

# Big Block Chain Data

A. Kumaravel, A. Rama, G. Ayyappan

**Abstract:** The expression 'Big Block Chain Data' is coined to convey the tight correlation and interlinking between two recently advancing technologies namely Big Data and Block Chain. Issues such as data protection and enormous information scalability inevitably require tradeoff to address more issues. This element is more crucial when considering big data, which is a progressive aspect of data science and cloud security in the network. Because it deals with extremely large information volume and cannot be addressed using standard techniques of data processing. The aim of this paper is to filter the requirements presented by the multi-disciplines areas like business, computing, law, etc in the form of feasible and satiable implementations for combining these two core technologies.

**Keywords:** Block Chain, Big data

## I. INTRODUCTION

With the support given by the big scale adoption of block chain induced bit coin, they are gaining ground in entire financial and business worlds. The yielding of secure and fast transactions could potentially change the patterns growing big data systems. According to a survey by KMPG [1-3], initially one-third of decision making administrators trust their company's data. With the introduction of block chain technologies, the scenario changed and the trust considerably strengthened, moreover real time applications become much more commonplace. The nature of block chain as listed below that also shows the demand for big data processing as well.

- The speed of data storage, data sharing, the quality of data and also the quality of data analytics is assured in block chain when all parties involved in a transaction have access to the same data.
- A detailed register of all transactions is kept in a single "file" which is the most important requirement of block chain. Elimination for the needs for multiple systems and complete overview of a transaction from start to finish is provided.
- A third-party mediator or centralized repository is eliminated for individuals to manage and control their personal data [4-7]

## II. BLOCKCHAIN TECHNOLOGY

Now a day's Crypto currency has given its popularity in both industry and academics. Bit coin is one such crypto currency that was very successful since it reached 10 million dollar capital marketing. Block chain is a decentralized digital ledger which consists of a collection chain of blocks. Each block is encrypted using a SHA 256 algorithm and a hash of

every block is stored in the successive block in block chain. The contents in a block cannot be altered without altering the contents in all other blocks [8-10].

Block Chain is a time-tested sequence of unchanging information recording that is operated by a computer cluster and not held by a single individual. Each block is well protected by cryptographic chain values and is linked to each other. Since the block chain has no central authority, the immutable record can be stored by everyone and their actions are accountable to everyone involved. There is no transaction cost in a block chain, but it includes infrastructure costs. The block chain is a easy way to fully automated and securely transfer data from A to B [11-13].

## III. ARCHITECTURE

The block structure of block chain is as shown in Figure 1 contains a block header and transaction details.

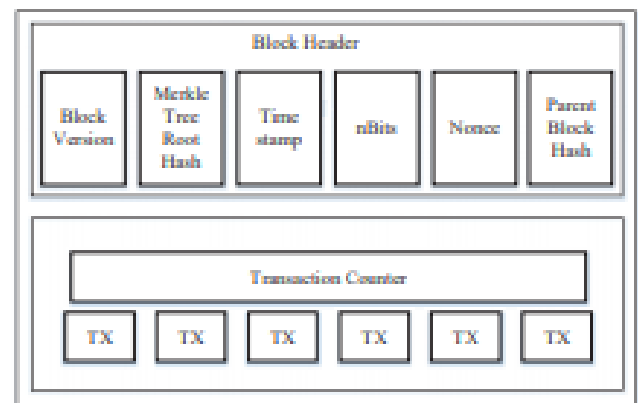


Figure 1 Block structure of Block Chain

The block header - contains details about the block version which specifies a set of validation rules used by the block, merkle tree root hash –the hash value of all the transaction in the block, time stamp – current time of creation of the block in seconds, n bits- threshold set for a valid block, Nonce – a number that is used by consensus algorithm for creating a block which can varied by the miner in order of consensus process. Parent block hash is a 256 bit hash value that points to the previous block [14-18].

**Revised Manuscript Received on August 22, 2019.**

**A. Kumaravel**, Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.

**A. Rama**, Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.

**G. Ayyappan**, Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, Tamilnadu, India.

### IV. STEP-BY-STEP PROCEDURE IN BLOCK CHAIN TRANSACTION

- BOB is going to make a transaction to Alice. The transaction will be triggered by BOB and broadcasted to all the parties in the network.
- The transaction has to be validated and for that the consensus algorithm is used between parties. The parties in the network is allowed to generate a hash code based on their nonce value and if it is less than threshold that party is allowed to create a new block and communicated to all the participating nodes
- Once the transaction is valid it is fed into the new block along with the hash of the succeeding block.
- The new block is the appended to the chain of all the communicating parties [19-21].

### V. BLOCKCHAIN AND BIG DATA

Big data is the compilation of a huge datasets. As the information is highly large and complicated, traditional data processing system is not appropriate for big data. It is possible to use these information sets to assess behavioral patterns and assess market trends. In day-to-day activities, the demand for big data analytics is very large. In ever-increasing real-world apps, the demand for big data is very high. Much transactional data is collected on the internet in various types of ledgers. To record and store all the information, many companies use cloud-based web storage. But for businesses and organizations, this is expensive. One of the most significant tasks in Big Data Analytics is to analyze the information to forecast business reaction.

Blockchain is an effective data storage method internet. The decentralized network can also provide many customers with access. Different participants in a transaction shop in distinct ledgers their transaction data. Block Chain enables all sides to a single network All transactions can be registered as blocks in the network that can be further validated by all sides involved. Since all the data will be deposited in the block chain, accessing these details will be more comfortable. Due to the Blockchain technology architecture, consumers can easily view historical operations including timeline details and other data. Following the roots of any transaction becomes very easy and effortless.

Using the Blockchain technology to store Big Data can save business costs, as it totally prevents third-party job. Blockchain has the ability to store large amounts of data for a very long time as well. Companies now have a day to choose to store information on a decentralized network system due to the benefits that prevail in the block chain. Companies and MNCs will not have to incur expenses for data storage systems due to these techniques. The added benefit of using the Blockchain to manage Big Data is the ease of applying Smart Contracts. The smart contracts can carry out operations automatically by coding the needed data. By preventing the third party, this can have a significant effect on lowering transaction costs [22-25].

Cloud computing and cloud storage servers help many organizations to collect and manage a huge volume of data. These data comes from corporate companies, online forums, objects of IOT or even from many unstructured sources.

Many data analytics tools like Hadoop is already available in the market. These tools can help only in analyzing the data but in no ways they can check the validity of data. There are uncountable number of sources that feed data into storage location and there are also many possible ways that accidentally or purposely introduce errors into the data. This contamination in data not only restricts the usage of data by enterprises but also avoids sharing it with others. Block chain technology for managing this big amount of data is a solution to this problem.

Block chain technology gained its popularity through bit coin when most companies where look for an alternate payment mode. Block chain characteristics can be implemented not only to economic fields but also to any sector and can be used to validate information and provide data security and big data sharing. Block chain is a digital ledger-forming distributed database. Every participant in the block chain must agree on changes in the folder, which is accomplished through a compromise system such as work proof or stake proof. The hash algorithm SHA256 and timestamp guarantees the immutability, verification and traceability of information on a block chain. But if the fed data is of low quality then block chain cannot transform it into a good quality data. So the companies that want to use block chain has to ensure the quality of the data being handled by providing necessary standards [26].

### VI. APPLICATION IN WHICH BLOCKCHAIN COULD IMPROVE YOUR BIG DATA

The first application of block chain to big data is in finance. Bit coin comes under crypto currency and all the transactions in bit coin are held within block chain. The transaction in bit coin is not private. By getting sufficient information about the patterns in bit coin transaction it is also possible to forge the data. There are many solutions already available in market like Skry which is recently purchased by Bloq help companies to avoid transactions with criminal actions.

Block chain if viewed as a database can be applied to many industrial domains and not just finance alone. In Industry it can be used to store the interaction details between a customer and supplier. Block chain can be very useful when sensors are added to the products to check its lifetime and durability Walmart is an American based multinational company that operates on food products is using block chain to trace the food products from its origin to the consumer. Walmart is producing a 40 petabytes of daily data efficiently stored and managed using block chain. This increase the customer satisfaction and they are sure that their product is reliable. Block chain improves big data in health information and hospital management too. The transaction between doctors and patient proves to be a high volume of data that has to be kept private. By introducing suitable security measures these data can be stored using block chain. This transaction can also store the details regarding treatment given to the patient and can be kept private and secure [27].

### VII. BIG DATA TOOLS FOR BLOCKCHAIN

Many companies are nowadays interested in developing big data tools for block chain. One



such example is BigChainDB. It is built over MongoDB.

Why big data not suitable for block chain?

- o Unlike bigdata (distributed database), block chains are not scalable
  - o Real time transactions are not supported by block chains
  - o Block chains do not have high transaction writing throughput.
  - o Searching and querying a particular data is not possible in block chain
  - o Block chains are not suitable to keep large data payloads
- So by including the highlights required for big business advancement, including versatility, inquiry capacity, and review trails, it will end up simpler for associations to assemble blockchain-based applications that satisfy corporate guidelines. The spread of these apparatuses, alongside the proceeded with push for advanced change that expects organizations to utilize carefully created and gathered information, will drive organizations to embrace blockchain so as to successfully out-contend others in their enterprises.

Due to the wide spread application of block chain in many domain the demand for block chain developers is increased in two fold. Many freelancing platforms has kept block chain skills as most demanding skill. In professional sectors also the demand for block chain developer has be increased widely.

### VIII. BLOCKCHAIN USE CASES IN BIG DATA

#### A. Ensuring Trust (Data Integrity)

Information recorded on the blockchain are reliable in light of the fact that they more likely than not experienced a check procedure which guarantees its quality. It likewise accommodates straightforwardness, since exercises and exchanges that happen on the blockchain system can be followed. A year ago, Lenovo exhibited this utilization instance of blockchain innovation to identify false reports and structures. The PC goliaths utilized blockchain innovation to approve physical reports which were encoded with advanced marks. PCs prepare the sophisticated marks and confirm the archive's reality through a blockchain record. For the most part, information uprightness is guaranteed when subtleties of the cause and associations related to an information square are placed on the blockchain and therefore checked (or approved) before it can be followed up very well [28].

#### B. Preventing Malicious Activities

Because blockchain utilizes consensus algorithm to check operations, it is inconceivable that the data assemble poses a risk to a lonely unit. A node (or device) that begins to behave exceptionally can be acknowledged and cancelled from the scheme without much of a stretch. Since the scheme is so distributed, a lone meeting makes it virtually incomprehensible to generate enough computing ability to alter the consent requirements and allow undesirable data within the structure. To alter the laws of blockchain, more hubs need to be pooled together to create an arrangement. This is not going to be possible for accomplishing a lone horrible on-screen personality [29].

#### C. Making Predictions (Predictive Analysis)

Blockchain information, much the same as different sorts of information, can be investigated to uncover profitable

experiences into the practices, patterns and in that capacity can be utilized to foresee future results. Additionally, blockchain gives organized information assembled from people or individual gadgets. In prescient examination, information researchers' base on huge arrangements of information to decide with great exactness the result of get-togethers like client inclinations, client lifetime esteem, dynamic costs, and stir rates as it identifies with organizations. This is, in any case, not restricted to business bits of knowledge as practically any occasion can be anticipated with the correct information investigation whether it is social conclusions or speculation markers. Also, because of the circulated idea of blockchain and the enormous computational power accessible through it, information researchers even in littler associations can attempt broad prescient examination undertakings.

### IX. REAL-TIME DATA ANALYSIS

As has been shown in monetary and installment frameworks, square chain sets aside a few minutes cross outskirts exchanges. A few banks and fintech pioneers are presently investigating square chain since it bears fast—actually, genuine time—settlement of enormous aggregates independent of geographic obstructions. In a similar way, associations that require continuous investigation of information in huge scale can approach a square chain-empowered framework to accomplish. With square chain, banks and different associations can watch changes in information progressively causing it conceivable to settle on speedy to choices—whether it is to hinder a suspicious exchange or track unusual exercises [30].

#### A. Manage Data Sharing

The information got from data structure can be organized in a blockchain. Along this way lines, venture groups don't rehash information investigation previously did by different groups or improperly reuse information that is as of now been utilized. Likewise, a blockchain stage can help information researchers adapt their work, most likely by exchanging investigation results put away on the stage.

### X. CONCLUSION

Big data needs the handling of extremely big information. It will be very costly to follow the block chain to manage big dat. Compared to large information gathered per second in big data analytics, the block chain treats comparatively tiny amounts of information.

How block chain is going to address is huge requirement is a quite interesting fact that has to be considered. But it is very well possible by just improving the scalability, privacy and security of the existing block chain it can be easy adopted to handle big data. When executed on a Blockchain arrangement, Big Data Analytics can have an extra layer of security. The main safety layer can be attributed to the blockchain's system engineering. Unlike the usual strategies, the system data is secure and can not be reproduced. Using Blockchain to store Big Data can also create the



information structured and meaningful. It turns out that investigating data is much more efficient and simpler. Using the development of Blockchain can also assist identify extortion. Banks and money related establishments can follow an exchange to its inceptions and recognize any oddities.

### REFERENCES

- [1] Gowri Sankaran, B., Karthik, B. & Vijayaragavan, S.P. 2019, "Weight ward change region plummeting change for square based image huffman coding", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 4313-4316.
- [2] Gowri Sankaran, B., Karthik, B. & Vijayaragavan, S.P. 2019, "Image compression utilizing wavelet transform", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 4305-4308.
- [3] Kandavel, N. & Kumaravel, A. 2019, "Offloading computation for efficient energy in mobile cloud computing", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 4317-4320.
- [4] Vinoth, V.V. & Kanniga, E. 2019, "Reversible data hiding in encrypting images-an system", International Journal of Engineering and Advanced Technology, vol. 8, no. 6, pp. 3051-3053.
- [5] Selvapriya, B. & Raghu, B. 2019, "Pseudocoloring of medical images: A research", International Journal of Engineering and Advanced Technology, vol. 8, no. 6, pp. 3712-3716.
- [6] Senthil Kumar, K. & Muthukumaravel, A. 2019, "Bi-objective constraint and hybrid optimizer for the test case prioritization", International Journal of Engineering and Advanced Technology, vol. 8, no. 6, pp. 3436-3448.
- [7] Kavitha, G., Priya, N., Anuradha, C. & Pothumani, S. 2019, "Read-write, peer-to-peer algorithms for the location-identity split", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 445-447.
- [8] Kaliyamurthie, K.P., Michael, G., Anuratha, C. & Sundaraj, B. 2019, "Certain improvements in alzheimer disease classification using novel fuzzy c means clustering for image segmentation", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 599-604.
- [9] Kaliyamurthie, K.P., Sundarraj, B., Geo, A.V.A. & Michael, G. 2019, "RIB: Analysis of I/O automata", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1019-1022.
- [10] Velvizhi, R., Rajabhushanam, C. & Vidhya, S.R.S. 2019, "Opinion mining for travel route recommendation using Social Media Networks (Twitter)", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 508-512.
- [11] Kavitha, R., Sangeetha, S. & Varghese, A.G. 2019, "Human activity patterns in big data for healthcare applications", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1101-1103.
- [12] Pothumani, S., Anandam, A.K., Sharma, N. & Franklin, S. 2019, "Extended VEOT framework - Implemented in a smart boutique", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 762-767.
- [13] Kaliyamurthie, K.P., Michael, G., Krishnan, R.M.V. & Sundarraj, B. 2019, "Pseudorandom techniques for the internet", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 915-918.
- [14] Aravindasamy, R., Jeffrin Rajan, M., Rama, A. & Kavitha, P. 2019, "Deep learning provisions in the matlab: Focus on CNN facility", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 990-994.
- [15] Theivasigamani, S., Linda, M. & Amudha, S. 2019, "Object sensing and its identification & motion sensing", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 545-549.
- [16] Mary Linda, I., Vimala, D. & Shanmuga Priya, K. 2019, "A methodology for the emulation of IPv4", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 848-852.
- [17] Velvizhi, R., Priya, D.J., Vimala, D. & Linda, I.M. 2019, "Increased routing algorithm for mobile adhoc networks", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1606-1608.
- [18] Sangeetha, S., Anuradha, C. & Priya, N. 2019, "DNS in real world", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 937-940.
- [19] Geetha, C., Vimala, D. & Priya, K.S. 2019, "Constructing multi-processors and spreadsheets with SKIVE", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 516-519.
- [20] Yugendhar, K., Sugumar, V. & Kavitha, P. 2019, "A novel method of univac using fuzzy logic", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 435-437.
- [21] Kaliyamurthie, K.P., Michael, G., Elankavi, R. & Jijo, S.A. 2019, "Implementing aggregate-key for sharing data in cloud environment using cryptographic encryption", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 957-959.
- [22] Jeffrin Rajan, M., Aravindasamy, R., Kavitha, P. & Rama, A. 2019, "A novel method of object orientation variation in C++ and java", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 708-710.
- [23] Nayak, R., Dinesh, S. & Thirunavukkarasu, S. 2019, "A novel method improvement of rapid miner for the data mining applications", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 457-460.
- [24] Sivaraman, K., Krishnan, R.M.V., Sundarraj, B. & Sri Gowthem, S. 2019, "Network failure detection and diagnosis by analyzing syslog and SNS data: Applying big data analysis to network operations", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 883-887.
- [25] Vimala, D., Linda, I.M. & Priya, K.S. 2019, "Decoupling online algorithms from erasure coding in DNS", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 950-953.
- [26] Rama, A., Kumaravel, A. & Nalini, C. 2019, "Preprocessing medical images for classification using deep learning techniques", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 711-716.
- [27] Sangeetha, S., Srividhya, S.R., Anita Davamani, K. & Amudha, S. 2019, "A procedure for avoid overrun error in universal synchronous asynchronous receiver transmitter (usart) by utilizing dummy join and interrupt latency method", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 657-660.
- [28] Aravindasamy, R., Jeyapriya, D., Sundarajan, B. & Sangeetha, S. 2019, "Data duplication in cloud for optimal performance and security", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 1156-1158.
- [29] Aravindasamy, R., Jeffrin Rajan, M., Sugumar, V. & Kavitha, P. 2019, "A novel method on developing superblocs and the transistor using apodyral", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 9 Special Issue 3, pp. 982-985.
- [30] Sasikumar, C.S. & Kumaravel, A. 2019, "E-learning attributes selection through rough set theory and data mining", International Journal of Innovative Technology and Exploring Engineering, vol. 8, no. 10, pp. 3920-3924.

## AUTHORS PROFILE



**A. Kumaravel**, Dean Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, India



**A. Rama** Assistant Professor, Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, India



**G. Ayyappan** Assistant Professor, Department of Information Technology, Bharath Institute of Higher Education and Research, Chennai, India