iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** III **Month of publication:** March 2023

DOI: https://doi.org/10.22214/ijraset.2023.49860

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

Decentralized E-KYC Blockchain Network for Central Bank

Tanuj Gupta[1], Devanshi Wadkar[2], Tejas Shinde[3], Samarjeet Kunjeer[4], Anuradha Deokar[5]
Department Of Computer Engineering, Aissms Coe, Pune

Abstract: ***Know Your Customer or KYC processes are the backbones of a financial institution's anti-money laundering efforts. KYC processes are mandatory in most of Europe and India. Almost all countries have their own process for identifying and keeping a record of their citizens like the USA has Social Security Protocol. According to current estimates, the amount of KYC spending rose to up to $1.8 Billion in 2021 on a global level. Despite the importance of the process, KYC continues to operate inefficiently. KYC processes are labor-intensive and time-consuming tasks. It is estimated that 80% of KYC efforts go in gathering information and processing while only 20% of efforts are evaluating and monitoring focused. This centralization of data is causing an inefficient KYC process and creating issues like :- Misidentification of fraudulent data , Inability of tracking customers, customers entering fake data(address, age etc) , delayed processing time and very high cost of processing. Reimagining KYC Using Blockchain Technology will enable seamless and secure data exchange at a fraction of the cost. A decentralized blockchain network will Facilitate near real-time data exchange. Blockchain when used with other technologies can indicate great potential to help organizations reduce the cost and time linked with the KYC process. The introduction of blockchain in KYC brings data on a decentralized network which also has its own pros like distributed data collection ,higher operational efficiency , validation of accuracy of the information. The customer will build their profile on the DCT KYC network. There are multiple options to store the data like Centralized encrypted servers, DCT platforms like IPFS(Interplanetary File System). Customers will perform transactions with Financial Institutions by giving them access to users profile(restricted non-appendable access),Using hashing and cryptographic processes the data would safely be called and returned to the DCT network . Thus making sure the whole process is immutable .***

Keywords : Decentralized (DCT) , Hashing , Cryptography, Blockchain, DCT KYC network

I. INTRODUCTION

A. Centralised e-KYC

Know your customer (aka KYC) is the regulatory and compliance obligation on the conventional banking and financial system, to capture customer information before onboarding and providing any financial services to the customer. To say it in another way, banks must assure themselves that their clients are genuinely who they claim to be. Banks may terminate business relationships with a client if it refuses to meet binding KYC requirements. Banks need to oblige with KYC regulations to prevent money laundering and understand the nature of the customer's activities.

Traditionally, the KYC verification process has to be carried out by Individual banks Independently. For KYC, customers are typically required to be physically present at the bank's branch or on a video call to provide personal identification information, such as passport or ID cards.

Procedures for identity verification include documents, non-documentary styles, or a combination of both.This process is problematic for banks because it is highly cost – intensive, time consuming for the banks. The burden is suffered by the client, who must respond to each request for KYC information or threat detainments to their deals. This is especially true for global and multi-banked corporates who can admit large volumes of individual KYC requests from each of their different banks, putting strain on their business connections.

India has made a commendable progress in digitization of this process. However, lack of bank standards and bank reservations about sharing customer information with competitor banks limits the reusability of the data. Having a central utility collecting all the data seems like a good idea. However, recent reports of leaks and misuses of personal data have lowered the confidence of both banks and customers in solutions that involve central data collection. At the same time, if a client operates multiple accounts in multiple banks, the disagreement grows indeed wider.

B. Integrating Blockchain in e-KYC

Blockchain is an intricate peer-to-peer network that uses distributed databases to identify and record information, also known as the value internet. It is a system of recording information that makes it insolvable or delicate for the system to be changed, addressed, or manipulated. Blockchain uses distributed tally technology, which publishes data across millions of bumps connected to the chain. Hence, there's no need for a centralized data garçon or network. This brings down the costs drastically. So, banks can now be free of the need to store and cover data, which is automatically taken care of.

II. LITERATURE SURVEY 1

MedRec: Patient Control of Medical Record Distribution Andrew Lippman, MIT; Nchinda Nchinda, Candidate for MS in Computer Science at MIT; Kallirroi Retzepi, Graduate Student, Viral Communications group at MIT; and Agnes Cameron, Master's Student, Viral Communications group at MIT IEEE Blockchain Technical Briefs, July 2018 Introduction Increasingly, people in the United States are required to manage their own healthcare and associated information. The days of the lifetime family doctor are over. At the same time, healthcare providers have to make that data available. This situation opens the door for innovative approaches to patient management. One obvious solution is a "Swiss bank" for healthcare records. These deposits engaged patient management to a cross-provider intermediary. In return for that convenience, we risk the addition of new data silos and commercial control points.

As an alternative, we can use a non-commercial, distributed system that allows patients to control who can access their records and thereby create a network solution where providers join that network and make data available on-demand at the request of patients.

MedRec is a network rather than a service. The advantage of this is that we can provide a cross-provider, patient-oriented interface and interaction mechanism. We constructed it using an Ethereum blockchain and we have tested it with diverse databases provided by our research partner, the Massachusetts-based Beth Israel Deaconess Medical Center. Further development will be done by a new, non-profit research endeavor called the Health Technology Innovation Center operated at BIDMC with continued participation by a team at the MIT Media Lab1 . We note three features of the system that are potentially significant. First, the system is designed to accommodate access to data for clinical researchers and serve as a point of entry for socially valuable epidemiological research for example to understand the propagation of disease and epidemics. MedRec is about more than patients or doctors, it is a component of a general healthcare environment.

Second, the architecture of MedRec is general. There are few that are health specific. We envisage that it can be a model for the management of individual identity and permissions in many situations where end-user control of identity and personal information across applications is crucial. This can be a base for social networks and as a benefit for individuals who want to ease who knows what about them. There is no coinage or transaction inherent in MedRec; it is designed to be free and open.

Third, in keeping with the design ethos of the Viral Communications Research Group at the MIT Media Lab, the system can be adopted incrementally, organization-by-organization. It is beneficial for internal management of records by hospital networks that consist of many independent providers and it scales to multiple, large-scale healthcare organizations. MedRec was inspired by original work by Ariel Ekblaw and Asaf Azaria2 . The current version, which is a new architecture, is supported by a grant from the Robert Wood Johnson Foundation. We use a blockchain that is maintained by medical providers who originate records to archive "smart contracts" that define access rights. Other data is also stored on chains. The goal of the program is to create a disinterested, non-profit, university-based system for patient control.

A. Design

The architecture of MedRec is easily understood by analogy to the World Wide Web. The web consists of three elements: An HTTP server that provides access to local data, the HTML protocol by which access is obtained and web elements are defined, and a browser that forms the interface.

Ideally, anyone and everyone could be a server and web browsers can draw from multiple ones to create a presentation. The World Wide Web is by design a network rather than a client server architecture even though in practice there are dominant servers. In MedRec, the language is a set of contracts commenced by patients that define what entities or parties can access which records . There are at present three types of contracts and more can be created. The simplest is one that asserts that entity B can access the records of patient A. More complicated ones allow for intermediary healthcare proxies, or allow a pharmacy to access all prescription records for patient A from any healthcare provider. We call the server equivalent a "full node." Full nodes are administrative members of the network. They can append blocks to the chain, admit new administrative members, and distribute notifications submitted to or originated by them. Examples include requests for participation in a clinical or epidemiological study or record changes.

We use proof of authority to append blocks and the addresses of holders of that authority are also stored on chain4 . New members with those rights are voted in by a majority of existing members. This facility is part of the Ethereum Blockchain5 . The interface is a local app run on a PC or phone . It allows generation of contracts and polls providers for notifications. There is an interface for a provider and one for a patient. Patient interfaces are light nodes and may or may not contain a copy of the blockchain. Third parties can also run an equivalent light node. That may include research organizations, pharmacists, patients' relatives, etc.

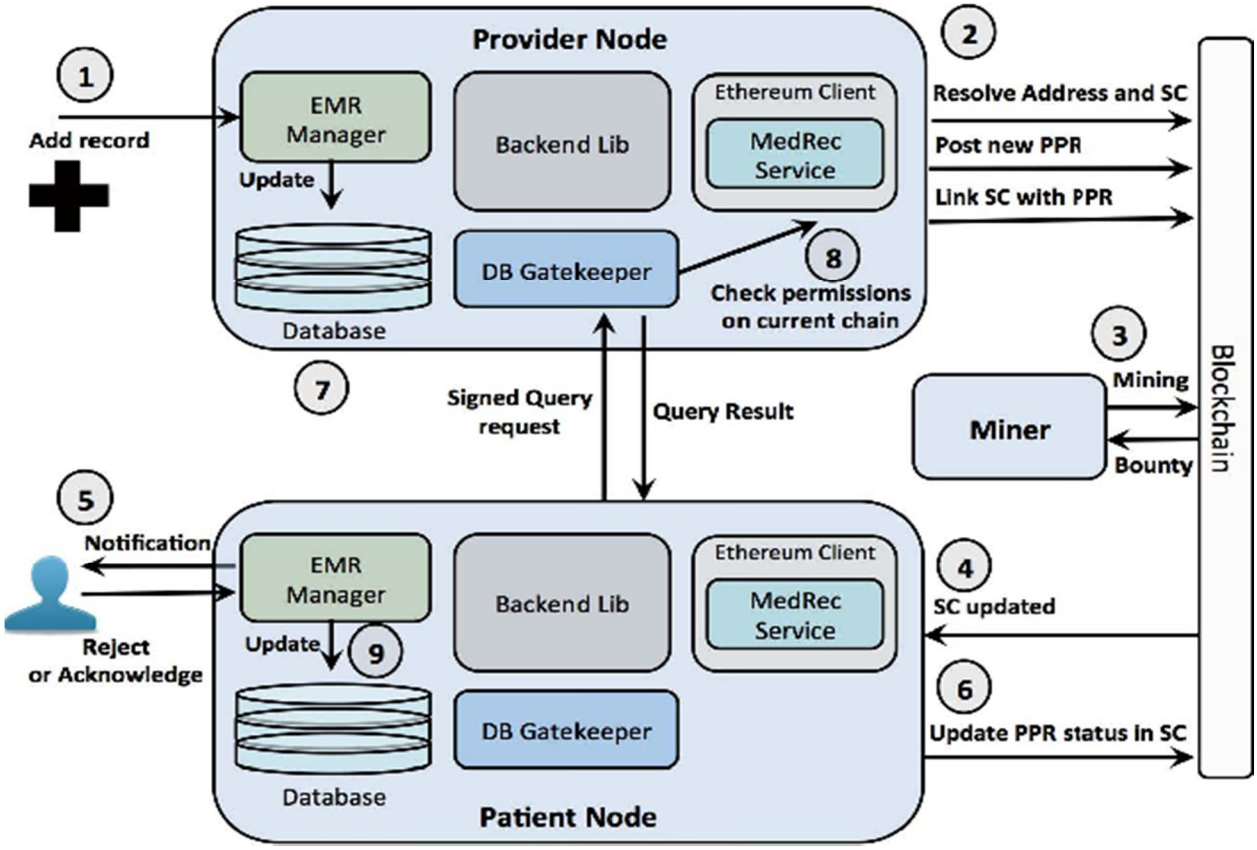


FIGURE 1.1

B. Operation

In this section, we show the work flow for three potential network constituents: healthcare providers, patients, and third parties such as pharmacies and research organizations.

Every user in the MedRec network installs the software and creates a login account. New providers make proposals to a special smart contract that orchestrates the addition and removal of providers to the network. Existing providers vote on whether to accept these proposals. Patients form relationships by sharing their account ID (an Ethereum address) with medical providers. Once a relationship with a provider is formed, patients can enable other accounts with the power to view portions of the medical data stored by that provider.

C. Caveats, Assessment, and Future Work

There are several elements to adoption of a MedRec network that are subjects of further work and development. Most crucial is the means by which providers adopt an interface to the system. A provider, as operator of a full node, commits to run a program that grants access to their databases under the rules of MedRec contracts. This entails an interfacing investment that can be noteworthy .

For large providers who already use an existing patient management application, this needs to be done once for that system and others can then use it. For smaller providers such as group practices, one must build an interface for each system that is in use.

As with any blockchain implementation, important questions include who maintains the blockchain, what the trust model is, what threats are to be defended, and what consensus scheme is to be used. In this case, the network is semi-public.

Anyone can join as a light node and be the predicate in a contract. But only providers can authorize contracts and append to the blockchain. Providers are trusted entities but we inoculate the system against intrusions of their internal systems by requiring majority voting. We argue that Ethereum-supported proof of authority mechanism is a robust solution. The overhead of running a full node is small both in terms of management and allocation of resources. Conversely, the advantages are large. The open-source model allows us to evolve with needs and community desires. These issues are assertions that will be tested at scale in real use. A second issue is the nature of the patient interface. We suspect that individual management of personal data is a task akin to management of a retirement plan. They are similar in that when we are young and healthy, we likely dedicate little energy to either retirement or healthcare. It has been extensively demonstrated that people devalue long term or low probability events. A good interface may ameliorate this.

To date, the interface we have implemented is optimized to be simple and encouraging. It allows for contract creation and deployment, visualization of the user's network and the ability to fetch and view data from the remote database. As we add features that are common in commercial healthcare interfaces, we have to keep an eye that the system does not become a task to use. This will evolve in time.

D. Conclusion

Proposed is a blockchain-based system that serves a societal need without the imposition of visible transactions or an application-specific coinage. We fill-in a network for a service and use the blockchain to manage that service. There is no definite economics associated with the work, nor any view of how society is organized. The general nature of the solution is biddable to other cases where an open-source, distributed model is useful.

III. LITERATURE REVIEW 2

Blockchains reworking the Media Experience: the instance of Bloomen .

The increasing use of digital technologies for media content production, distribution and delivery has enabled the planning of latest solutions. Blockchain technology will dispense a completely unique framework permitting privacy preservation,copyright protection and new models for direct compensation of certified content creators. associate degree example of such associate degree approach for disrupting the media trade with blockchain technology is that the Bloomen project

(European Union's Horizon 2020 analysis and innovation program on the base of grant agreement No. 762091 Bloomen: Blockchains within the new era of democratic media expertise. website: http://bloomen.io/).

The main goal of the Bloomen project is to increase the utilization of the blockchain technology to handle totally different on-line user transactions, providing associate degree innovative method of content creation, sharing, customized consumption, substantiation and copyrighting. an important a part of this approach is that the KYC method, so as to link

the cryptanalytic identities to real-world identities and to own access to a permissioned blockchain network. Identity management problems handled inside Bloomen contains each biometric authentication across totally different networks (outside and within the blockchain) and identity management over closed silos systems and open public ledgers.

Given the localized nature of the design projected by the project, it addresses a way to distribute and deconcentrate the identity management practicality, as well as the choices for users and alternative third parties to line this practicality themselves, as opposition betting on central systems directors. Also , it's into the thanks to distribute the identity management practicality across the various levels of the Bloomen design (e.g., inside the blockchain and therefore the application level) and investigates however, with the quality iatrogenic by the decentralization, identity management may be created to be universal, i.e., however it will extend across terribly dissimilar use cases. The surroundings of the analysis work bestowed during this paper has been enforced within the frame of the Bloomen project as how to implement a light-weight, economical and localized KYC method on a gathering blockchain so as to facilitate the availability of media applications .

IV. SECURITY ALGORITHM

A. Cryptography

Distributed computing, medium design, and cryptography algorithms form the holy trio of blockchain technology. Distributed computing utilizes a decentralized network of computers and was before blockchains in the form of torrenting networks.

Cryptography is what serves as security for guarding those impulses. The seminal Bitcoin white paper explained how these three scientific principles could play together to form a secure, peer- to- peer exchange of value that would exclude the need for a third party in fiscal deals.

A hash function is a fine function that converts a numerical input value into another compressed numerical value.
The input to the hash function is of an arbitrary length but affair is always of a fixed length. For eg MD, SHA1, SHA- 2, SHA- 3, RIPEMD, and Whirlpool.

B. Hashing in blockchain

Blockchains hash each sale before speeding them together into blocks. Hash pointers link each block to its precursor, by holding a hash of the data in the former block. Because each block links to its precursor, data in the blockchain is inflexible. The mincing function means that a change in any sale will produce an entirely different hash, which will alter the hashes of all posterior blocks. To propagate a change across the blockchain, 51 of the network would have to agree to it. Hence, the term " 51 attack ".
Different blockchains use different cryptography algorithms. The Bitcoin blockchain uses the SHA256 algorithm, which produces a 32- byte hash. Dogecoin and Litecoin both use Scrypt, which is one of the briskly and lighter cryptography algorithms.
Important Characteristics For A Strong Hashing Algorithm:-
A cryptographic hashing algorithm must fulfill these specific criteria to be effective:
The same input must always induce the same yield. Anyhow of how numerous times you put the data through the mincing algorithm, it must constantly produce the same hash with identical characters in the string.
The input can not be derived or calculated using the affair. There should be no way to reverse the mincing process to see the original data set.
Any change in the input must produce an entirely different affair. Indeed changing the case of one character in a data set should produce a hash that's significantly different.
The hash should be of a fixed number of characters, anyhow of the size or type of data used as an input.
Creating the hash should be a fast process that does n't make heavy use of calculating power.

C. Working of SHA 256

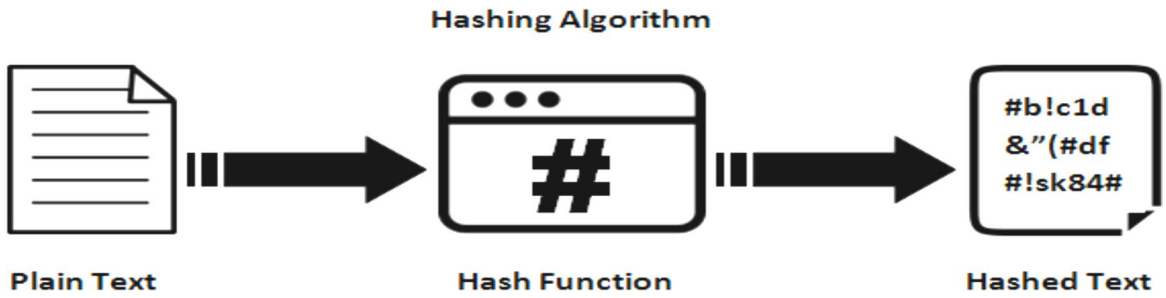


Figure 1.2

Append : Padding bits
First step of our hashing function begins with appending bits to our original message, so that its length becomes the same as the standard length required for the hash function. We proceed by adding a few bits to the message that we have in hand. The length of the message should be exactly 64 bits less than a multiple of 512 after addition of the calculated bits.
M + P + 64 = n x 512
i.e M = length of original message
P = padded bits

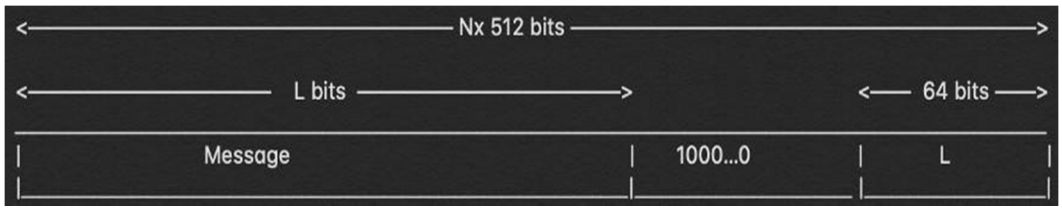


Figure 1.3

The bits that are appended to the message, should begin with '1' and the following bits must be '0' till they are exactly 64 bits less than the multiple of 512.

Append : Length bits

we now append our length of bits which is equivalent to 64 bits, to the overall message to make the entire thing an exact multiple of 512.We add the remaining 64 bits by taking the modulo of the message given.

i.e. the one which is without the padding, with 2^{32}.

The message obtained is then appended to the padded bits and we get the entire message block, which must be a multiple of 512.

Initialize the buffers

A message block is now created on which computations are carried out to figure out the final hash. we need certain default values to be initialized for this step.

a = 0x6a09e667
b = 0xbb67ae85
c = 0x3c6ef372
d = 0xa54ff53
e = 0x510e527f
f = 0x9b05688c
g = 0x1f83d9ab
h = 0x5be0cd19

There are more 64 values that need to be included which will act as keys and are denoted by the word 'k'.

Courtesy - SHA-2 Wikipedia

Now ,these values are utilized to compute the hash.

```
k[0..63] :=
   0x428a2f98, 0x71374491, 0xb5c0fbcf, 0xe9b5dba5, 0x3956c25b, 0x59f111f1, 0x923f82a4, 0xab1c5ed5,
   0xd807aa98, 0x12835b01, 0x243185be, 0x550c7dc3, 0x72be5d74, 0x80deb1fe, 0x9bdc06a7, 0xc19bf174,
   0xe49b69c1, 0xefbe4786, 0x0fc19dc6, 0x240ca1cc, 0x2de92c6f, 0x4a7484aa, 0x5cb0a9dc, 0x76f988da,
   0x983e5152, 0xa831c66d, 0xb00327c8, 0xbf597fc7, 0xc6e00bf3, 0xd5a79147, 0x06ca6351, 0x14292967,
   0x27b70a85, 0x2e1b2138, 0x4d2c6dfc, 0x53380d13, 0x650a7354, 0x766a0abb, 0x81c2c92e, 0x92722c85,
   0xa2bfe8a1, 0xa81a664b, 0xc24b8b70, 0xc76c51a3, 0xd192e819, 0xd6990624, 0xf40e3585, 0x106aa070,
   0x19a4c116, 0x1e376c08, 0x2748774c, 0x34b0bcb5, 0x391c0cb3, 0x4ed8aa4a, 0x5b9cca4f, 0x682e6ff3,
   0x748f82ee, 0x78a5636f, 0x84c87814, 0x8cc70208, 0x90befffa, 0xa4506ceb, 0xbef9a3f7, 0xc67178f2
```

D. Compression Function

This step is the main part of the hashing algorithm . The entire message block of 'n x 512' bits long is divided into 'n' chunks of 512 bits and each of these 512 bits are then put through 64 rounds of operations and the output obtained is fed as input for the next round of operation repeatedly.

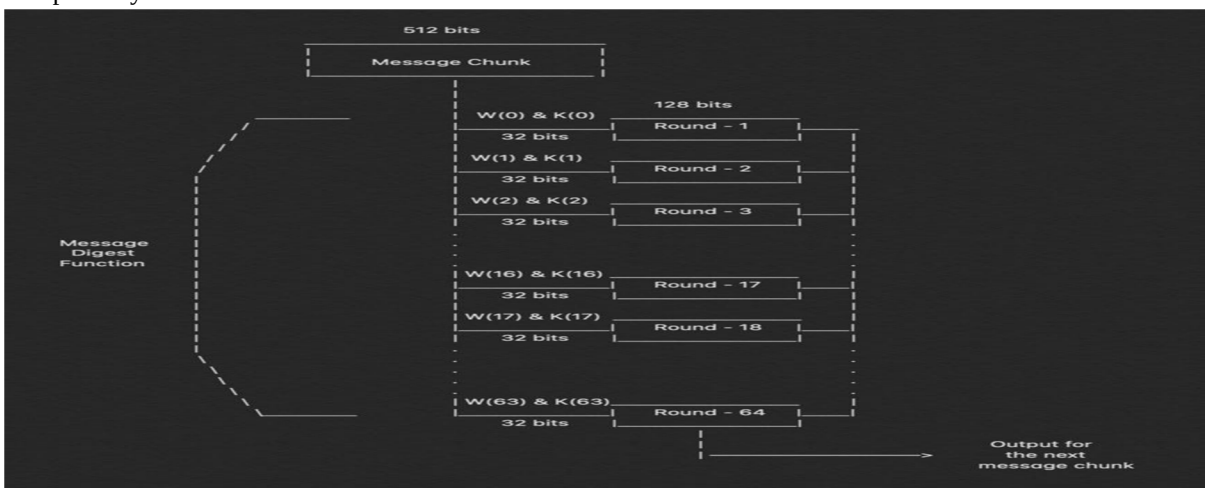


Figure 1.4

ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue III Mar 2023- Available at www.ijraset.com

In the image above it is clearly seen that the 64 rounds of operations are performed on a 512 bit message. It is observed that two inputs that are sent in are W(i) & K(i), for the first 16 rounds we further break down 512 bit messages into 16 parts each of 32 bit but after that we need to calculate the value for W(i) at each step.

$W(i) = W^{i-16} + \sigma^0 + W^{i-7} + \sigma^1$

where,

σ^0 = $(W^{i-15}$ ROTR$^7(x))$ XOR $(W^{i-15}$ ROTR$^{18}(x))$ XOR $(W^{i-15}$ SHR$^3(x))$

σ^1 = $(W^{i-2}$ ROTR$^{17}(x))$ XOR $(W^{i-2}$ ROTR$^{19}(x))$ XOR $(W^{i-2}$ SHR$^{10}(x))$

ROTR^n(x) = Circular right rotation of 'x' by 'n' bits

SHR^n(x) = Circular right shift of 'x' by 'n' bits

Here , we establish a method to create the W(i) for any given of the 64 rounds.

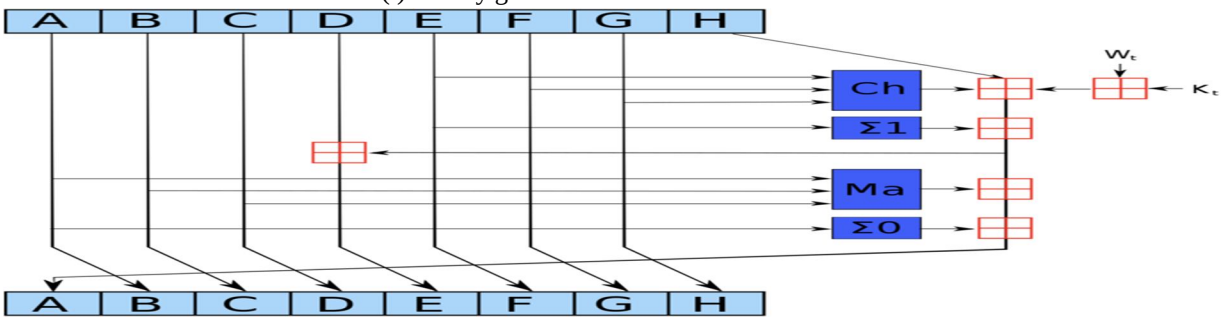


Figure 1.5

Depiction of a single “round”

The above image exactly shows what happens in each round and now that we have the values and formulas for each of the functions carried out we can perform the entire hashing process.

Ch(E, F, G) = (E AND F) XOR ((NOT E) AND G)

Ma(A, B, C) = (A AND B) XOR (A AND C) XOR (B AND C)

∑(A) = (A >>> 2) XOR (A >>> 13) XOR (A >>> 22)

∑(E) = (E >>> 6) XOR (E >>> 11) XOR (E >>> 25)

\+ = addition modulo 2^{32}

These are the functions that are performed in each of the 64 rounds that are performed repetatively for ‘n’ number of times

Output

The output from each round acts as an input for the next round and this process keeps on going till the last bits of the message remains and the result of the last round for the n^{th} part of the message block will provide the result i.e. the hash for the entire message. The length of this output is 256 bits.

Conclusion

The SHA-256 hashing algorithm is presently one of the most popularly used hashing technique as it hasn’t been cracked yet and the hashes are calculated quickly in comparison to the other secure hashes like the SHA-512. It is very well established but the industry is trying to progressively move towards the SHA-512 Which is more secure as experts claim SHA-256 might be vulnerable very soon.

V. PROOF OF WORK CONSENSUS ALGORITHM

A. What is a Consensus Algorithm?

A consensus algorithm is a process in computer science used to achieve agreement on a single point value among distributed systems. These algorithms are useful to achieve reliability in a network involving multiple users or nodes. Solving the consensus problem issue is important in distributed computing and multi-agent systems such as those seen in cryptocurrency blockchain networks.

Consensus algorithms play an important role in large-scale, fault-tolerant systems because they enable a set of distributed nodes to work as a collective group and agree on system state, even in the presence of failures or outages. The algorithm sets a threshold, or a specific number of member machines that must reach consensus.(51 % in POW consensus)

As consensus over a problem takes place, consensus algorithms take into consideration that some processes and systems will act maliciously and that only a portion of the nodes will respond. They also consider some communications will be lost during transmission of data over the network. However, the available nodes should be able to carry out the consensus process
,even if some of the nodes don't work . For example, an algorithm may require that at least 51% of nodes respond to achieve consensus or agreement on a data value or network state.
This ensures consensus is achieved with the available nodes or resources. The mechanism also ensures the integrity of decisions made by the remaining nodes in the fault-tolerant system.

B. Proof of Work

The PoW algorithm is one of the first algorithms created for consensus. First introduced in 1993, however it was reintroduced in 2008 by Satoshi Nakamoto founder and creator of bitcoin. In Proof of Work nodes solve complex mathematical puzzles as fast as possible. The miner who solves it in the least amount of time gets the reward and his block is added to the blockchain
Blockchains based on the PoW algorithm, miners who are also known as participant nodes must solve a complex mathematical problem by finding a cryptographic hash of a particular block in order to prove the work done by them is legit , only then the block is confirmed.
The miners do this by taking data from a block header as an input, and continuously running this data through a cryptographic hash function. Small changes are made to the input data to increase or decrease the difficulty of the mathematical puzzle by including an arbitrary number called a nonce. Every block has a unique nonce. Difficulty of the mathematical puzzle is decreased when there are less miner nodes active on the network, on the other hand the nonce value is increased when traffic on the network is high.
When the miner finds the solution that leads to consensus, they get a block reward, which generally is tokens of that native network. Doing all this work results in high energy consumption . Therefore many networks are trying to switch from POW to Proof of Stake, POW is very hazardous for the environment.
Nonetheless, the PoW algorithm continues to be one of the main consensus algorithms because

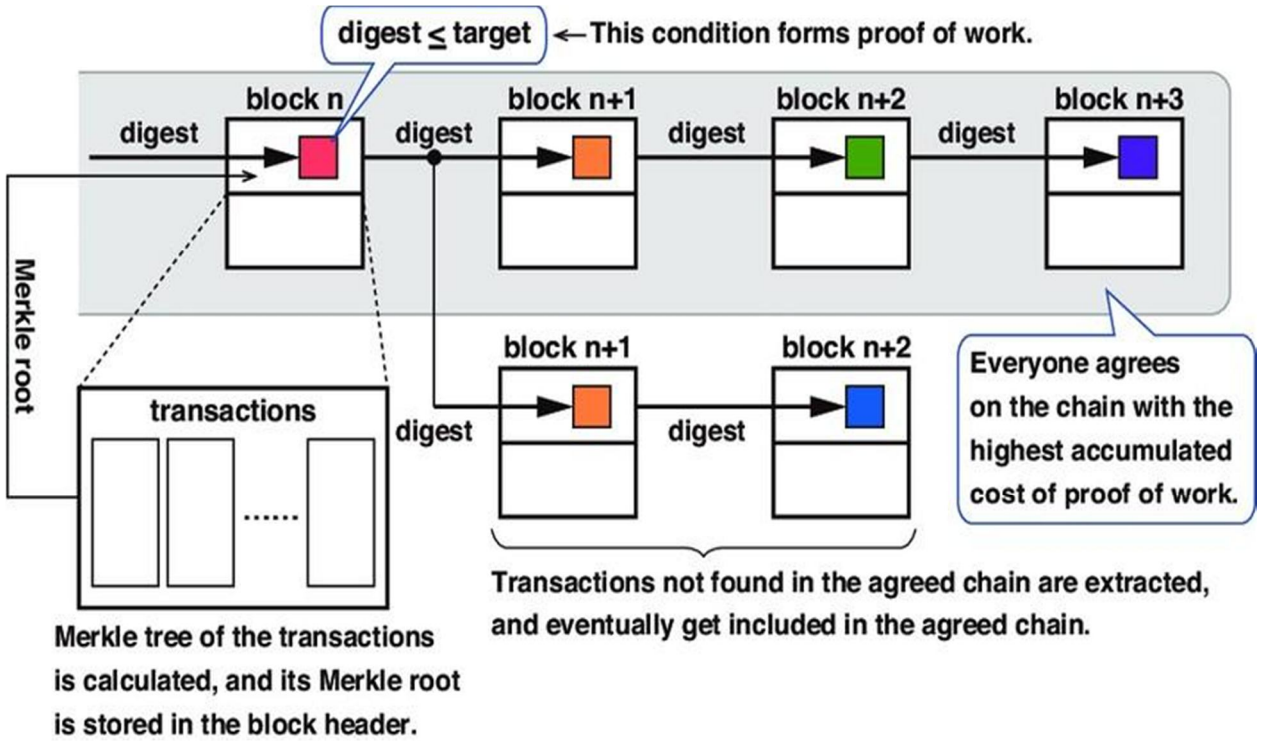


Figure 1.6

It maintains network security and is highly resistant to attacks from hackers like DDoS attacks, sybil attacks, replay attacks etc. It's also one of the first consensus algorithms and has proven to be an excellent choice for maintaining a high level blockchain.

C. Disadvantages of Proof of Work Consensus

POW is found to be inefficient with slow transaction speeds on the blockchain network

One of the major drawbacks of POW is that it consumes a large amount of energy. As POW works on a system where the fastest miner gets the reward, all miner nodes try to set up a high powered lab in order to be more efficient.

Miner nodes are required to buy very expensive equipment in order to stay in the competition.

Transaction fees can get very high on a busy day on a POW blockchain network.

Bitcoin is one of the first and largest cryptocurrency network in the world that runs on POW consensus. Bitcoin consumes more power than entire nations.

Nonce: Nonce is a random number that can be used just once in a cryptographic communication and is added to a hashed or encrypted block in a Blockchain that, when rehashed, meets the difficulty level restrictions

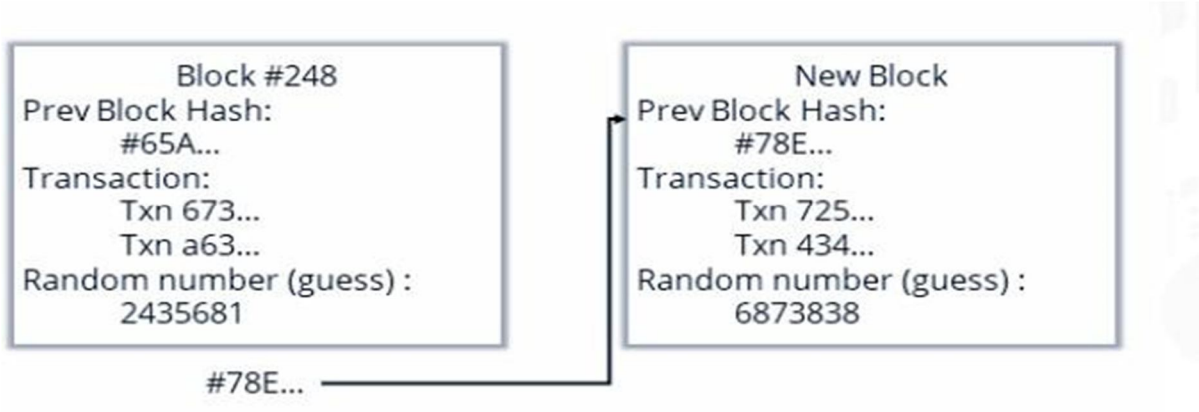


Figure 1.7

D. Merkle Trees

Merkle trees, which are also known as Binary hash trees, are an extensive sort of data structure in computer science. In bitcoin and many other cryptocurrencies, they're used to encrypt blockchain data a lot more efficiently and securely. It's a mathematical data structure that is made up of hashes of various data blocks that collectively summarize all the transactions in a block.

It also enables fast and secure content verification across big datasets and verifies the steadiness and content of the data.

E. Merkle Root

A Merkle root is a simple mathematical technique for confirming the facts on a Merkle tree. They're used in cryptocurrency to ensure that data blocks sent through a peer-to-peer network are intact ,damaged- free , and not altered. They play a very important role in the computation required to keep cryptocurrencies like bitcoin and ether running smoothly.

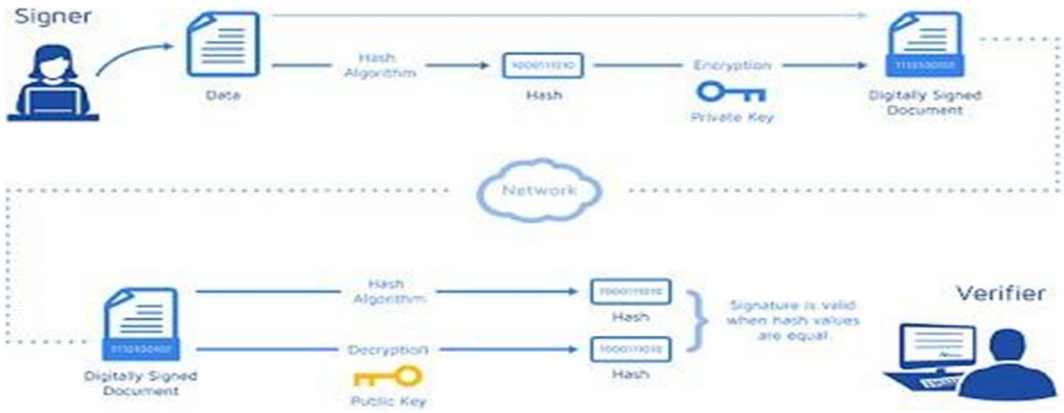


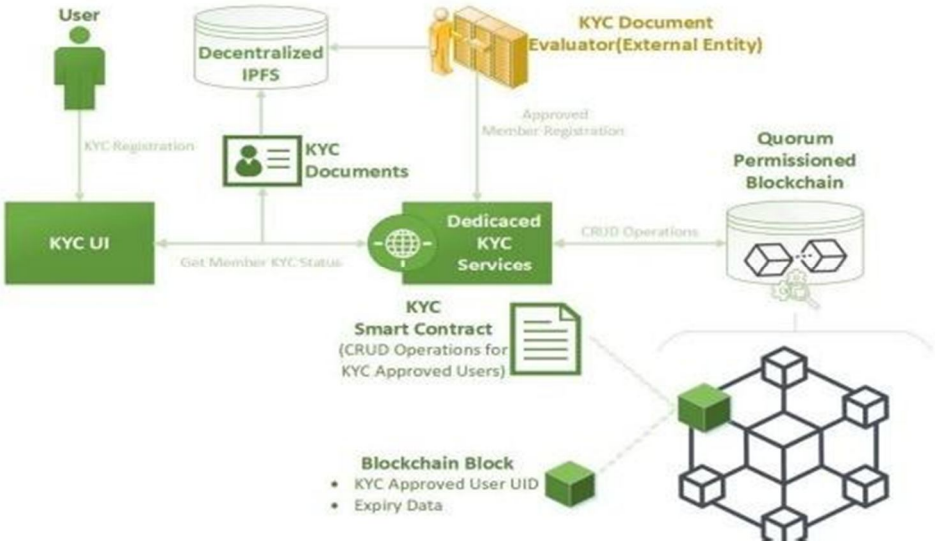
Figure 1.8

Working of Merkle Tree: A Merkle tree sums up all transactions in a block and generates a digital fingerprint of the entire set of operations, allowing the user to verify whether it includes a transaction in the block or not.

Merkle trees are made by hashing pairs of nodes repeatedly until one hash remains; this hash is known as the Merkle Root .
They're built using the bottom-up approach, using Transaction IDs, which are hashes of individual transactions.
Each non-leaf node is the hash of its previous hash, and every leaf node is a hash of transactional data present in the block
Merkle trees and blockchain work simultaneously .If a blockchain network didn't include Merkle Trees, per se, every node on the network would have to retain a complete copy of every single Bitcoin transaction ever made.Any authentication request on the network would require a huge amount of data to be transferred over the network.
Merkle Trees are a solution to this issue. They hash records in accounting, separating the proof of data from the data itself.
Proving that giving a tiny amount of data across the network is all that is required for a transaction to be valid.
Furthermore, it enables you to demonstrate that both ledger variations are alike in terms of nominal computer power and network bandwidth.
Validate the data's integrity: It can be used in validating the data's integrity effectively.
Takes little disk space: Compared to any other data structures, the Merkle tree takes up very little disk space.
Tiny information across networks: Merkle trees can be broken down into small pieces of data for verification purposes.

VI. SYSTEM ARCHITECTURE

The whole system's architectural approach focuses on simplicity of operations. Firstly, the user, a candidate system customer, submits their Know Your Customer personal documentation during the "KYC Registration" process through the dedicated and simple user interface called "KYC UI" . Obviously, it is the responsibility of the user to properly describe the requested information; it is necessary to notice that a new scanning mechanism will be sensitive for misbehaving users and will exclude them early in the process. After proper document submission, beyond the focus of this paper, an easy and simple-step procedure, the KYC Documentation is stored in a decentralized peer-to-peer blockchain-compatible repository, IPFS (InterPlanetary File System https://ipfs.io). IPFS protocol creates a resilient system of file storage on a decentralized peer-to-peer network. Though, IPFS has simple security over its data.
The core security tools behind IPFS technology consist of cryptographic hashing techniques and mechanisms. Every chunk of data that is stored on IPFS has a dedicated address that originates from a special data-hashing procedure. It is known that a hashing process consists of a one-way function that takes as input the data and outputs a single hash, while there is no reverse function.
Similarly, IPFS content addressing associates a single hash (content id or CID, e.g., "QmbFULfor2mTXvsoBPS9EkoC5Xzt56geutPpm49dqpjz8Y") that routes to the content data. To access the content data, its CID is needed. Following the successful storage and validation of the KYC data, a procedure is triggered in order to store specific information to the blockchain. The data that will be stored in the blockchain is only the essential information in order to identify the KYC-approved user, and also the time period in which the KYC approval is valid. Therefore, there is no sensitive information stored in the blockchain. In this way, sensitive information of the user cannot be appended but only verified.
It can also be accessed under special circumstances , with user consent only.



VII. RESULTS

Different research papers are studied to understand how different system architectures can provide better or worse efficiency for the verification process

Based on 0ur research done during the proposed work following interpretations are made :

The cost for KYC verification is drastically reduced due to the use of a decentralized platform

The processing time for the data is very low as compared to traditional KYC verification process

D – KYC provides a transparent and extremely high security platform which is also capable to protect data privacy for KYC verification, which is missing in the current system

D-KYC when used in combination with other technologies such as AI can showcase high potential

An industry's future lies in total digital transformation, which can only be accomplished through infrastructure changes. To improve operational efficiency, core processes must be modified. This can only be achieved by being welcoming to new and riotous technologies. The main goal of the solution proposed by us was to reimagine the existing traditional KYC process. This proposed paper gives a solution to the problem of redundancy and inefficiency in the current KYC process, lowering the system's operational costs drastically.

We also eliminate the presence of a single point of failure by utilizing a blockchain-based approach. Blockchain is a game-changing technology, and its applications are expanding exponentially. Implementing a blockchain application for kyc document verification provides proof of identity of the customers on banks and pellucid access to all or any of the banks in the blockchain network, ensuring quick access to the kyc document while also providing greater security. By doing so, we can lower the cost of maintaining the document from the centralized organization.

REFERENCES

[1] Vincent Schlatt, Johannes Sedlmeir , Simon Feulner, Nils Urbach, “Designing a Framework for Digital KYC Processes Built on Blockchain-Based Self-Sovereign Identity.” This is the accepted version of https://doi.org/10.1016/j.im.2021.103553, published in the Special Issue “Blockchain Innovations: Business Opportunities and Management Challenges” in Information & Management, 2022

[2] O. Novo, “Blockchain Meets IoT: An Architecture for Scalable Access Management in IoT,” IEEE Internet of Things Journal, vol. 5, no. 2, pp. 1184 - 1195, 2018.

[3] C. Lin, D. He, X. Huang, K.-K. R. Choo, A. V. Vasilakos, “Bsein: A blockchain-based secure mutual authentication with fine-grained access control system for industry 4.0,” Elsevier Journal of Network and Computer Applications, vol. 116, pp. 42-52, 2018.

[4] M. Ober, S. Katzenbeisser, K. Hamacher, “Structure and Anonymity of the Bitcoin Transaction Graph,” MDPI Future Internet, vol. 5, no. 2, pp. 237 - 250, 2013.

[5] European Central Bank (2012) Virtual currency schemes.https://www.ecb.europa.eu/pub/pdf/other/virtualcurrencyschemes201210en.pdf, 31 Oct 2017

[6] Sunitha N V, P Ashwinǀ, Sandhya, Shriraksha Bhat, Tushara Sasi, “ KYC Verification Using Blockchain” International Journal for Research in Applied Science & Engineering Technology (IJRASET) ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538 Volume 10 Issue VII July 2022

[7] Xingtong Chen and Gang Kou, “A systematic review of blockchain Min Xu” Correspondence: xumin@swufe. edu.cn Southwestern University of Finance and Economics, Chengdu, China .

[8] Moyano, J. P., & Ross, O. “KYC optimization using distributed ledger technology.” Business and Information Systems Engineering, 59(6), 411-423. https://doi.org/10.1007/s12599-017-0504-2 , 2017

[9] Syed Azhar Hussain and Zeeshan-ul-Hassan, “BLOCKCHAIN-BASED DECENTRALIZED KYC (KNOW-YOUR-CUSTOMER)”, The Fourteenth International Conference on Systems and Networks Communications ICSNC 2019.

[10] S. Nakamoto. “Bitcoin: A peer-to-peer electronic cash system”, 2008.

[11] Kosba, A. Miller, E. Shi, Z. Wen, and C. Papamanthou. Hawk: “The blockchain model of cryptography and privacy-preserving smart contracts.” University of Maryland and Cornell University, 2015.

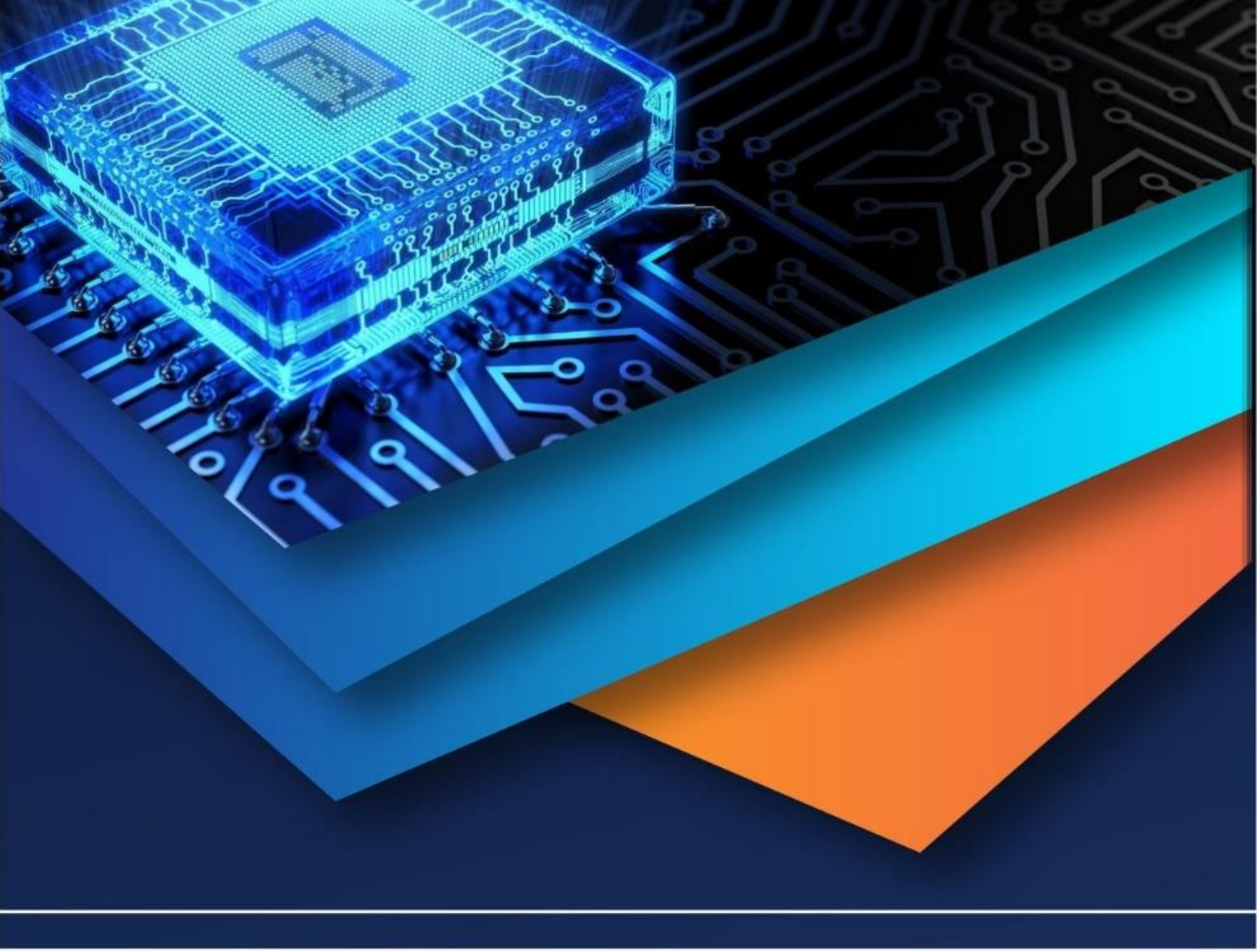
[12] European Security and Markets Authority, “The distributed ledger technology applied to securities markets.”, 2016

[13] https://www.esma.europa.eu/sites/default/files/library/2016-773_dp_dlt.pdf. Accessed 31 Oct 2017

[14] Aitzhan, N. Z., & Svetinovic, D. “ Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams.” IEEE Transactions on Dependable and Secure Computing. 2016

[15] Nash, K. S. , “IBM pushes blockchain into the supply chain.”, The Wall Street Journal. Available online: https://www.wsj.com/articles/ibm-pushes-blockchain-into-the-supplychain-146852882 4. 2016

[16] Peters, G. W., and Panayi, “Understanding modern banking ledgers through blockchain technologies: future of transaction processing and smart contracts on the internet of money,” in Banking Beyond Banks and Money, 2016



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY