Study on learning resource authentication in MOOCs based on blockchain

Yonghui Dai*

Management School,
Shanghai University of International Business and Economics,
Shanghai, 201620, China
Email: daiyonghui@suibe.edu.cn
*Corresponding author

Guowei Li

School of Information Management and Engineering,
Shanghai University of Finance and Economics,
Shanghai, 200433, China
Email: liguowei@mail.shufe.edu.cn

Bo Xu

Management School,
Shanghai University of International Business and Economics,
Shanghai, 201620, China
Email: brianxubo@163.com

Abstract: In MOOCs, the learning resources authentication is a matter of great concern for the learners and teachers. Its construction is faced with the challenges of information security and privacy protection. Considering that the blockchain has the advantages of decentralisation, autonomous and non-tampering, this paper provides a solution to implement the construction based on blockchain technology, which includes the system architecture, experimental validation and key technologies such as decentralised transaction and tamper-resistance. The results prove the technical feasibility and safety reliability of blockchain to learning resource management in MOOCs.

Keywords: learning resource; massive open online courses; MOOCs; blockchain; decentralised transaction; tamper-resistant.

Reference to this paper should be made as follows: Dai, Y., Li, G. and Xu, B. (2019) 'Study on learning resource authentication in MOOCs based on blockchain', *Int. J. Computational Science and Engineering*, Vol. 18, No. 3, pp.314–320.

Biographical notes: Yonghui Dai received his PhD in Management Science and Engineering at Shanghai University of Finance and Economics, China in 2016. He is currently a Lecturer in the Management School, Shanghai University of International Business and Economics, China. His interests include machine learning and big data analytics. His works have appeared in international journals more than 20 papers.

Guowei Li received his MS in the University of Science and Technology of China in 2006. He is a PhD candidate at School of Information Management and Engineering at Shanghai University of Finance and Economics. His research interests include system simulation and data mining.

Bo Xu received his PhD in Management Science from the Fudan University, China in 1997. He is a Professor at the Management School, Shanghai University of International Business and Economics, China. His interests include strategic management and educational technology.

1 Introduction

With the development of educational information and the deepening of education reform, the education of massive open online courses (MOOCs) has developed rapidly, and more and more scholars are involved in this learning style (Watson et al., 2017). As an important part of the construction of MOOCs, the learning resource authentication and how to build the security of learning resources is the key to MOOCs security.

In the construction of traditional MOOCs resource, the learners and teachers regularly upload the learning resources to a trusted centre node for sharing (Wu et al., 2017). This centralised data storage mode is faced with information security problems and such as centralised malicious attacks, distributed denial-of-service (DDoS) attack, malicious data tampering of data centres and so on (Rottondi et al., 2016). In order to solve the above security problems, cloud computing (Kong et al., 2018), big data analysis (Lin and Chu, 2017) and other technologies are applied to ensure the safety of MOOCs. In general, the identity authentication includes password or the unique physical characteristics, such as dynamic password, USB Key, fingerprint, voice, appearance and other biological characteristics. However, the above authentication has some disadvantages, for example security, high cost or low efficiency, so the authentication method based on new technology is constantly explored. Therefore, blockchain was used for our study on authentication.

This paper is organised as follows. In Section 2, related research of blockchain technology and application are introduced. In Section 3, the system architecture of learning resource authentication is shown. In Section 4, the key technologies are illustrated. In Section 5, the realisation of authentication based on blockchain is introduced. Section 6 is the conclusion of this article.

2 Related work

2.1 Blockchain technology

The blockchain is a large-scale to decentralisation network without relying on the trust centre. It includes a collection of various technologies, such as peer-to-peer (P2P) transactions (Feld et al., 2016), decentralised databases, consensus operations, fault-tolerant mechanisms and so on (Kraft, 2016). P2P technology is different from the traditional client/server mode, and the nodes in the network have the same status and have the dual characteristics of the server and the client. Each node also provides services to the outside world while using services. P2P technology can assign successfully the burden of traditional servers to every node in the network, and every node in the network has certain storage and computing power. Therefore, the more nodes there are, the higher the quality of service will be in the network. Blockchain can be seen as a chain of data structures that combine blocks together (Wang et al., 2016): it consists of the block header and the block body. The block header is used to identify the block and is responsible for connecting to the next block, and the block body is used to store transaction information and is responsible for verifying transaction information. The blockchain structure is shown in Figure 1.

As the core of the blockchain, consensus algorithm is an algorithm for reaching the agreement in the whole network after proposing a node or multiple nodes in a distributed system. The initial application of distributed consensus algorithm is the Paxos algorithm. This algorithm has a certain ability of fault tolerance, but Paxos problem is a consensus problem in the scenario of distributed system failure but no malicious nodes. Therefore, the fault tolerance of Paxos is not Byzantine fault tolerance. Byzantine error means that in distributed system, nodes may have any errors, including system failure, timeout, sending messages repeatedly, intentionally misleading, forgery of signatures, etc. In the PBFT algorithm, the Byzantine node that does not exceed the total node 1/3 of the whole network can be tolerated, that is, if there is more than 2/3 normal nodes in the whole system, the whole system can run normally. In order to adapt to application scenarios, scholars have also improved the PBFT algorithm (Sukhwani et al., 2017), mainly focused on modifying their underlying network topology requirements, and dynamically adjusting the total number of nodes in the network.

Digital signature technology is an application of asymmetric key encryption and digital summarisation, and it guarantees the integrity of the digital information transmission, the authenticity of the sender's identity and the prevention of the occurrence of the denial in the information transaction. Digital timestamp technology is a transformation application of digital signature technology its generation process is shown in Figure 2.

It can be seen from Figure 2 that the process of the digital timestamp is as follows. Firstly, users add a timestamp to a digital file with some Hash algorithm to compute the digits of the file. Then, digest is sent to the third party timestamp service organisation (TSA). And the service organisation begin to sign it after adding the date and time of the digits, and then sends the generated digital time stamp file (*.tsa) to the user. Finally, users get authoritative proof of this digital file at a specific time and state by using this digital time stamp.

Figure 1 The blockchain structure

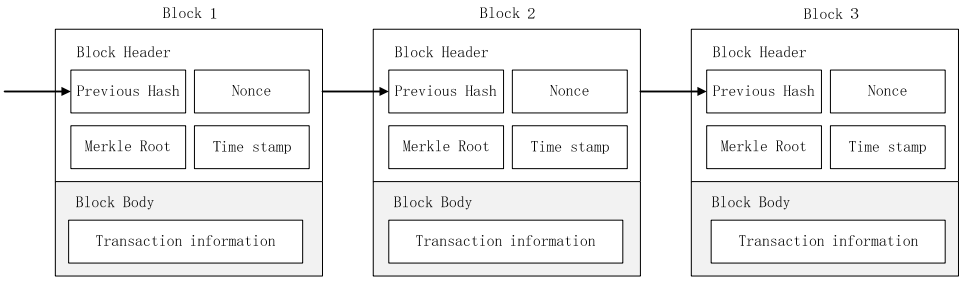
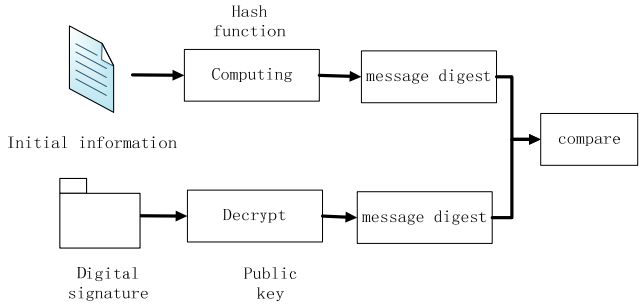


Figure 2 The digital signature process (see online version for colours)



2.2 Blockchain application

The concept of blockchain can be traced to the scholar Nakamoto, he pointed out that the time stamp can be used in the form of block in order to solve the security problem of electronic money (Nakamoto, 2008). Nowadays, blockchain has been highly concerned by governments, financial institutions and technology companies (Beck et al., 2017). For example, more than 40 major international financial institutions collaborated to develop blockchain technology, such as HSBC, UBS and Bank of America. The USA has launched a securities trading platform based on blockchain technology, which has become an important milestone in the centralisation of the financial securities market.

In the field of science and technology, the IBM joint Linux foundation has established a special blockchain with open source project Hyperledger, which has entered the substantive development stage. They will provide Watson API on IoT platform of Watson to help business customers and developers develop, so as to realise the autonomy of internet of things with blockchain (Yang et al., 2017). In the field of food, the research report of British Thad University pointed out that if the blockchain technology is applied to the food supply chain, it can solve the increasingly serious global issues such as food waste (Badia-Melis et al., 2015). In the medical field, some scholars have proposed a medical record security storage scheme, which is combined with blockchain and cloud storage technology (Mei, 2015).

The application of blockchain in education is still in the early stage, it mainly focuses on the blockchain transcript, learn ledger, and building a trust system for education. blockchain transcript is based on blockchain technology, including learners' learning process and results. This transcript can record detailed data of learners' growth experience, learning process, skills, completed learning projects, teachers' evaluation and so on. At the same time, blockchain technology can be used for distributed learning records and storage in the field of education, and it allows all educational institutions and learning organisations to record learning behaviours and learning results across systems and platforms, and permanently store them in cloud servers (Li and Zhang, 2017).

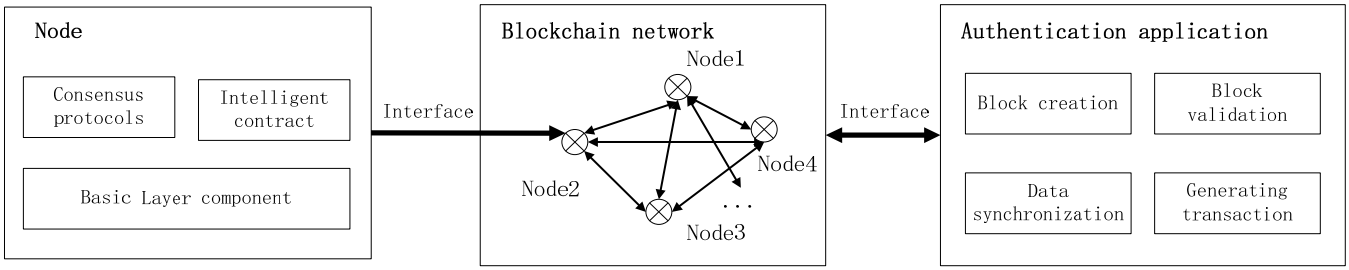
Ownership of copyright is also a major scenario of blockchain application. Because the security and reliability of the copyright protection based on asymmetric encryption algorithm is higher, and the blockchain has the characteristics of openness and transparency, any resource creation information can be queried, tracked and obtained by users, which can solve the problem of copyright ownership from the source. In addition, important information such as student achievement, personal files and academic credentials can also be stored on the blockchain, so as to build a safe, credible and tamper-proof credit system for students, and help them to solve the problems of current student credit loss and academic fraud (Hyvärinen et al., 2017).

In the construction of open education resources, the distributed technology of blockchain can help education resources be distributed stored in different blocks, all nodes can be directly shared learning courseware by specific consensus protocol software, which not only helps to improve the efficiency of sharing, but also can solve the problem of resource island. In particular, through the transparent and automatic execution of intelligent contracts, we can realise the automatic execution of resource uploading, authentication, circulation and sharing, reduce the cost of resource sharing, improve the efficiency of resource sharing, and build a new form of network resource transfer (Yang et al., 2017).

3 System architecture

In our study, the learning resource authentication system is composed of blockchain network with students and teachers, and the framework of the system is shown in Figure 3.

Figure 3 The framework of the system



It can be seen from Figure 3 that all nodes are linked by interfaces to form blockchain network and authentication application.

3.1 Blockchain network

The blockchain network is composed of many nodes with teachers and students, and these nodes implement the function of related authentication, such as block creation, block validation, data synchronisation, and so on.

Compared to the traditional centralisation system, the centralisation architecture used consensus algorithm to support the interoperation between nodes and the unified management of node behaviour. Therefore, the design of consensus mechanism is an unavoidable problem in the application and development of blockchain. The DPoS mechanism integrates the resources of the community, and its centralisation relies on a certain number of Representatives. Usually, a stable node will be recommended as a managed node, and it participate in maintaining blockchain transaction data, which can make the system more secure and centralised.

Consensus protocol is used to solve the problem of trust in the network topology, which is also the key technology in the blockchain network. In the bitcoin network, a consensus mechanism of 'A minority obeys the majority' is adopted, because more than 50% nodes in the default network are trusted nodes. However, because the evaluation system designed in this paper is a private blockchain network, the number of nodes cannot reach the order of bitcoin network, so we need to improve the trust threshold of consensus protocols. By applying the Byzantine algorithm to the consensus protocol, the trust threshold can be raised from 50% to more than 66%. This greatly strengthens the credibility and security of the blockchain network.

Intelligent contract is the interface of each node and external interaction in the blockchain. Different intelligent contracts can undertake different functions and maintain uniform agreement by consensus protocol. In order to guarantee intelligent contract work in different environment, intelligent contract uses Docker application container engine, which has good cross platform properties. As the key to intelligent contract database, ROCKSDB uses NoSQL to its database, and it has two advantages over the traditional relational database. One is the read speed, and the other is its flexibility. Because the consensus of the blockchain is based on the cost of sacrificing storage time and it uses NoSQL to form a database, which can reduce the consumption time of a large number of read data. In addition, unlike relational databases, NoSQL implements specified data structures and fields, which make blockchain networks more flexible and generality.

3.2 Authentication of interface and procession

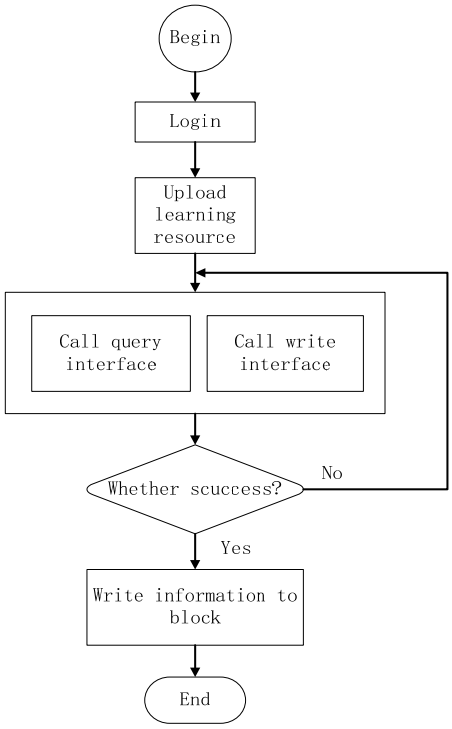
Authentication mainly includes block creation, block validation, data synchronisation, interface, and so on. Of these, interface consists of a series package function, For example, some of management interface of blockchain is shown in Table 1.

Table 1 Interface of blockchain

Interface	*Parameters and description*
GetAccountInfo	Parameters:id
	Get user's account information
GetBlockInfo	Parameters:id
	Query information of id blockchain
GetBlockNum	Parameters:id
	Query the number of blocks of id blockchain
GetActiveNode	Parameters:none
	Return the reliable node considered by the whole network

In addition, the processing of blockchain creation for learning resources is shown in Figure 4.

Figure 4 The processing of blockchain creation



It can be seen from Figure 4 that the user first logs into system, and then uploads the learning resource. In order to link these learning resources to blockchain, the transaction will be executed, and the query interface and write interface will be called. If the interface is called successfully, the information will be written to block, and the transaction is ended.

4 Key technology

4.1 Decentralised transaction

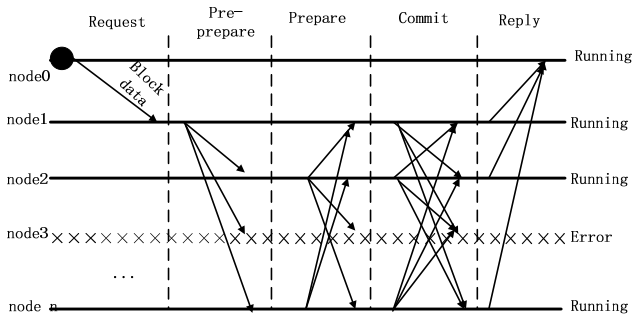
In blockchain, the rights and obligations of all nodes of the whole system are equal, and any node's damage or failure

will not affect the operation of the whole system, and it implements the decentralisation of intelligent contracts. In order to achieve the centralisation, a large number of hash operations are needed for the whole network node, the practical Byzantine fault tolerance (PBFT) consensus mechanism algorithm is applied to this study.

PBFT is a consensus agreement that tolerates malicious behaviour. The system only needs to tolerate the mistake and agree through more than half of the process, and the consensus can be reached. Generally, if there are 2^n processes, the number of erroneous processes can reach consensus as long as it is less than n. But Byzantine errors are more complex, it can tolerate a number of erroneous processes less than the number of processes that are not Byzantine systems. That is to say, the total number of nodes R must be greater than or equal to 3F plus 1, where F is the total number of problem nodes.

PBFT can be considered as three stages: pre-preparation, preparation, and confirmation. The execution is shown in Figure 5.

Figure 5 The execution of PBFT



In the pre-preparation stage, the main node assigns an integer sequence number for the request after receiving the request message and generates prep messages.

In the preparatory phase, the backup node makes sure that the message is the original message after receiving the pre-prepared message.

In the confirmation stage, when the replica node verifies that the preparation message is true, the confirmation message is broadcast to the other nodes. After verifying the confirmation message, the base point of the copy can be executed, and the result is returned to the requester.

4.2 Tamper-resistant

Timestamp refers to the data generated by digital signature technology, which includes digital signature, including original file information, signature parameter, signature time and other information, in order to prove that the original file existed before signing time. The timestamp service consists of three parts: a trusted time source, a signature system, and a timestamp database.

The trusted time source is the time source of the timestamp system. The time of all components in the TSA system must be based on this trusted time source, especially in the time stamped time, which must be filled strictly according to the trusted time source. The self-trusted time source is the time obtained by the hardware and methods approved by the national authority of time.

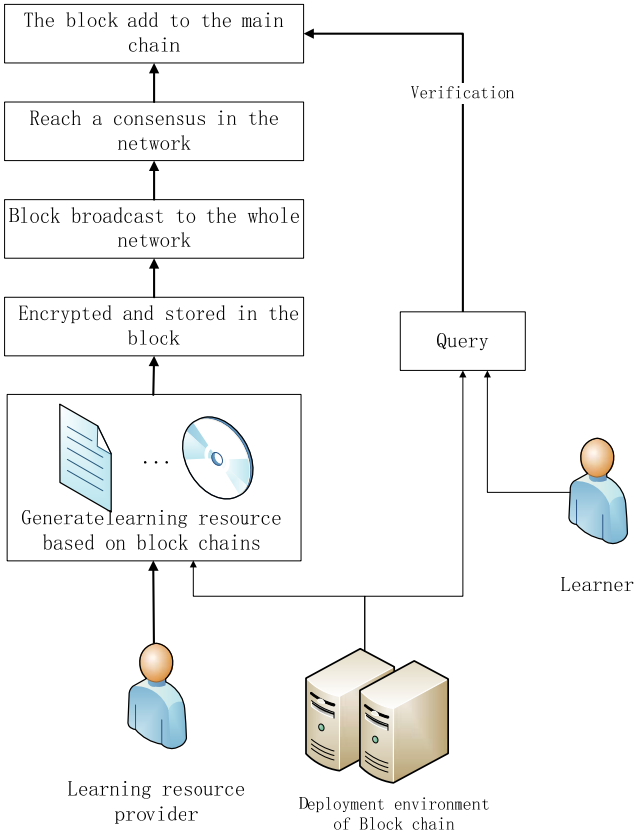
The signature system is responsible for receiving timestamp applications, validating application legality, and producing and issuing timestamps, and finally storing the timestamp in the database. In this process, the application of the message and time stamp format, the generation and issuance of the timestamp must conform to the requirements given in the specification. User sends the timestamp application to the signature system and the signature system gets the user's file data digest, and then verify the validity of the application. Finally, the current time and the file digest are bound in a certain format, and the signature is returned and stored in the database.

The timestamp database is responsible for preserving the timestamp issued by the TSA system, and must be regularly backed up so that the user can apply for a timestamp from it.

5 Experimental verification

In our study, experimental validation of the learning resource authentication in MOOCs is shown in Figure 6.

Figure 6 The framework of experimental validation (see online version for colours)



It can be seen from Figure 6 that the learning resources were generated based on deployment environment of blockchains firstly, and then these learning resources were encrypted by using the private key and were stored in the block. Next, it broadcasts the above learning resources to the whole

network and waits for authentication. Finally, if the consensus condition is reached, the block is added to the main chain. Overall, authentication nodes begin to trade in blockchain networks, and then the authentication node collects all transactions in the chain through the transaction pool at intervals. After collecting all the transactions of the above period, the mining was started, and once the authentication was passed, the node would be broadcast to the whole network, and all the authenticate nodes of the whole network stored the block on their local chain. After logging in the website, the MOOC's learners query the learners' resources, and get information from the blockchain to verify. If their validation is passed, learners can learn.

The hardware deployment of our experiment consists of ten servers, and the software includes JDK version 1.8 and Apache Maven version 3.3. Each server acts as an authentication node and runs the blockchain codes. The sample of code is shown as in Figure 7, which includes an index, a timestamp, a transaction list, and the hash value of the previous block and so on.

Figure 7 The sample code

```
block = {
    'index': 1,
    'timestamp': 1530589716.9490936,
    'transactions': [
        {
            'sender': "6bb07987b1c8828ef0aa8eb73bd00285",
            'recipient': "a7705e8a6c0a4ba0bfb25cce7b5ba390",
            'amount': 10,
        }
    ],
    'proof': 562389212361,
    'previous_hash':
"60afb923be15d0dda627c53bc714800852fc7f5cb8baee6f1a1af51531113fe5"
}
```

6 Conclusions

MOOCs model has been favoured by learners in recent years, some new information technologies are applied to study on MOOCs such as affective computing, deep learning, and big data analysis. By deepening the reform of education, MOOCs has become a trend and more technologies will be used in it. In order to promote the application of blockchain better in education, educational practitioners and technicians need to better understand the blockchain and actively participate in learning and using it.

In the future, blockchain will become another disruptive technology to promote the reform of information technology. As internet plus education involves a lot of data uploading, blockchain technology can effectively prevent data tampering, enhance data protection, and be widely used. In short, blockchain technology brings both challenges and opportunities for educational institutions, and teaching institutions should pay close attention to the research and application of the blockchain, and explore its application.

Acknowledgements

This work is supported by the project of Shanghai Higher Education Society under Grant (No. GJEL1851), Shanghai Philosophy and Social Sciences Plan (No. 2016BGL004, No. 2018BGL023), Liberal arts and Social Sciences Foundation of Ministry of Education in China (No. 16YJA630011) and International Cooperative Program of Shanghai Municipal Science and Technology Commission of China (No. 16550720500).

References

Badia-Melis, R., Mishra, P. and Ruiz-García, L. (2015) 'Food traceability: new trends and recent advances, a review', *Food Control*, Vol. 57, No. 11, , pp.393–401.

Beck, R., Avital, M., Rossi, M. et al. (2017) 'Blockchain technology in business and information systems research', *Business and Information Systems Engineering*, Vol. 59, No. 6, pp.381–384.

Feld, S., Schönfeld, M. and Werner, M. (2016) 'Traversing bitcoin's p2p network: insights into the structure of a decentralized currency', *International Journal of Computational Science and Engineering*, Vol. 13, No. 2, pp.122–131.

Hyvärinen, H., Risius, M. and Friis, G. (2017) 'A blockchain-based approach towards overcoming financial fraud in public sector services', *Business and Information Systems Engineering*, Vol. 59, No. 6, pp.1–16.

Kong, W.W., Lei, Y. and Ma, J. (2018) 'Data security and privacy information challenges in cloud computing', *International Journal of Computational Science and Engineering*, Vol. 16, No. 3, pp.215–218.

Kraft, D. (2016) 'Difficulty control for blockchain-based consensus systems', *Peer-to-Peer Networking and Applications*, Vol. 9, No. 2, pp.397–413.

Li, Q. and Zhang, X. (2017) 'Blockchain: promoting the opening and public trust of education by technology', *Journal of Distance Education*, Vol. 35, No. 1, pp.36–44.

Lin, W.T. and Chu, C.P. (2017) 'A fast and parallel algorithm for frequent pattern mining from big data in many-task environments', *International Journal of High Performance Computing and Networking*, Vol. 10, No. 3, pp.157–167.

Mei, Y. (2017) 'Research on blockchain method for safe storage of medical records', *Journal of Jiangxi Normal University (Natural Science)*, Vol. 41, No. 5, pp.481–487.

Nakamoto, S. (2008) *Bitcoin: a Peer-to-peer Electronic Cash System* [online] https://bitcoin.org/bitcoin.pdf (accessed 29 August 2017).

Rottondi, C., Panzeri, A., Yagne, C.T. et al. (2016) 'Detection and mitigation of the eclipse attack in chord overlays', *International Journal of Computational Science and Engineering*, Vol. 13, No. 2, pp.111–121.

Sukhwani, H., Martinez, J.M., Chang, X. et al. (2017) 'Performance modeling of PBFT consensus process for permissioned blockchain network (hyperledger fabric)', *Reliable Distributed Systems*, IEEE, pp.253–255.

Wang, H., Chen, K. and Xu, D. (2016) 'A maturity model for blockchain adoption', *Financial Innovation*, Vol. 2, No. 1, pp.2–12.

Watson, W.R., Watson, S.L. and Janakiraman, S. (2017) 'Instructional quality of massive open online courses: a review of attitudinal change MOOCs', *International Journal of Learning Technology*, Vol. 12, No. 3, pp.219–240.

Wu, Z.Q., Liang, Y., Kang, J. et al. (2017) 'Secure data storage and sharing system based on consortium blockchain in smart grid', *Journal of Computer Applications*, Vol. 37, No. 10, pp.2742–2747.

Yang, X.M., Li, X., Wu, H.Q. et al. (2017) 'The application model and challenges of blockchain technology in education', *Modern Distance Education Research*, No. 2, pp.34–45.