JOURNAL OF SOFTWARE ENGINEERING & INTELLIGENT SYSTEMS
ISSN 2518-8739
31 August 2019, Volume 4, Issue 2, JSEIS, CAOMEI Copyright © 2016-2019
www.jseis.org

SYSTEMATIC LITERATURE REVIEW OF BLOCKCHAIN IN REQUISITES OF BIG DATA AND ITS APPLICATIONS

UMAIR ASLAM, SAAD MANZOOR, MUHAMMAD IMRAN BABAR

Bahria University Islamabad, Pakistan
Army Public College of Management & Sciences, Rawalpindi, Pakistan
Email: umair.aslam@shifa.com.pk, xardar786@gmail.com, drimran@apcoms.edu.pk

Academic Editor: Hikmat Ullah Khan, COMSATS University, Wah, Pakistan.
Published: 31st August, 2019
Funding: Authors have no funds.
Article Type: Research Article

ABSTRACT

The word Blockchain has introduced a new technology which helps to utilize the distributed ledger system using distributed transactional databases. The blockchain works on three main categories. The first one, decentralization which contains no centralized authorities that dictates to other blocks. The second one is transparency, which dictates easy validated transactions however it allows us verifiable transparency in each mechanism. The last one is immutable which restrict us to temper the data because once the data has been changed it will agitate the whole chain of data. The immutable pillar is actually used for security purpose in blockchain technology. Blockchain technology uses peer to peer network that allows to join new block in chain without affecting other blocks. In this paper, we are discussing about recent approaches, methodologies and recent trends which are being used in blockchain technology and big data.

Keywords: blockchain; big data; applications; systematic literature review; ledger system; distributed transactional databases;

1. INTRODUCTION

Blockchain contains the set of blocks which carries information [1]. In 1991, this technology was introduced by researchers and in 2009 *Satoshi Nakamoto* was adopted in the form of crypto-currency bit-coin [2, 3]. It is the collection of distributed system that has completely access to anyone [4]. The most frequent feature of this technology is, once the data has been entered in the chain of block then it's very tricky to be changed in the data [5]. Each block has three different specification, data, hash and hash of previous block. Data can be of variant form that can be stored into block for example medical data, transactional data and stock market data etc. Each block has its own hash which is known as an address of the block [6]. Previous hash is known as the address of previous block. In the blockchain, first block does not have its previous block so its previous hash should be zero that is why it's known as genesis block. In blockchain technology, proof of work mechanism is used for creation of new blocks, proof of work calculates the hashing (hash, previous hash) of block in few seconds.

Blockchain technology uses peer to peer methodology which allows the block to join the chain [5]. When new block adds in the chain, all other blocks receive copies of the whole blockchain. In peer to peer, other blocks verify the newly added block and checks out everything. In this technology, all the blocks create consensus and the other blocks assure that newly added block is either valid or not [7]. Due to high security in blockchain, if someone tries to temper the block then that will be rejected by other networks of the block. If you want to successfully change in block [6], you must need to get acknowledgement from other blocks of the chain and also you will need to redo proof of the work, so it is very harsh to temper the block, blockchain technology data is evolving with the passage of time [3]. In this paper, we have identified how blockchain technology managed big data [8]. Big data is also a main trending technology currently, so it's very complicated issue that how can big data survive by the usage of blockchain technology [1]. Distinguish researchers are working on this technology to make big data secure (hospital, educational, military etc) and can be processed well [9, 10].

This paper has 5 sections in it and in its f Section 1 we briefly described about blockchain and big data technology and described that how this technology is growing day by day. Section 2 is about literature review in which we gathered data from well-known resources based on our research questions. Moreover, a brief description of our methodology is also presented in this section. As well as Section 3 presents critical analysis, in this section we analysed the usage of blockchain technology based on the data of the research questions designed in section 2. our research questions by using briefly references. Moreover, Section 4 is about conclusion of research paper that concludes this research.

Corresponding Author: Umair Aslam
Email Address: umair.aslam@shifa.com.pk

2. LITERATURE REVIEW

This research consists of scholarly research papers from reputed journals, conferences and books which are consisting of 40 well known resources top world universities journals by IEEE, ACM, Elsevier and ScienceDirect. The purpose of this literature review is to identify the existing tools, approaches, methodologies and actually big data and blocks chains which are consistently working in industry. In literature review we observed its latest technological trends and researchers and industrialists who are simultaneously working on it and improving this technology based on the existing shortcomings.

Our research is based on four critical research questions in order to find out the appropriate research directions in the domain of blockchain and big data. The research is based on following four key questions.

RQ1: What is blockchain technology and how it works?

RQ2: What are the block chain processes and methodology?

RQ3: What are the pros and cons of blockchain technology?

RQ4: How blockchain handles big data in real life and how they both work in industry and their applications?

Our research is focus on these above stated questions, so in this paper we are going to evaluate them.

RQ1: What is blockchain technology and how its works?

Blockchain is a technology, a universal online database that can be used by anyone and anywhere using internet connectivity. Blockchain does not allow supporting faking documents, transactions and other information [1]. Blockchain is a spreadsheet a revolutionary picture which has generated thousands of clones across computer network [9]. After that computer network which is designed using blockchain, to update the spreadsheet regularly and you should have the basic perception about the understanding of blockchain concept [9].

Blockchain contains the collective information and frequently submissive database. The Blockchain is a distributed database in which every block can get the copy of newly entered record which means that the information it keeps is truly publically accessible for everyone but no one can change any of the record [8]. Each block has an enticement for processing and verifying the transactions [8]. So, that's why these characteristics make the blockchain wonderful to keep the records transparent and everyone can use it [9].

A blockchain consists of several blocks in a chain where each block contains data, the hash (address) of the block and the hash of preceding block [8]. Data section contained smart contracts, transaction and other confidential records. Hash itself is an identification of the block. Hash of block is automatically developed when new block is formed in chain [11]. The third element of the block is preceding hash that is also known as previous hash. It is an address or hash of previous block in chain [12]. The benefit of blockchain is that it removes the function of a middle man between two parties [8, 11]. Private signing is mandatory for transaction in blockchain [13]. Transaction can be in various ways like smart contracts, money etc. The process of transaction in data consists of rules, logics and source and destination as well as other validated information.

RQ2: What are the blockchain process and methodology?

Blockchain technology used the concept of peer to peer network in which every node get the copy of other node [9], In this chain there is no central authority, every node of chain or network performs as a client or server. The blockchain transactions must be validated and authorized [12]. Every node of network must verify the other node, and due to this, middle man effort is removed from this technology [1]. Blockchain is also known as decentralized database that stores confidential record of customer and also managed it efficiently [1]. When new transaction occur in this network, then existing node validate it and keep the record of new transaction then new block is formed in chain that is also known as mined of transaction, and after creation of new block proof of work performs automatically [12] in which hash of block is generated for the identification of block, then previous hash also develop to engage block with chain, each block has unique pointer or parent block address [9]. So, in this way chain of blocks formed that's also known as blockchain.

Blockchain technology is also considered as secured, when someone wants to do tempering in data, hash of block automatically change and then next block previous hash is considered different, invalid block will formed due to tempering in any block then chain of block also get disturbed [8]. If someone wants to do legal change then they will perform proper methodology that's known as proof of work, every node of network validates the node and get the copy of this node also then change will occur [1]. So, in short tempering in data is not easy in this technology [13]. The ledger in this technology is used to remove the impurity or transparency, the main concept of ledger is publically access of chain for everyone [12]. Invalid operation is automatically rejected and creates

no single block when other nodes of network not verified or reject it. All data or information in network contains encrypted [11].

RQ3: What are the pros and cons of blockchain technology?

Every technology has its own advantages and disadvantages, we've distinguished the pros and cons of blockchain technology but here we are going to discuss only major advantages and disadvantages of this technology. Once a new block is validated and added into chain, this addition is permanent and can't be tempered or removed and this feature of blockchain is known as immutable. This feature of blockchain leads to increase security in network and is an easiest accountability of network [8]. For enlarged security of this technology encryption and decryption is used to keep the data secure from unauthorized person/ user. In this technology, transactions also required validation from the other nodes of network [8]. Blockchain technology eliminates the process of middleman, due to high security and confidential process anyone can easily perform transaction [8].

In this technology, every transaction must have acknowledgement of other nodes and this step leads to the accountability of blockchain [1]. In blockchain technology reversible transaction process cannot be performed This step removes fraud transaction so; transaction can't be reversible in this technology [11]. Irreversible is also considered as disadvantage if someone accidently performs transaction in this chain then it's up to next block/ customer to return back or not. In different studies, the too much verification is considered a disadvantage, if someone wants to change his/ her information or data they need to perform proof of work fully [11].

In blockchain technology data is accessible to everyone and anyone can read it easily, So, we can't store data in such a way that only validates user has access on data. Researcher proposed emerging solution but still not sure that it will efficiently work or not [13]. The major cons are considering the transaction delay. Transaction delay sometimes occurs due to proof of work, verification or validation of transaction from all nodes of network. In short, blockchain technology make possible to perform a lot of things with data like prevent data from fraud, tempering and permanent records [7].

RQ4: How blockchain handles big data in real life and how they both work in industry?

Blockchain and big data both are independent technologies but now days, both are in full stoke [12]. As we already described in previous sections about blockchain, that it is distributed database that store records permanently [15]. Blockchain and big data enhance the quality of data in various terms. Big data indicates the huge amount of data that can't be managed or processed easily [13]; it requires a lot of efforts and techniques. In earlier age, huge data managed through various processes but they took a lot of time for processing so, with the passage of time new methods and technologies introduced but they are rejected due to efficiency, security and some for other facts. Blockchain technology entered in market the main intension of this technology is security, removing scam, fraud and managing data very well, with the passage of time this technology is evolving [13]. Moreover, by using this emerging technology companies are saving their big data in blockchain and saving a lot of cost and time too [11]. Blockchain technology has ability to store the variety of data and can manage and process very well [11]. This technology also allows smart contracts to perform automatically transactions within just few seconds, it also reduces the transaction effort [12].

Now a day different sectors are using this technology. Banking sector used this technology to maintain user accounts information. In health care sector, it also used to keep the record of patient and accessible to staff [16]. In education field, educators adopted this technology for providing material to students efficiently and securely, decentralized network allows authenticated user to access data from anywhere [17]. While, the most popular use of this approach is Cryptocurrency. In 2009 first time Satoshi used this approach and developed bitcoin technology [16]. While, with the usage of this currency every transaction is made by proper non removable record and can easily be tracked [16]. Each transaction record is permanently and easily traceable in this process so that no one can change it [18]. In Japan banking sector nearly 50 banks made agreement on Ripple. Ripple is a publically accessible network formed on blockchain technology concept [19]. This system allows minimum cost and reliable transactions [20]. The association of these banks shrunk the transaction time [18, 21].

3. DISCUSSION AND ANALYSIS

3.1 Blockchain in medical care

Big data and blockchain in terms of health comes in distinct traditions like patient records, hospital, clinical, medical reports, strategies and many more, from wide resources data comes in different frequencies and terms for observation and also for analysis [7]. Hospitals require efficient tools for research to perform analysis on this data [22]. In medical care we have also observed that sometimes we need confidential and secure data, for this purpose Blockchain technology is used in this sector to make data secure, accessible to hospital staff and no one can change or play with data [23]. Medical data is also considered as critical that no one can use it with wrong aspects.

Blockchain can solve this kind of problem for big data [24]. In literature review we have observed that blockchain technology enabled securely data transfer to its consumers and also keep the record secure [25]. In blockchain no replication of data is observed, each transaction has unique hash for identification of block and for user records [26]. Medical reports are encrypted so one can easily understand and can play with data and also digitally signed [16]. In medical care we can easily store huge of data using "MedRec" approach in blockchain into three distinguish contracts. The first one is "Registrar Contract" [27]. It stores only patient unique identity, basic description about patient and allows accessing only authorized institution [28]. Second is "Patient Provider Relationship Contractor" [23]. It provides the required information to other nodes and is also known as contract between patient and custodian. The last one is "Summary Contract" [27]. This maintains proper patient's medical records, which also has brief description of patient disease [29]. Moreover, this contract makes all transactions validated and authorized by patient due to access of unauthorized person [27]. MedRec approach enables patient to view his/ her own data and also allows medical research for mining or analysing data to discover new solution of diseases [29].

3.2 Blockchain in education

Blockchain is a technology that also transforms the education sector [30, 31]. Blockchain is sole link evidence of digital event, creates duplicate on each node of network [32]. The whole record is distributed overall the networks of nodes [32]. In education sector, it stores the achievement of degrees, credit hours which are also known as degree certificates [33]. Each student's degree certificate create blocks in chain and authorized industrialist will have access to it, student can also share it with distinguish companies and can add on CV [34]. The blockchain technology assures publically student awarded degree from university and it is validated and authorized [35].

Consider a system if anyone have a big idea and wants to publish it securely without any fear so that no one can change it, they need to use this technology [36]. In this technology they can create permanent portfolio and publish the research or any kind of proof of work with proper identification [37, 38]. Anyone can also protect the work from copyrights issue.

3.3 Blockchain in crypto currency

Everyone is familiar with this technology and it's growing day by day. In world there are huge amount of Cryptocurrency formed but the most famous is Bitcoin and Ethereum [21]. It also worked as a traditional currency; users can easily buy products and services [39]. People starting trading in Cryptocurrency and saving their assets through this technology [40], Cryptocurrency is itself considering a well formed business [16, 41]. Cryptocurrency is the medium of exchange transaction developed and stored in chain mechanically, these transactions are considered secure because of using encryption and decryption methodology in this technology to examine financial units and recording of all transactions securely [16, 41]. This technology is not physically considered rather it's only available in networking chain [17]. This crypto currency is not managed or centralized by any bank but the blockchain network is overly decentralized which manages it [41]. In blockchain, blocks accumulate the information related to customer transactions, when any transaction is performed in chain, the transaction must be verified to other block either it's valid or not, after that new block is formed in chain and transaction record is stored in new block [42], whenever the creation of new block is done for unique identification, hash of block is created [19] through proof-of-work automatically, and previous hash also updates to previous block [42]. Meanwhile, in some Asian and European countries it is considered as legal but mostly it is banned in several countries by government due to some state issues [18]. Cryptocurrency have plenty of advantages it increases the transparency of every transaction which is recorded permanently, due to public access everyone can view and observe it [16]. There is an accurate tracking of transaction due to proper block hash and previous hash, no one can send back his transaction, it is also called as forward transaction [20]. Permanent ledger, in this technology, is used for permanent record of transaction [39]. Transaction cost reduces due to direct trading which saves both time and cost [18]. Faster and inexpensive, settlement saved excessive amount of money [43]. Voting can also be performed using this technology through smart phones and tablets with accurate and immediately results [43].

4. CONCLUSION

In this paper, we discussed about blockchain and big data, both are two distinguish technologies but now a day's both are working concurrently in industry and playing important role in data mining. We reviewed 40 distinct research papers of different researchers and gathered data from well-known sources and reviewed about blockchain and big data. Blockchain is basically decentralized database with the network of numerous computers around the world, so it is very complicated to temper or remove data in this technology. Blockchain ensures automatically that copy of data is stored to all the computers of network in each second. Blockchain is most trended technology that is under use in different sectors as we have discussed above briefly. Blockchain is considered secure technology, which stands in decentralization and peer to peer networks. Moreover, this technology allows us that we can merge it with big data with high security. Blockchain allows storing excessive

amounts of data due to complex network architecture. It's a perfect source for further research using big data. Meanwhile, distinct industries are using this technology to make the data secure and earn money by giving the permission of public access of data for research or analysis. Blockchain technology is also used in cryptocurrency for monetary purpose and for making transactions more secure and efficient. Blockchain is also playing an important role managing big data in education, banking, supply chain hospitals and many others with the assurance of security and is available on open source ledger for researchers and analyst.

ACKNOWLEDGEMENT

Special thanks to Bahria University Islamabad, Pakistan for providing resources to complete this research.

REFERENCE

1. Karafiloski, E., *Blockchain Solutions for Big Data Challenges* 2017.
2. Biswa, K. and V. Muthukkumarasamy, *Securing Smart Cities Using Blockchain Technology.* 2016.
3. Victoria Louise Lemieux, *Trusting records: is Blockchain technology the answer?* 2017.
4. Ammous*, D.S., *Blockchain Technology: What is it good for?* August 8, 2016.
5. Sun1, J. and J.Y.a.K.Z.K.Z. , *Blockchain-based sharing services: What blockchain technology cancontribute to smart cities.* 2017.
6. Hou, H., Sun Yat-sen University, and C. Guangzhou, *The Application of Blockchain Technology inE-government in China.* 2016.
7. Asaph Azaria, A.E., Thiago Vieira and Andrew Lippman, *MedRec: Using Blockchain for Medical Data Access and Permission Management.* 2016.
8. Yuan1, Y. and a.F.-Y. Wang2, *Towards Blockchain-based Intelligent Transportation Systems.* 2016.
9. Zibin Zheng1, S.X., Hongning Dai2, Xiangping Chen4, and Huaimin Wang3, *An Overview of Blockchain Technology: Architecture, Consensus, and Future Trends.* 2017.
10. *Big Data:The Management Revolution.* 2015.
11. Guo, Y., *Blockchain application and outlook in the banking industry*
12. Friedlmaier, M., *Disrupting industries with blockchain: The industry, venture capital funding, and regional distribution of blockchain ventures*
13. Ekblaw, A., *A Case Study for Blockchain in Healthcare:"MedRec" prototype for electronic health records and medical research data*
14. Xindong Wu, F., *Data Mining with Big Data.* 2014.
15. Labrinidis, A., *Challenges and Opportunities with Big Data.* 2016.
16. Eyal, I., *Blockchain technology: Transforming libertarian cryptocurrency dreams to finance and banking realitiesm* 2017.
17. Herbert, J., *A novel method for decentralised peer-to-peer software license validation using cryptocurrency blockchain technology* 2015.
18. Zheng, Z., *Blockchain challenges and opportunities: A survey* 2017.
19. Scott, B., *How can cryptocurrency and blockchain technology Play a role in building social and solidarity finance?* . 2016.
20. E.R.Aruna, A.R.M.R., K.V. N. Sunitha, *VALIDATION OF APPLIED SECURITY DESIGN PATTERNS USING TEST TEMPLATES IN SECURE AWARE SDLC*
21. MH Miraz, M.A., *Applications of blockchain technology beyond cryptocurrency*
22. Mertz, B.L., *Hospital CIO Explains Blockchain Potential.* 2015.
23. Esposito, C., *Blockchain: A Panacea for Healthcare Cloud-Based Data Security and Privacy?* 2017.
24. Lin2, A.Z.X., *Towards Secure and Privacy-Preserving Data Sharing in e-Health Systems via Consortium Blockchain.* 2018.
25. Sun1, L.Z.L.W.Y., *MIStore: a Blockchain-Based Medical Insurance Storage System.* 2018.
26. Imran Ihsan, F.F.K., *INTELLIGENT CLASSIFICATION OF LEARNING OBJECTS USING INFORMATION CONTENT, INTRA DOCUMENT TERMS AND DOMAIN VOCABULARY.*
27. Dubovitskaya, A., *Secure and Trustable Electronic Medical Records Sharing using Blockchain.* 2017.
28. Gräther, W., *Blockchain 4 education.* 2016.
29. Peterson, K., *A Blockchain-Based Approach to Health Information Exchange Networks.* 2015.
30. Grech, A., *Blockchain in education.* 2017.
31. Hoy, M., *An introduction to the blockchain and its implications for libraries and medicine.* 2017.
32. Sharples, M., *The blockchain and kudos: A distributed system for educational record, reputation and reward.* 2016.
33. Gräther, W., *Blockchain for education: lifelong learning passport.* 2018.
34. Turkanović, M., *EduCTX: A blockchain-based higher education credit platform.* 2018.
35. Skiba, D., *The potential of Blockchain in education and health care.* 2017.
36. B Xu, *Exploring blockchain technology and its potential applications for education.* 2018.

37. Yang, X., *The application model and challenges of blockchain technology in education.* 2018.
38. Fenwick, M., *Legal education in the blockchain revolution.* 2017.
39. Sun, S., *RingCT 2.0: A compact accumulator-based (linkable ring signature) protocol for blockchain cryptocurrency Monero.*
40. Gainsbury, S., *How blockchain and cryptocurrency technology could revolutionize online gambling*
41. Fanning, K., *Blockchain and its coming impact on financial services* 2016.
42. Li, X., *A survey on the security of blockchain systems* 2017.
43. Ali Abdrhman Ukasha, E.S.A., Rasim Amir Ali, Mabrukah Edrees Fadel, *COMPUTERIZED SYSTEM DESIGN FOR THE DETECTION AND DIAGNOSIS OF LUNG NODULES IN CT IMAGES.*

AUTHORS PROFILE