Review of Blockchain Technology to Address Various Security Issues in Cloud Computing

Parin Patel and Hiren Patel

Abstract Being distributed in nature, Cloud is an in-use technology over the last decade where the data is stored on the Cloud service provider's site where the data owner does not have direct access to the storage where her data is stored. This leads to severe security issues. Out of various such issues, data correctness or integrity is one of the major aspects of Cloud data storage security. Researchers have been exploring the options of using blockchain technology to address the data storage correctness concern in Cloud computing. Blockchain technology has provided a new dimension to address the traditional security issues in computing technologies, where every transaction is treated as an atomic operation, and once performed, it becomes immutable. The transactions are stored in a block, and such blocks are linked together to form a chain through a cryptographic hash code. Such chains are distributed in nature in the sense that there is no single central administrative authority to manage them, and any node in the network can verify the correctness of data or transaction. In this paper, we intend to make an exhaustive study on various alternatives to solve the data integrity issues in Cloud computing using blockchain technology. Being distributed, immutable, transparent, cost effective and efficient to use, blockchain is definitely a solution to many issues including Cloud data storage security problems.

Keyword Cloud computing · Blockchain technology · Security · Storage · Smart contracts integrity

P. Patel (✉)
Research Scholar, Kadi Sarva Vishwavidyalaya, Gandhinagar, Gujarat 382015, India
e-mail: patelparinv@gmail.com

H. Patel
Vidush Somany Institute of Technology and Research, Kadi Sarva Vishwavidyalaya, Kadi 382715, India
e-mail: hbpatel1976@gmail.com

K. Kotecha et al. (eds.), *Data Science and Intelligent Applications*,
Lecture Notes on Data Engineering and Communications Technologies 52,
https://doi.org/10.1007/978-981-15-4474-3_38

1 Introduction

The Cloud computing is utilized by numerous individuals in their everyday lives. Cloud computing is extremely helpful in business advancement as it gets surprising outcomes in an opportune way. Figure 1 shows the architecture of Cloud computing. There are three different types of Cloud services viz. software as a service (SaaS), infrastructure as a service (IaaS) and platform as a service (PaaS). SaaS is the service which provides online software services on pay and use basis. Client can pay to use services or applications that are hosted on the Cloud. Microsoft has Cloud-based online services like Office, Gmail, Outlook.com and Salesforce [1]. IaaS is the service which provides all infrastructure to build their application on Cloud-like servers, storage, workspace, etc. Amazon Web Services (AWS), Cisco Metapod, Google Compute Engine (GCE), Joyent and Microsoft Azure are the examples of it [2]. User can deploy their own software using platform as services (PaaS). DigitalOcean, Google Apps, Salesforce, Workday, Concur, Citrix GoToMeeting and Cisco WebEx are the examples of PaaS [2, 3]. There are three deployment models of Clouds like private Cloud, public Cloud and hybrid Cloud. Private Cloud is operated and maintained by specific organization. Private Cloud can be hosted by third party service provider or onsite of the organization. It provides more security, scalability and flexibility. Private Cloud examples are Citrix, Cisco, CSC, Dell, EMC, HP, IBM, Mirantis, Rackspace [4]. In public Cloud, Cloud service provider (CSP) makes it available for public and commercial use. Cloud provider manages all supporting infrastructure, hardware and software. Clients can access all services through Web browser. Public Cloud provides many features like scalability, law maintenance, reliability and lower cost services. Public Cloud examples are Amazon Elastic Compute, Google App Engine, IBM's Blue, Microsoft Azure, Salesforce Heroku and others

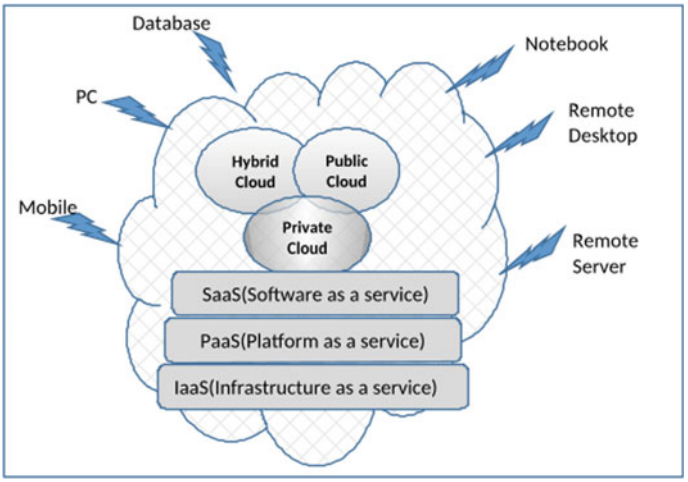

Fig. 1 Cloud architecture [3]

[5]. Community Cloud infrastructure is shared by the organizations which are having similar interest. Examples of community Cloud are Google Apps for Government, Microsoft Government Community Cloud [6]. Hybrid Cloud is the combination of the private and public Cloud. In this infrastructure, applications and data can easily move between private and public Clouds. Hybrid Clouds are more flexible, cost effective and provide control over the data. Management tools such as Cisco Cloud Center, Egenera RightScale Cloud Management and Scalr Enterprise Cloud Management Platform help businesses handle workflow creation, PAN Cloud Director service catalogs, billing and other tasks related to hybrid Cloud [7].

As the user data is stored on the Cloud service provider's site which is not under the control of the Cloud user, severe security issues may arise. Being a non-trustworthy element, CSP may play malicious and modify, destroy or use the data. Apart from this, several other issues that are mentioned beneath need also to be addressed. Multiple serious attacks like virus attack, Man-in-the-Cloud Attack, Denial of Service (DoS) Attacks, OpenStack Components Attacks and hacking of the client's data are the biggest Cloud computing data security issues [8]. When user is transferring important data on to the Cloud, it is important to ensure about security of the system. Choosing the ideal Cloud setup is also the major concern when you are adopting Cloud storage. If you are not choosing the correct Cloud, then it may be you have to face some serious hazards. Some companies have small data, so they prefer public Clouds, while huge data organizations usually use private Clouds. Cost barrier is also major obstacles for small organizations. For effective working of Cloud computing, you need to hold up under the high charges of the data transfer capacity. Data backup is the major issues faced by Cloud service in the case of data loss. There must be a proper backup policy for the recovery of data to deal with such kind of loss. Cloud management is the complex functionality of Cloud. It consists of lot of technical challenges.

At the point, when you store your information on Cloud, another person is having overall control on data. Data privacy and integrity are major issues in Cloud. You lose your privacy control on your data. Someone else is accessing your data without your permission, so it can lead to data leakage. Data is stored on the Cloud hacked or harmed by someone, so it may lead to data vulnerability. Devices are accessing your data that might not be safe. Storage gateways and API help to migrate data onto the Cloud. Data may be hacked when you store data on the Cloud.

There are multiple facets to address the security concerns of Cloud computing. Traditional cryptographic techniques are also widely used. However, researchers have started exploring the option of using recently introduced blockchain technology to solve the security problem of Cloud computing due to its decentralization and distributed nature.

Blockchain was fancied by an individual (or group of people) using the name Satoshi Nakamoto [9] in 2009 to serve as the general public transaction ledger of the cryptocurrency bitcoin. Blockchain is the distributed ledger which can be accessed by everyone on the network. In blockchain technology, anyone can join the network. A blockchain is the chain which contains data in every block. There are blocks in the blockchain, and every block consists of list of transaction, timestamp, nonce, previous hash value and Merkle root as shown in Fig. 2. Hash is the unique mathematical

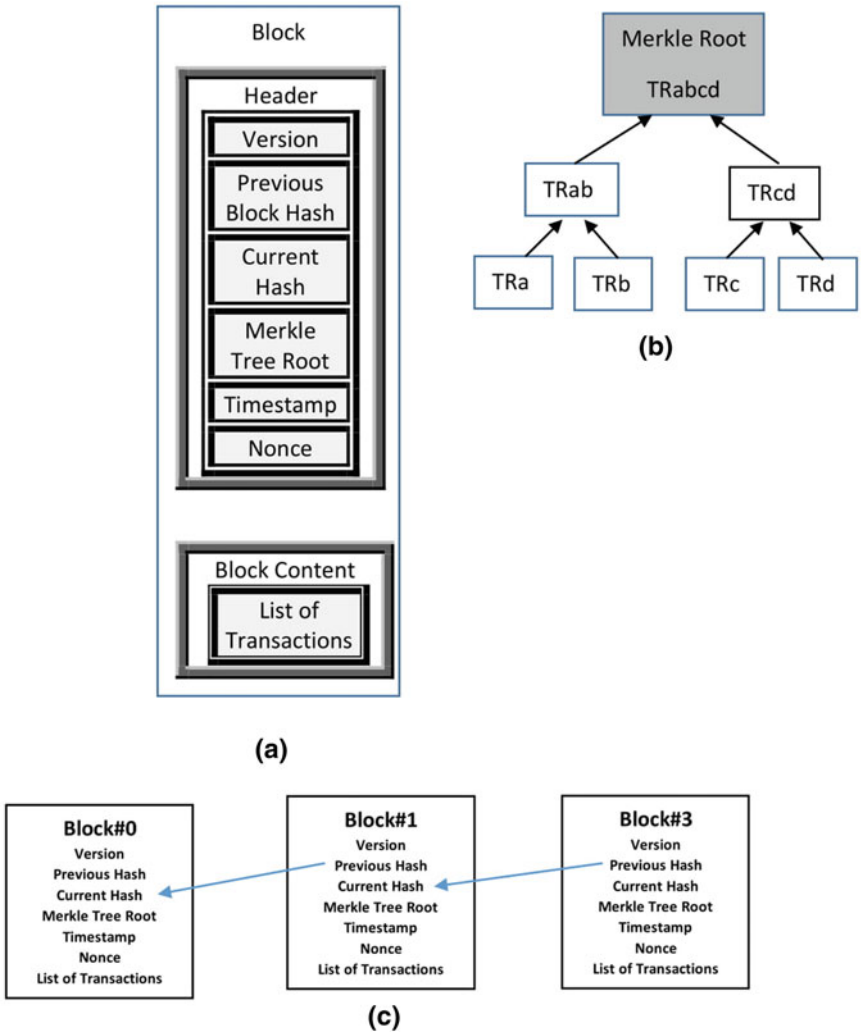

Fig. 2 **a** Blockchain block structure. **b** Merkle tree. **c** Blockchain structure [10]

code which is store with every block generated in the chain. Merkel tree is built continuously by hashing pair of nodes until only one hash left. Every leaf node is a transaction of data, and every non-leaf node is a hash of a previous hash values. Merkle tree requires even number of leaf that is why it is called as binary tree. After creation of the block, it is broadcasted through network, and then, validation process starts. Miners solve mathematical puzzles and lend their computing power. The process of solving puzzle and creating a hash is called mining. Miners required huge computational power to solve puzzles [10]. All nodes in the network check whether the calculation is correct or not. Once the majority of nodes in the network

Table 1 Blockchain platform comparison

Platform	Ethereum [12]	Hyperledger [13]	Bitcoin [14]
Smart contract	Yes	Yes	No
Data access	Public/private	Private	Public
Consensus	Proof-of-work	PBFT	Proof-of-work
Performance scalability	Low	High	Low

come to consensus and agree to the solution, the block is added to the chain. When new block is added to the chain, it updated in all existing copies into network.

Nonce value must be solved by miners. It is a complicated algorithm which is solved by miners in the network. Timestamp value is the time when block is added to the chain. The previous hash value is the hash value of the previous block. Using this previous hash value, new hash value is generated for new block. Transaction must be added when it was approved by the miners.

In a blockchain, every single node holds the copy of each and every transaction. If someone decides to alter the transaction than everyone in the network can verify that altered transaction. If someone is trying to cheat the system, then everyone in the network will not allow to perform the transaction. Noting down transaction in the ledger makes sure that double spending is not allowed. Moreover, it removes the trust we put into centralized system. Trust is developed through consensus where everyone is connected to network and making sure that the system is working fruitfully. Blockchain provides secured and tamperproof platform [11]. Comparison between blockchain platforms is described in Table 1. Ethereum and Hyperledger support smart contracts where bitcoin does not support. Ethereum supports both private and public data accessibility. Hyperledger supports private and bitcoin supports public data access. Ethereum and bitcoin use POW. Hyperledger uses PBFT. Ethereum and bitcoin have low scalability, and Hyperledger has high performance scalability.

In this paper, we study about blockchain technology and solution provided to security issues of Cloud storage. From the study, researcher can address various security issues by detailed study and understanding of this paper.

The rest of the paper is organized as follows. In Sect. 2, we discuss related work which addresses security issue in Cloud storage using blockchain technology. Section 3 focuses on the current security challenges and issues. Section 4 provides knowledge about the current tools available to implement blockchain using Cloud. We conclude our study in Sect. 5.

2 Related Work

Cloud storage uses Internet to store data over the network. Traditional Cloud storage technology handles information encryption, access control, data duplication and network management, etc. Cloud storage services cannot handle the issue of security and cannot be able to deal with the demands of users. Public key cryptography and identity-based cryptography play an important role for the security in Cloud. In public key cryptography, public–private key pair is used to encrypt and decrypt the data. Identity-based encryption used user identification, such as mobile number and email address for public key generation.

In the aspect of Cloud storage security for health data, authors of [15] have proposed a blockchain-based solution. In this proposed work, they used elliptic curve cryptography. All encrypted patients' data is stored on the blockchain using private accessible unit. Authors in [16] have proposed an architecture that protects data from unauthorized access using blockchain technology. In this system data stored in a distributed manner and manage access permission of files. Holt [17] has proposed architecture where any attempt to modify data cannot be accepted. In this paper, they have implemented log verification process using cryptographic solutions. Authors [18] have presented the work on privacy of audit logs using blockchain technology. They have used linked database technique which provides privacy and immutability to logs. Authors in [19] have discussed about the issue of log storage and proposed a secure log storage using blockchain which helps to crate immutable logs. It uses blockchain consensus algorithm for validation for logs.

Authors in [20] used genetic algorithm for file block replica placement problem between multiple users and multiple data center. Use blockchain to store metadata like URLs and hash values. In the proposed work [21], authors have implemented smart contract to store file of owner's data on blockchain network. They have also implemented verification and missing file recovery using smart contract on Cloud. Using blockchain technology, authors in [22] record unalterable timestamp and generate blockchain receipt for each and every data block for validation of provenance data. Blockchain-based provenance system enhances the integrity and availability of data. Authors in [23] proposed a secure solution for data provenance using two-folder encryption method to enhance security. SPROVE [24] architecture is implemented using digital signature and encryption to provide confidentiality and integrity to provenance data. Dstore [25] allows data owners to hire the local disks of other peers to save their data in a distributed manner without using third party. In this solution, periodically, audit happens with challenge verification solution based on Merkle hash tree. It uses smart contract to maintain lease relationships. Paper [26] has proposed an architecture named as Storj which provides end-to-end encryption of data without using third party. In blockchain-based P2P, Cloud storage network user can share and transfer date without relying on third party service providers.

3 Current Challenges and Opportunities

Secure Cloud storage has implemented using blockchain technology for different domains. From the study, we have identified some security issues that must be still addressed for security of the Cloud storage. As per the current data breaches, blockchain technology still needs more security concern. As per the report from Coindesk within the first 9 months of 2018, over $927 million were stolen by hackers from cryptocurrency exchanges [27]. In 2018, numerous exquisite cryptocurrencies such as ZenCash, Verge and Ethereum Classic fell victim to 51% attacks [28]. As per the current survey, there are many data breaches which occur due to bugs in the system. System does not provide integrity unless there is a problem of key theft or loss by hackers [29]. An attacker tries to access user's key by doing various attempts on user's smartphone or personal computer. There is still key theft and loss which much be addressed. Outsourcing of data security is significant key issue. Management of access control is also one important key issue. At the point, when somebody needs to get information on Cloud, then how to give access to that clients is a likewise challenge in Cloud and blockchain. Access control violations are likewise should be tended to. In the current system, transactions are scattered; it will increase the chances of security attacks [29]. Despite the fact that the security of blockchain is ceaselessly improved, issues have proceeded to be reported.

4 Tools and Technologies

Solidity is the language which was intended to target Ethereum Virtual Machine (EVM). Developers can execute business logic in smart contract with the use of solidity. Ethereum node can be implemented using go programming language in Geth. To implement different tasks on Ethereum blockchain, Geth is used [30]. Truffle is a framework which is used to develop Ethereum-based application [30]. To convert solidity program into EVM readable format, Solc compiler is used. Ethereum-based dApps can be developed using Embark framework [30]. Metamask is a wallet which used to give connection between Ethereum blockchain and Internet browsers like Firefox or Chrome and so forth [30]. Hyperledger provides blockchain platforms and networks for software developers. It is open-source software development approach that gives transparency, longevity and interoperability [31]. Remix IDE is one of the easier browser-based tools which is used for creation and deployment of smart contracts. It provides writing, debugging, testing and deploying smart contract written in programming language called solidity [32]. Solium tool makes sure that the code is formatted and resolves security and vulnerability issues in your code. It ensures that code is free from security holes [33]. Ganache is the tool which allows to create private Ethereum blockchain and test dApps. It allows to test without paying any gas. It also manages mining speed and gas cost [34]. Mist provides you the place where

you can store ether tokens and run smart contract. It is an official Ethereum wallet which is available in Linux, Mac and Windows operating systems [35].

5 Conclusion

Data security in the Cloud environment has been a burning issue for quite a while. Apart from traditional network security approaches, the issue can also be addressed using blockchain technology. Very few researchers have plunged into this domain. We intend to address the issue of data storage security in Cloud computing with blockchain implementation. Usage of the third party between cloud user and service provider can be eliminated (due to severe concern of trust) by the notion of the smart contract in Blockchain. Another property of blockchain viz. immutability has affirmative outcomes in terms of preserving storage correctness through maintaining every transaction intact. The distributed nature of blockchain technology allows flexibility in different ways viz. (i) any node can verify correctness of transaction (ii) malicious nodes cannot add transaction to network as it has to be verified by selected number of legitimate nodes (iii) to add a block containing valid transaction in a network, the node (miner) has to prove its credibility through consensus mechanisms such as PoW and PoS. In this research, we have explored various recently proposed mechanisms to address the issue of Cloud data storage correctness through blockchain technology. We have studied such approaches in detail and analyzed them with their pros and cons. We further have discussed the current challenges and opportunities in this domain along with tools and technologies which can be used for further implementations. In the future, we intend to propose a security model for Cloud data storage with the use of blockchain technology.

References

1. Avoyan H (2019) 3 types of cloud computing services—monitis blog. Monitis blog. Available at: https://www.monitis.com/blog/3-types-of-Cloud-computing-services/. Accessed 30 Nov 2019
2. IaaS P (2019) SaaS (Explained and compared). Apprenda. Available at: https://apprenda.com/library/paas/iaas-paas-saas-explained-compared/. Accessed 30 Nov 2019
3. Lifewire (2019) Take 5 minutes to learn the basics of cloud computing. Available at: https://www.lifewire.com/what-is-Cloud-computing-817770. Accessed 30 Nov 2019
4. Anon (2019) Available at: https://www.networkworld.com/article/3007991/3-types-of-private-Clouds-which-one-s-right-for-you.htm. Accessed 30 Nov 2019
5. SAM Solutions (2019) 4 best cloud deployment models (An overview with examples). SAM Solutions. Available at: https://www.sam-solutions.com/blog/four-best-Cloud-deployment-models-you-need-to-know/. Accessed 30 Nov 2019
6. Techno-pulse.com (2019) Techno-Pulse. Available at: https://www.techno-pulse.com/. Accessed 30 Nov 2019

7. Search Cloud Computing () What is hybrid Cloud? - Definition from WhatIs.com. [online] Available at: https://searchCloudcomputing.techtarget.com/definition/hybrid-Cloud. Accessed 30 Nov 2019
8. Jabir RM, Khanji SI, Ahmad LA, Alfandi O, Said H (2016) Analysis of cloud computing attacks and countermeasures. pp 1–1. https://doi.org/10.1109/icact.2016.7423295
9. Bitcoin.org (2019). Available at: https://bitcoin.org/bitcoin.pdf. Accessed 30 Nov 2019
10. Zheng Z, Xie S, Dai H, Chen X, Wang H (2017) An overview of blockchain technology: architecture, consensus, and future trends. Available at: https://doi.org/10.1109/bigdatacongress.2017.85
11. Cheng S et al (2017) Research on application model of blockchain technology in distributed electricity market. IOP Conf Ser: Earth Environ Sci 93:012065. https://doi.org/10.1088/1755-1315/93/1/012065
12. Ethereum GitHub implementation. Available at: https://github.com/ethereum/goethereum. Accessed 13 Feb 2018
13. Hyperledger GitHub implementation. Available at: https://github.com/hyperledger/fabric-sdk-py. Accessed 13 Feb 2018
14. Bitcoin GitHub implementation. Available at: https://github.com/bitcoin/bitcoin. Accessed 13 Feb 2018
15. Al Omar A, Bhuiyan MZ, Basu A, Kiyomoto S, Rahman MS (2019) Privacy-friendly platform for healthcare data in Cloud based on blockchain environment. Futur Gener Comput Syst 95C:511–521. Available at: https://doi.org/10.1016/j.future.2018.12.044
16. Zyskind G, Nathan O et al. (2015) Decentralizing privacy: using blockchain to protect personal data. In: Security and privacy workshops (SPW), IEEE, pp 80–184
17. Holt JE. Logcrypt: forward security and public verification for secure audit logs. In: Proceedings of the 4th Australasian workshops on grid computing and e-research (ACSW'06), Tasmania, Australia, pp 203–211
18. Sutton A, Samavi R (2017) Blockchain enabled privacy audit logs. In: d'Amato C et al (ed) The semantic web—ISWC 2017. ISWC 2017. Lecture notes in computer science, vol 10587. Springer, Cham
19. Kumar M, Singh AK, Kumar TVS (2018) Secure log storage using blockchain and cloud infrastructure. In: 9th ICCCNT, IISC, Bengaluru, India, IEEE
20. Li J, Wu J, Chen L (2018) Block-secure: blockchain based scheme for secure P2P Cloud storage. Inf Sci 465:219–231. https://doi.org/10.1016/j.ins.2018.06.071
21. Li J, Wu J, Chen L, Li J (2018) Deduplication with blockchain for secure cloud storage. In: 6th proceedings of CCF conference, big data, China. Available at: https://doi.org/10.1007/978-981-13-2922-7_36
22. Liang X, Shetty S, Tosh D, Kamhoua C, Kwiat K, Njilla L (2017) ProvChain: a blockchain-based data provenance architecture in cloud environment with enhanced privacy and availability. In: 2017 17th IEEE/ACM international symposium on cluster, cloud and grid computing (CCGRID), Madrid, pp 468–477. https://doi.org/10.1109/ccgrid.2017.8
23. Asghar MR, Ion M, Russello G, Crispo B (2012) Securing data provenance in the cloud. In: Camenisch J, Kesdogan D (eds) Open problems in network security. Lecture notes in computer science, vol 7039. Springer, Berlin, Heidelberg
24. Hasan R, Sion R, Winslett M (2009) Sprove 2.0: A highly configurable platform-independent library for secure provenance, ACM, CCS, Chicago, IL, USA
25. Xue J, Xu C, Zhang Y, Bai L (2018) Dstore: a distributed cloud storage system based on smart contracts and blockchain. In: 18th international conference, ICA3PP 2018, Guangzhou, China, Part III. https://doi.org/10.1007/978-3-030-05057-3_30
26. Andrew A (2019) Blockchain security: how far have we come in 2019? Forbes, Forbes Magazine, 27 Mar 2019. Available at: https://www.forbes.com/sites/andrewarnold/2019/03/27/blockchain-security-how-far-have-we-come-in-2019/#26feb29b2457
27. Chandhok A (2019) Top five blockchain security issues in 2019. LedgerOps, LedgerOps—Elite Cybersecurity, 28 Mar 2019. Available at: https://ledgerops.com/blog/2019/03/28/top-five-blockchain-security-issues-in-2019

28. Wilkinson S, Boshevski T, Brandoff J, Buterin V (2014) Storj a peer-to-peer cloud storage network
29. Park JH, Park JH (2017) Blockchain security in cloud computing: use cases, challenges, and solutions. Symmetry 9:164. https://doi.org/10.3390/sym9080164
30. Blockchain Technology Explained: Introduction, Meaning, and Applications. Hackernoon. Available at: https://www.hackernoon.com/blockchain-technology-explained-introduction-meaning-and-applications-edbd6759a2b2 Accessed 30 Nov 2019
31. Hyperledger (2019) About—Hyperledger. Available at: https://www.hyperledger.org/about Accessed 30 Nov 2019
32. SitePoint (2019) Remix: develop smart contracts for the ethereum blockchain—SitePoint. Available at: https://www.sitepoint.com/remix-smart-contracts-ethereum-blockchain/ Accessed 30 Nov 2019
33. Dzone.com (2019) 10 tools for blockchain development—dzone security. Available at: https://dzone.com/articles/10-tools-for-blockchain-development Accessed 30 Nov 2019
34. Codementor.io (2019) Developing for ethereum: getting started with ganache. Codementor. Available at: https://www.codementor.io/swader/developing-for-ethereum-getting-started-with-ganache-l6abwh62j Accessed 20 No 2019
35. Ltd, Attores Pte (2019) Step-by-step guide: getting started with ethereum mist wallet. Available at: https://www.medium.com/@attores/step-by-step-guide-getting-started-with-ethereum-mist-wallet-772a3cc99af4 Accessed 17 Nov 2019