Simulation model of Blockchain System in the Higher Education

Shmatko Olexandr [1], Serhii Yevseiev [2] Vladyslav Khvostenko[3]

[1] *National Technical University "Kharkiv Polytechnic Institute" st. Kirpichova, 2, Kharkiv, 61000, Ukraine*
[2,3] *Simon Kuznets Kharkiv National University of Economics, ave. Nauki, 9-A, Kharkiv, 61166 Ukraine*

Abstract

This article presents a mathematical model of a distributed ledger for higher education. The main components of this network are considered, as well as their formal presentation. The model of peer-to-peer network is visualized, the research of the parameters of the centralized and decentralized data processing network is carried out. Based on the data obtained, simulation models were built and investigated. The results of the simulation simulations were analyzed and the most optimal parameters were selected.

Keywords

distributed ledger, blockchain, mathematical modeling, simulation, probability theory, theory of random processes.

1. Introduction

At present in the world there is a revolutionary transition from informatization of the main spheres of human activity to their digitalization.

If informatization involves, in essence, the modernization of certain human activities through the use of information and communication technologies, the digital transformation (or digitization) in its turn involves their qualitative transformation, departure from the usual types and forms of activity to the new ones, based on digital models and technologies [1,2].

The development of the digital environment requires the support and development of both existing conditions for the emergence of promising end-to-end digital platforms and technologies, as well as the creation of conditions for the emergence of new platforms and technologies.

The main end-to-end digital technologies are:

– big data;
– neurotechnology and artificial intelligence;
– distributed registry systems (blockchain);
– quantum technologies;
– new production technologies;
– industrial internet;
– components of robotics and sensors;
– wireless communication technologies;
– virtual and augmented reality technologies.

Continuing the cycle of work on the digital transformation of education [3,4], the paper conducts research on the use of blockchain technology (blockchain) for the tokenization of educational assets and promising areas of its implementation in education.

2. Literature review

In [5,6] possible scenarios for using blockchain technology in the field of education are considered. Methods and technologies of tokenization of assets, related to the educational process, are investigated. It is concluded, that the blockchain technology is decentralized and transparent with a high degree of reliability, which

III International Scientific And Practical Conference "Information Security And Information Technologies", September 13–19, 2021, Odesa, Ukraine
EMAIL: oleksandr.shmatko@khpi.edu.ua (A. 1); serhii.yevseiev@hneu.net (A. 2); vladyslav.khvostenko@gmail.com (A. 3)
ORCID: 0000-0002-2426-900X (A. 1); 0000-0003-1647-6444 (A. 2); 0000-0000-0000-1234 (A. 3)

ensures the equality of all users of the chain's services. The transparency of the technology guarantees the participants in the process against abuse and forgery of documents. The study of the features of smart contracts made it possible to form the advantages of smart contracts in the field of education

In [7], provides a critical analysis of application of the blockchain technology considering with its applicability opportunities and restrictions in education; it also aims to identify the consequences of its influence upon the development of education.

The article [8] provides an overview of the use of blockchain for academic transcripts. The aim is to find, among the proposed models, overlapping aspects that solve common problems and can lead to a universally accepted de facto standard. In addition, since academic institutions will serve as oracles for specific blockchain applications, a robustness study is underway to see if the proposed applications effectively solve the oracle problem.

The paper [9] is a Systematic Bibliometric Review of the Literature on Blockchain Applications Research in Higher Education. The review includes 37 articles that provide up-to-date knowledge on the current implications of using blockchain technology to improve higher education processes. The LRSB findings show that blockchain is being used to create new interventions to improve the prevailing ways of sharing, delivering and protecting student knowledge data and personal records.

The relevance of this work is due to the increasing popularity of distributed registry systems, in connection with which it is necessary to assess the quantitative parameters of this network and determine the most optimal parameters.

The general network model is a peer-to-peer network in which each participant has m client applications, an application server S, an N node (a server for communicating with other network nodes)

3. Simulation model

Simulation is a method of research in which the studied system is replaced by a model, with sufficient accuracy describes the real system from which experiments are conducted in order to obtain information about this system.

In favor of using the methods of simulation in this situation is the impossibility of experimenting on a real object, because then we would have to develop two full-fledged systems. Also models will allow to demonstrate work of two architectures in time and to calculate indicators for decision-making in favor of one of them.

The main parameter of the study will be the average transaction processing time of the system.

To simulate the model you need to know the following parameters:

1 Average processing time of one application;

2 Number of customers sending applications;

3 Number of servers processing these requests.

Many transactions related to smart contracts circulate on the Ethereum platform. To calculate the average processing time of one application, you need to include several assumptions and simplifications:

1) The generation time of a new block is subject to the exponential law (the covariance coefficient for this law is a constant equal to one) [7].

2) The Ethereum blockchain platform does not have the maximum possible block size and limit on the number and size of transactions, but there is a limit on the maximum amount of gas (gas, transaction fees) used in the block. This value can be reduced or increased in the next block by 20 percent [6].

When developing a mathematical model, it is assumed that the maximum number of transactions in the block will be 77. This number is taken from the average number of transactions in the block of the real network Ethereum [5], obtained as of November 2017

3) The emergence of new transactions (in other words, applications) is subject to the simplest law of distribution, namely Poisson's. In the developed mathematical model it is considered that the flow of incoming applications is the simplest, because it corresponds to the properties of stationary, ordinary and no aftereffects in the considered conditions.

Each transaction is processed sequentially and has a strict order of writing to the decentralized blockchain; this ensures the ordinary flow of applications.

A centralized system can also be considered in the context of queuing theory, because the server is a single-phase queuing system.

AnyLogic software environment is used to build a simulation model and conduct experiments. Simulation models of two systems were built using AnyLogic tools.

Input parameters of the model:

1 Number of customers sending requests

2 Number of miners in the blockchain network

3 Number of requests per 10 minutes from one client

4 Number of requests from one client

Figures 1 and 2 show simulation models of decentralized and centralized networks.

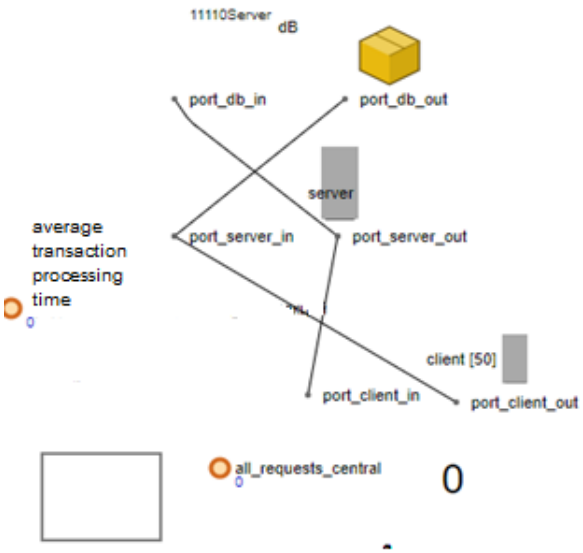

Figure 1: Simulation model of centralized network.

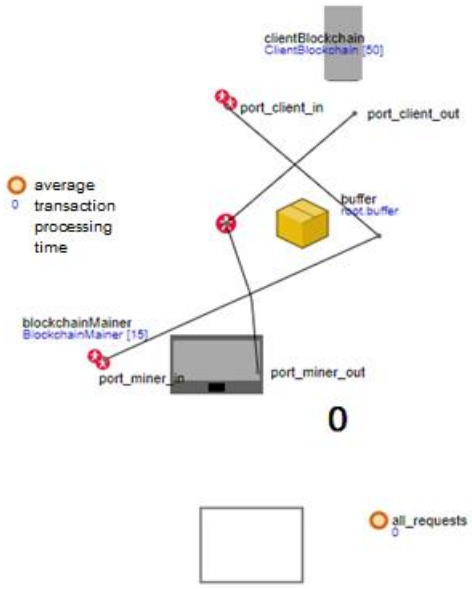

Figure 2: Simulation model of decentralized network.

The algorithm of centralized work is as follows:

1. Requests with a given intensity come from customers

2. Requests are queued on the application server, where they are processed and sent to the database server

3. After processing on the database server, the transactions again fall on the application server, where the result is sent back to the client

4. The client receives a response from the application server regarding the processing of its payment transaction

In the decentralized model, transaction processing has a different form:

1. Customers send transactions with a given intensity

2. Transactions fall into the buffer, where they are collected in blocks

3. When the block is filled with transactions, the miner begins the Mining block process

4. When the first of the miners completes the process, the block is closed and placed in the chain chain, and the transactions in this block are considered processed, so the responses are sent back to customers.

4. Results of modeling

Consider the Hinchin-Polachek formula for calculating the average waiting time of the application:

$$\omega = \frac{\lambda \times b^2 \times (1 + v^2)}{2 \times (1 - \lambda \times b)}$$

where λ - the intensity of the flow of applications,

b is the average processing time of one application,

v is the coefficient of variation of the law of distribution of the average processing time of one application.

If the denominator of the formula is greater than or equal to one, the average waiting time for the execution of one application goes to infinity. Indeed, if the intensity is too high, the application will never be processed at an infinite interval. The calculated values corresponding to the blockchain system considered in the work. The average processing time of one application.

$$b = \frac{the\ average\ mining\ time\ of\ the\ block}{the\ number\ of\ transactions\ in\ the\ block}$$

The average mining time of the block and the average number of transactions in the block were

obtained from the average indicators of the actually working Ethereum network in November 2017 [9, 10].

$$b = \frac{15}{77} \approx 0.195\ sec$$

The coefficient of variation for the exponential law, which determines the processing time of one application, is equal to one. Thus, we obtain the formula of the average waiting time for processing one application, which depends on the intensity of the input stream:

$$\omega = \frac{\lambda \times 0.038}{1 - \lambda \times 0.195}$$

For the centralized model:

b = average time of application processing on the application server + average time of application processing by the database server = 200ms + 103ms. = 0.303 sec.

Then the formula for the average waiting time for processing one application, which depends on the intensity of the input stream for the centralized network model:

$$\omega = \frac{\lambda \times 0.091}{1 - \lambda \times 0.303}$$

It is proposed to conduct several experiments, with different indicators of the intensity of the flow of requests and the number of customers.

Parameters of first experiment.

Number of clients: 5

Miner's number: 10

Number of transactions from the client per minute: 0.2

Number of requests: 10

First of all, you should calculate the intensity of the flow of applications per second:

$\lambda = 0.2 / 60 = 0.003\ sec$

The next step is to calculate the average waiting time for processing one application for a centralized system:

$\omega = (\lambda \cdot 0.091) / (1-\lambda \cdot 0.303) = 0.00027\ sec.$

And for centralized respectively:

$\omega = (\lambda \cdot 0.038) / (1-\lambda \cdot 0.195) = 0.00011\ sec.$

The experiment will run for 10 minutes. The centralized system processed requests in 3190,767 seconds, and the decentralized system in 66,880 seconds. A total of 50 requests were processed, as evidenced by the green colors of both rectangles.

Conduct experiment 2 with another data set:

Number of clients: 20

Miner's number: 15

Number of transactions from the client per minute: 1

Number of requests: 20

Let's calculate the values for modeling:

$\lambda = 0.2 / 600 = 0.016$ sec.

$\omega = (\lambda \cdot 0.091) / (1-\lambda \cdot 0.303) = 0.0014$ sec.

And for centralized respectively:

$\omega = (\lambda \cdot 0.038) / (1-\lambda \cdot 0.195) = 0.00060$ sec.

The experiment will run for 10 minutes. In the decentralized system, this experiment ends at 79.833 seconds of simulation, and the centralized system completed its work in 6673.53 seconds, processing only 124 applications.

Based on this, we can conclude that the processing of transactions in the decentralized network model is almost 47 times faster than in the centralized. At the same time, the centralized system has less fault tolerance than the decentralized one, as experiment 2 showed. In addition, the centralized system is vulnerable to DDoS attacks, while in the decentralized model, one of the nodes would have to take at least 51% of the load, which is completely unrealistic. That is why the confidentiality of data in a decentralized system is an order of magnitude higher than in a centralized one.

In order to clearly demonstrate the importance of the data, it was decided to conduct 23 experiments on different data sets and to track how each of the systems will behave as the number of queries increases. A constant number of clients was selected for the experiments - 5 pieces and the range of requests from 5 to 205. This means that each client will send 1,3,5,7 ... 41 requests. The results of these experiments are presented in Figure 3.

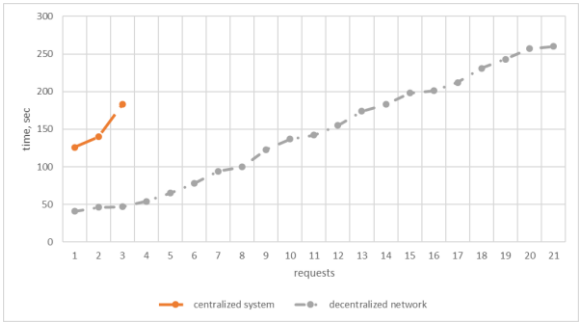

Figure 3: Graph of fault tolerance of systems

As can be seen from the figure, after 25 requests, the centralized system does not process the total number of requests coming into the

system. This means that the load of 5 requests from each of the 5 customers per minute for her was the maximum. The decentralized system processed all incoming requests.

The graph clearly shows that the curve of the centralized system breaks at the coordinate (183,132; 25). And the curve of the decentralized system is growing

5. Conclusions

The experiment showed that the performance of the network depends on the intensity of the appearance of applications, while for the correct operation of the blockchain technology of the presented type, it is possible to vary the values of the intensity of nodes and the values

buffer size.

The authors did not consider internal connections between network elements when building the models, which could affect the results. Also, the simulation model does not provide the possibility of obtaining a point estimate of the investigated parameter, but allows one to obtain interval estimates, the accuracy of which depends on the methods and scope of observations, the initial state, and the pseudo-random number generator.

It should be noted that modeling the performance of blockchain technology using the AnyLogic system can be convenient for analysis when changing various parameters. However, for more accurate results, it is necessary to carry out

additional research in the field of blockchain modeling on the AnyLogic emulator.

The analysis of the models showed the applicability of separate simulation systems for assessing the impact of blockchain technology on data transmission and processing networks.

In this paper, an overview of solutions based on blockchain technologies in the field of higher education was carried out and presented, as well as simulation models with an emphasis on queuing systems were presented. The results of comparison of decentralized and centralized systems are presented.

In the future, it is planned to expand the system indicators to obtain more accurate results using the AnyLogic system and propose a methodology for calculating the network infrastructure, taking into account the characteristics of the traffic and the received data.

6. References

[1] Nakamoto, S., Bitcoin, A. 2008. Bitcoin: A peer-to-peer electronic cash system. Available at: https://bitcoin.org/bitcoin.pdf

[2] Tulchinsky, G. 2017. Digital Transformation of Education: Challenges for Higher School. Russian Journal of Philosophical Sciences, 6, pp.121–136.

[3] Antonova, D. A., Ospennikova, E. V., Spirin, E. V. 2019. TSifrovaya transformatsiya sistemy obrazovaniya. Proektirovanie resursov dlya sovremennoy tsifrovoy uchebnoy sredy kak odno iz ee osnovnykh napravleniy. Vestnik Permskogo gosudarstvennogo gumanitarno-pedagogicheskogo universiteta. Seriya: Informatsionnye kompyuternye tekhnologii v obrazovanii, 14. Available at: https://cyberleninka.ru/article/n/tsifrovaya-transformatsiya-sistemy-obrazovaniya-proektirovanie-resursov-dlya-sovremennoy-tsifrovoyuchebnoy-sredy-kak-odno-iz-ee

[4] Oleg Barabash, Andrii Musienko, Spartak Hohoniants, Oleksandr Laptiev, Oleg Salash, Yevgen Rudenko, Alla Klochko. Comprehensive Methods of Evaluation of Efficiency of Distance Learning System Functioning. International Journal of Computer Network and Information Security(IJCNIS), Vol. 13, No. 1, Feb. 2021. pp 16–28.

[5] Shmatko, O., Borova, T., Yevseiev, S., & Milov, O. Tokenization of Educational Assets Based on Blockchain Technologies. ScienceRise: Pedagogical Education,(3 (42), 2021 pp.4–10..doi: 10.15587/2519-4984.2021. 232321.

[6] Fedorova, Elena P., and Ella I. Skobleva. "Application of Blockchain Technology in Higher Education." European Journal of Contemporary Education 9.3 2020 pp. 552-571.

[7] Caldarelli, Giulio, and Joshua Ellul. "Trusted Academic Transcripts on the Blockchain: A Systematic Literature Review." Applied Sciences 11.4 .2021. pp.18-42.

[8] Kiffer, Lucianna, Dave Levin, and Alan Mislove. "Stick a fork in it: Analyzing the Ethereum network partition." Proceedings of the 16th ACM Workshop on Hot Topics in Networks. 2017.

[9] Gencer, Adem Efe, et al. "Decentralization in bitcoin and ethereum networks." International Conference on Financial Cryptography and Data Security. Springer, Berlin, Heidelberg, 2018.