra to change the order of letter pairs. The railFence cypher and columnar transposition are two examples of additional factors and cyphers that are currently in use. It is difficult to differentiate them from the ones that came before them. The permutation key type provides encryption and decryption speeds that are equivalent to those of the replacement key type.

Columnar transposition uses a manageable amount of RAM, while RailFence makes extensive use of the memory. The avalanche effect and columnar transposition caused by RailFence are almost nonexistent. Both the plaintext and the ciphertext may be attacked in the same way if they are both taken from the same source. They share a lot of characteristics, yet each one is unique in its own way. Both the RailFence encryption and the columnar transposition are unique in their own right due to the distinctive ways in which the plaintext is written and read. This is because the RailFence encryption and columnar transposition use unique techniques.

Blowfish Algorithm

Bruce Schneier created the Blowfish cipher in 1993 as an alternative to the DES Encryption Technique. It is considerably quicker than DES and gives a decent encryption rate, with no known viable cryptanalysis techniques. It is one of the earliest patent-free, secure block

ciphers, and is therefore publicly accessible for use by anybody.
BlockSize: 64-bits
KeySize: configurable size between 32 bits and 448 bits
18 sub-keys are included in [P-array]
Number of Rounds: SIXTEEN
Four Substitution Boxes [each containing 512 elements of 32 bits]

Algorithm

Step 1: Initialization of subkeys:

- The processes of encryption and decryption each need a total of 18 sub-keys, labelled P[0] through P[17]; nevertheless, these sub-keys are shared across the two methods.

- Each member of the array is an entry consisting of 32 bits, and the 18 sub-keys are saved in a P-array.

- It is initialised with pi(?) digits.

- Each sub-key's hexadecimal form is

P[0] is "353f6a58,"
.
.
.
.
P[17] is "3789tby4"

Step 2: Populate Substitution Boxes

- Encryption and decryption both involve the use of four substitution boxes, often known as S-boxes. Each S-box has 256 items in it (S[i][0]...S[i][255], 0&lei&le4). Each individual item is comprised of 32 bits.

- It is initialised with the digits of pi(?) once the P-array has been initialised. Here you may locate the S-boxes!

Step 3: Function F

Blowfish is a Feistel network consisting of 16 rounds. Theinput is a 64-bit data element, x. Divide x into two 32-bit halves:
xL, xR
For i = 1to 16: xL = XL XOR Pi
xR = F(XL) XOR xR
Swap XL and xR
Swap XL and xR (Undo the last swap.) xR = xR XOR P17 xL = xL XOR P18 Recombine xL and xR
It was during the Cambridge Security Workshop in December 1993 that Schneier first presented the Blowfish Algorithm (BA) as a replacement for the Data Encryption Standard (DES). Because of its robustness and ease of hardware implementation, it provided secure and reliable data encryption. The lack of patent protection means that no licence is necessary. The primary steps in the BA algorithm are a table search, an addition, and an XOR. Four S-boxes and a P-array make up the table.
The Blowfish Algorithm is a cypher that employs a reduced variation of DES's F-function architecture (Feistel rounds) to provide the same level of security at a faster rate and with more software efficiency. Performance and security are impacted by factors like as key and block sizes, dispersion, and confusion, and algorithm complexity. Assumptions like this are used to hide the plaintext-ciphertext connection.
This makes it more difficult to launch cypher text assaults. It is more difficult for an attacker to find the plaintext when the statistics about the plaintext are dispersed across the ciphertext. Increasing the dispersion of a cypher increases the likelihood that a change to a single bit in the input will result in a change to every bit in the output (1 in 12). (half).
The symmetric block cypher known as BA is composed of a Feistel network, iterating a simple encryption function sixteen times, and BA. Each Feistel structure has its own set of advantages, particularly in terms of hardware. Changing the sequence of the keys in the key schedule is all that is required to read the cypher text. To some extent, the BA might

include techniques for encrypting and expanding keys.

Blockchain

Blockchain is "an open, distributed ledger that effectively records transactions between two parties in a form that can be confirmed and stored forever," according to the official definition. Blockchain is a distributed ledger that can record and verify data transfers and transactions. allows individuals to make decisions together while having different degrees of system control. Simply said, it's a technology that allows individuals to meet and collaborate with a broad spectrum of strangers to achieve their goals. Continuously, the blockchain will add what are known as "blocks," which are just bits of encoded data.
Text and headers are unique to each block that the block header includes the current block hash, Nonce, Merkle root, timestamp, and hash of the preceding block. Each block, which ranges from 1MB to 8MB, contains 500 transactions. The block should be made using mining data. Blockchain block genesis is the initial block. Until miners seek a consensus mechanism, the block size rises with each transaction. If more than half of the network's nodes agree on the algorithm, the transaction may be added to the blockchain by reusing the hash from the previous block.

Components of blockchain:

Hash function:

The blockchain uses a cryptographic hash function to hash transactions, and this function returns a hash of a certain size (32 bits, 64 bits, 128 bits, or 256 bits) as its output. Depending on the situation, the size of this hash might be 32 bits, 64 bits, 128 bits, or 256 bits. Further, as it employs an OWHF, de-hashing or reverse-hashing it is not conceivable.

Mining

The creation of new blocks for the blockchain is what mining is all about. The process of adding new blocks to the current chain involves a competition between miners, who create these blocks.

Consensus:

Without a central authority, this is the process of obtaining a consensus on whether or not to add the block to the chain. In addition, it delivers solid security, is fault-tolerant, and guarantees the proper block is added to the chain.

Listed below are the components of the consensus algorithm.

- The Proof of Work consensus process chooses miners arbitrarily, via trial and error, to produce blocks containing a nonce.
- In PoS, sometimes referred to as the Proof of Stack Consensus Algorithm, the miner selection is based on the miner's financial and chronological status within the network.
- The Delegated Proof of Stack Consensus Algorithm selects a miner via a mechanism of voting and election.
- A miner is chosen using a realistic Byzantine Fault Tolerance consensus mechanism when a minimum number of nodes concur.

Digital Signature

The crucial aspect of peer-to-peer (P2P) networks referred to as Authentication is lost when encryption is used; as a result, digital signatures were created to address this issue. Digital signatures provide many benefits, including protection against forgery, assurance of data integrity, and privacy protection for sensitive information. A (the sender) may encrypt a file using B's public key, and then B (the receiver) would use B's private key to decode the encrypted file. But how can B be sure that A was the one who really delivered this message (or file)? The use of digital signatures offers the opportunity for this challenge to be successfully circumvented.

Merkle Root

It is a framework that enables, according to F. Merkle Root. efficient and safe content verification in a vast quantity of data.

Smart Contracts

Intelligent contracts Blockchain offers a

quicker, cheaper, and more secure more secure decentralised platform and circumvents the intermediate

Types of Blockchain

Essentially, there are two blockchain models. Permission-based network criteria constitute Both permission-less and permissioned blockchain networks exist. Moreover, it may be divided into three categories:

Public Blockchain

Public Blockchain is a decentralised ledger that requires no special permissions to access. The public ledger of a blockchain network may be seen by anybody, and anyone can join the network. This way, the whole network may become decentralised and run on a permission-less basis, made possible by blockchain technology. a consensus mechanism that is immune to manipulation of data

Exclusive Blockchain Technology

In this configuration, the ledger is protected from prying eyes. Within a blockchain, only people who have been predetermined to be a member of the network at the time consensus is obtained are able to confirm transactions and block others. As its headquarters, it provides a more thorough authentication and verification of all transactions.

Blockchain Consortium

Blockchain Consortium Sometimes referred to as a Federated.

The blockchain serves as an intermediary between public and private blockchains in which blocks may be shared or kept secret. Once validation is complete, more links may be added to the chain. performed by a small number of predetermined users. It is in part decentralised.

There are six distinct categories of security services that may be provided by a blockchain: privacy, authentication, integrity, non-reputation, provenance, and confidentiality. A number of the most powerful attacks, such as a Distributed Denial of Service (DDoS), collision attack, sybil attack, eclipse attack, injection attack, replay attack, or ransomware attack, have the potential to undermine services such as integrity, authentication, privacy, and secrecy.

A. Hash-based Attack

In order to execute this attack, the attacker must first get the hash values and then look for further messages with the same hash value. By modifying the hash value, an attacker is able to seize control of more than 51 percent of the network's mining power.

B. Centralisation Attack

Blockchain is an example of a decentralised network since it functions as a peer-to-peer system. In this scenario, the attacker makes an effort to disrupt decentralisation and gives the impression of centralisation by acting in a certain way.

C. Traffic Level Attack

An adversarial node launching a traffic attack attempts to disrupt a network by generating unnecessary congestion, thereby blocking access to the network for the legitimate users.

D. Network Level Attack

An illegitimate user is one who compromises a network by using stolen or otherwise unlawfully obtained credentials, hardware, or software. Behaviour analysis may also be used to identify fake users.

E. Injection or Insider attack

Unauthorized users provide untrustworthy data to programme interpreters. A system expert with administrative privileges altered with the data, creating a harder task. This is a common insider attack, where a hostile user with administrative privileges modifies genuine user accounts and other information to hide their tracks and make it harder for security employees to identify the breach.

F. Private key leakage attack

When the same key and nonce are used several times, it gives an adversary the ability to either memorise the keys off by heart or utilise

duplicate values to figure out the nonces and secret keys.

Avalanche Effect

Any method for encrypting data should, ideally, have the property that even a little change in the plaintext or the key leads in a considerable shift in the cypher text. This is an essential need for strong encryption. Even if just one character of the plaintext or the key is changed, the ciphertext should have a large amount of sensitivity to these changes. This kind of behaviour is referred to as the "Avalanche Effect."

Proposal Design

The findings of the paper include using blockchain technology in conjunction with encryption in order to beef up security and guarantee the authenticity of all transactions. Using a method known as substitution cryptography, the information will be encrypted. The plaintext is changed into an encrypted version via the use of a method known as substitution cryptography. In order to construct the encrypted text, the required character representations are converted into their ASCII equivalents. The result is then sent through the blowfish algorithm for processing. The resultant number is subjected to a transposition in order to divide it in half, and each portion of the number is thereafter placed into its own block on the Blockchain.

Conclusion

After assessing the algorithms based on the Avalanche effect, we decided that the Hill Cipher scored the best; thus, it is suitable for use in situations when maintaining one's privacy and reputation is of utmost importance. When both time and memory are critical factors, the Caesar cypher, the Playfair cypher, and the Polyalphabetic cypher are the ones that function the most effectively. There has been a significant amount of study conducted on this topic, and it has gradually gained acceptance as a reliable and trustworthy technique for encrypting data. This is in part owing to the fact that it is flexible and effective when implemented in hardware, among other benefits.

Users are now able to securely store data and communicate with one another, while making all of their interactions public and auditable thanks to the technology that underpins blockchain. Blockchain technology has a lot of potential benefits, but it is still vulnerable to being attacked.

Therefore, it is crucial to have an in-depth knowledge of the many dangers it poses in order to avoid them. More would be done to protect the blockchain so that its users may continue to reap its benefits if more people knew about the many threats it faces.

References

[1] Dr. Sumathy Kingslin, R.Saranya, "Evaluative Study on Substitution and Transposition Ciphers" 2018 IJCRT | Volume 6, Issue 1 January 2018 | ISSN: 2320-2882

[2] Anita. N, Vijayalakshmi. M, "Blockchain Security Attack: A Brief Survey" IEEE - 45670 | 10th ICCCNT 2019

[3] Ashwak ALabaichi, Faudziah Ahmad and Ramlan Mahmod "Security Analysis of Blowfish algorithm" | ISBN: 978-1-4673-5256-7 / 13 / ©2013 IEEE

[4] Xi Li, Zehua Wang, VIctor Wand, Victor C M Leung, Hong JI, Yiming Liu, Heli Zhang," Blockchain-empowered Data-driven Networks: A Survey and Outlook" ACM Computing Surveys Volume 54 Issue 3 | Article No.: 58pp 1–38

[5] Arkan Kh Shakr Sabonchi, Bahrye Akay"A survey on the Metaheuristics for Cryptanalysis of Substitution and Transposition Ciphers" January 2021Computer Systems Science and Engineering 39(1):87-106

[6] K. A. Rajan, "Use of Transposition Cipher and its Types", IJRESM, vol. 4, no. 11, pp. 164–165, Dec. 2021.

[7] U. Thirupalu, Dr. E. Kesavulu Reddy FCSRC (USA)"A New Cryptosystem for

Ciphers using Transposition Techniques" SSN: 2278-0181 Vol. 8 Issue 04, April-2019 (IJERT)

[8] Mr. Rajendra S.Navale1, Mr. Adilshah N.Jalgeri, Mr. Balkrushna B.Jagadale "Survey on various substitution techniques for Cryptography" Volume-7,Issue-4, (Apr-17) ISSN (O) :- 2349-3585