A Sketch of Blockchain Technology Simulation Model

Gasanov I.I., Ereshko F.I.
Dorodnicyn Computing Centre, FRC CSC RAS, Vavilova 40, 119333, Moscow, Russia
fereshko@yandex.ru

Abstract. The paper presents a sketch of the model for the system on blockchain principles. The Bitcoin project is adopted as a standard for illustrations and applications. Information propagation in the network is simulated.

Keywords: blockchain technology,information propagation, decentralization

Introduction

The blockchain technology is a specialized information-communication technology with some specific features (which is equivalently defined as techniques, means and methods of using computer equipment to perform the functions of collecting, storing, processing, transmitting and using data).
Two features distinguish the blockchain technology among the technologies for maintaining distributed databases: cryptographic data protection and decentralized procedure for ensuring coordination of the interests of all participants, i.e. for reaching a given consensus.
A wide interest in this technology has arisen recently due to the rush demand for cryptocurrencies and interest in the Projects in this area. Let us distinguish among them the Bitcoin Project and the Ethereum Project[1, 2].
The Bitcoin project originally created as a payment system now provides an increase in the wealth of the Project participants by producing a digital product (= bitcoin). The possibility of exchanging bitcoins for fiat money has become an additional incentive to attract new participants to the Project. At the same time, the volatility of this instrument is large.
The Bitcoin Project participants form an informal dynamic Coalition (its composition changes from time to time). They enter the Coalition by accepting the membership conditions and rules of interaction with other participants of the Bitcoin community. Some participants use the system only to perform their payment transactions. Another part of the community, so-called miners, takes part in maintaining the system registry (blockchain itself). In addition to the permanent verification of the registry correctness, they carry out mining, namely they participate in the competition to solve some scalar inequality and, if successful, increase their capital, i.e. the number of conventional coins in their virtual wallets.
Significant interest in cryptocurrencies has introduced the blockchain technology into the arsenal of scientific methodologies and applied developments as one of the effective technologies to support coalition solutions.

ICID-2019 Conference

Based on the ideology of system analysis, decision theory, operations research and game theory, the task to analyze a family of the procedures forming the blockchain technology is set.

Comment. In the original paper by S. Nakamoto and in subsequent comments in the media, the data protection process based on cryptographic algorithms and the mechanism for reaching a consensus among the participants with respect to the registry correctness took the prevailing place. This undoubtedly original design which has a fundamental character was used as the main argument in justification of the protection of private money and supported the interest in cryptocurrencies. On the back of the general interest, the Bitcoin Project gave rise to related solutions which together formed a new technological trend - the Blockchain technology.

1 Basic objective

As part of a systematic approach to study phenomena in nature and society and when designing application systems, the following steps are assumed: meaningful study and understanding of a particular object, construction of models, definition of information databases, implementation of software, elaboration of scenarios and performance of computational experiments in a simulation mode.

Thus, the primary ones are: problem statement, Project description and construction of the corresponding model.

This circumstance is fully understood by the developers of current digital platforms where there are the concepts of Project and Model in each description of the digital Platform.

For the problems associated with the use of the blockchain technology, the initial statement is undoubtedly a community of active agents each of which has its own goals, resources and awareness.

The community can arise spontaneously, but it can be created by some individual or legal entity - the Designer or the Center. Community organizers work out a common goal for which a coalition is being created and mechanisms for achieving it. Coalition management procedures may be centralized or decentralized.

The general agreement of the coalition includes various target conditions related to functioning of the community within the coalition, for example, issues of general awareness, distribution of the common good achieved, particular or general behavior strategies, etc.

The Bitcoin was created by a certain group of individuals on the basis of the papers by Satoshi Nakamoto [1, 2] who described the Project of creating an analogue of money as an exchange tool for a certain community of persons and created the primary software. The declared goal of the project was to build a sustainable secure payment system operating on the World Wide Web. The system was supposed to ensure the speed, cheapness and confidentiality of the participants' operations without engaging in the validation of the transactions of a certain single Center but solely due to the coalition members’ regulated reactions. Thus, the decentralized system was announced.

The next workable example of creating a coalition with using the blockchain technology to support functioning of the coalition is the Ethereum Platform. The targeted

aspiration of the community coalition in the Ethereum Project was to computerize the process of concluding contracts between the participants being active economic agents.

The final paragraph of the review [4] [https://en.wikipedia.org/wiki/Decentralization] which is directly related to the use of the blockchain technologies in the decision theory indicates that decentralized applications and decentralized organizations based on the blockchain can create competition for government organizations in the performance of managerial and regulatory functions [3,5].

The objective of the authors of the paper is that modeling of the blockchain system would allow mathematical methods to prove various statements about it that are now being made informally. Many effects can be investigated by simulating the work of the blockchain on computer implementation of the model. In addition, the models provide a convenient tool for creating and testing one's own inventions that modify elements of the structure under study.

It seems that it would be now difficult unequivocally to determine what a blockchain is, what else is a blockchain, and what is not a blockchain anymore. It seemed that the term itself appeared as a designation of the structure in which the bitcoin cryptocurrency functioned and then various modifications, currencies and structures began to appear. It is likely that the blockchain should be called exactly what is accepted by the society as a blockchain. However, the term may undergo changes in the process of evolution.

2 Description of the model sketch

In the specific version that can be found in the description of the Bitcoin and Ethereum Projects, the blockchain technology includes a sequence of actions to achieve a consensus, i.e. to achieve the general agreement of the participants on the correctness of the operations included in the blockchain database (in the chain of blocks). This agreement is stipulated by the conditions for the participants to join the coalition. It is ensured by the confirmation by all miners of the correctness of the next following block of data with transactions and by the solution by one of the miners of the basic inequality whose parameters depend on the contents of the block and which serves as an evidence of its work.

The blockchain is a technology on a peer-to-peer network designed to enter into contracts between the participants. The specificity of the contracts depends on implementation of the blockchain. Transfer of cryptocurrencies between the participants can also be interpreted as implementation of a contract for property transfer.

The network functions decentralized. In particular, this means that any contracts between the participants are concluded autonomously without any intermediaries.

All contracts are recorded in a single registry having the form of a chain of information blocks. The registry has many copies and is kept and maintained by the participants. The registry is public in the sense that any participant has access to it. However, the direct content of the contracts can be closed with a cipher from all the participants except the counterparties of the contract itself. Such multiplicity of the registry copies is one of the factors that protect the structure.

Information about the contracts concluded are distributed over the network as a chain reaction.
It is likely that the key property of the structure lies in the fact that it functions automatically insofar as its evolution is provided by the autonomous actions of the participants. These actions are performed by the participants voluntarily according to certain rules. (This is very similar to the coordinated behavior of the flocks of birds.)
We omit the description of some optional properties. In fact, the blockchains operate on the basis of computer networks. But for a blockchain as some abstract structure, this property is unessential. Although it would be inefficient, you can imagine a blockchain supported by courier service.
It seems quite promising to use for storage of the registries some other media than computer hard drives. Even now, private keys to digital wallets are recommended to be stored in paper form in view of the danger of hacking.
The information in the contracts is encrypted by using modern cryptographic methods or is hashed. It will be probably so in the future. But it appears that this fact itself and the choice of encryption methods are not integral properties of the blockchain.
It is possible that the blocking effect of the blockchain structure itself may disappear in some versions. Blocks are needed to record the general consensus of the participants in respect of new pieces of information. Why do not other forms of harmonization arise? It is unclear whether or not the structure would be still called a blockchain in this case. Now, any agreement or consensus is achieved through time-consuming unproductive work on each new block simultaneously done by many miners. We would like to think that other more natural ways of reaching a consensus will be invented. It is likely that the solution is on the path of introducing some elements of centralization into the structure (with a great many of large nodes) and modeling could exactly help to move forward in this direction. Would it be possible to call such a structure a blockchain? So far this is just a question, but why not?
But anyway, we found it right to start modeling with the Bitcoin as the most popular object underlying the phenomenon as a whole. Then, it is possible to go to the various existing and new modifications.

3 Project

When studying a complex object, it is not always advisable to build a single model that covers all its aspects. It is possible to consider some different models aimed at the study and optimization of various aspects of its functioning. Thus, you need to know the encryption systems used in the Bitcoin, but it is not necessary to include them in the model studying the stability of the structure. In this case, it is enough to make some necessary assumptions about the degree of reliability of such ciphers. The same applies to hashing. Hashing can be the subject of a separate study and (if necessary) modeling, but this is another separate topic.
In this model description, an attempt is made to sketch out only those aspects of functioning of the blockchain which are related to dynamics of the participants' progress towards a consensus, threats of appearance of so-called forks, threats of malicious attacks on the system. Forks mean the situation when two (or more) correct versions of the registry (blockchain) are simultaneously and immediately distributed in the

network and, therefore, there is no consensus. Thus, the model is not a real blockchain but a certain simplified system designed in the spirit of the Bitcoin.
The project of computer implementation of the model that simulates the operation of the system is described. The model assumes a variable number of the participants. Their quantities and equipment capacities vary according to some stochastic law from step to step of the model. Counterparties of operations, volumes of operations, moments of creation of new blocks by miners and successful miners are also randomly determined.
The description is made in a theoretic game (subject-oriented) form [6] in which the elements forming the model perform a separate independent work.
In the final form, the model is implemented as a program that sequentially processes the time steps of the model and autonomously modifies the status of each of the model elements (network nodes) [7].

Discrete time
We build a model with discrete time, i.e. step-by-step model. The steps (cycles) of the model is denoted by *t*. The model step is a rather small quantity. It is not determined yet what is worth staying at.

Set of objects
We consider a set of objects that will be called the nodes of the Network.
We assume that all the nodes being modeled combine mining and communication functions. This assumption is related to the dimension of the model. Simple nodes without the mining function are implicitly represented in the model as a random flow of transactions arriving at mining nodes.
Let us denote the number of simulated network nodes by N_t. This value is variable. n new nodes can be created at each model step with a certain probability $\alpha^+(n)$ specified in the model. And any of the existing nodes can be canceled with probability α^-. The appearance of a node means that it immediately begins to function by working with the current state of the blockchain. Since different nodes can from time to time support different versions of the blockchain, let us agree that any new node accepts the version of the correspondent node with the lowest number.
By varying the values α^+ and α^-, it is possible to choose the approximate maximum size of the network that is optimal for the model. In practice, already existing nodes may be temporarily disconnected and re-connected again. There is no point in complicating the model with this effect.

Node capacity
Each node n is assigned the conditional quantity P_n characterizing its computing capacity. This value is selected in a probabilistic way at the time of creation of the node. We believe that it does not change during the whole life of the node (which, of course, is very conditional).

Arcs (communication channels)
Network nodes are connected to each other by bidirectional arcs. Arcs simulate communication channels through which information is transmitted. We assume that in each step t, each node must have no less than $s^{\min}$ communication channels. Channels are selected randomly when creating a node and saved. In the step t, the node n may have less than $s^{\min}$ arcs. This is possible due to the termination of the correspondent node. In this case, a new correspondent is randomly selected for the node n. A new arc begins to function with a lag of time t^s.
Comment. In practice, not only miners but also simple nodes participate in the transfer of information over the Bitcoin network. The participants at their sole discretion take care of the existence of connections by making appropriate inquiries. In the model, some simplification is taken for the reasons of dimension.

Participant's (node's) information
Let us describe what information the participant (node) n has at the time t. He is aware of the chain of blocks formed by this time. Each block is characterized by its number and unique identifier (name). It contains the numbered names of the previous blocks and a set of the names of the transactions included in it.
The block name simulates the real block hash. In practice, it is enough to store only a link to the previous block for the purpose of verifying by a participant of the compliance of a new block with its chain version. This is achieved due to the uniqueness of the hashes. So far in this project, the hash apparatus is not intended as a mandatory element of modeling which leads to some simplifications. To ensure that the block and the chain are checked for compliance, the support of such block with a list of the names of all the preceding blocks is entered into the model. The processing by the participant of the newly acquired block is modified accordingly.
Any real block contains a set of data necessary for convenience of the network and for the protection of information. In this model, all this has not been considered yet. It is assumed that the information is sufficiently protected by cryptography from falsification. The same applies to transactions that contain keys, entrances, exits and any accompanying information in practice. All this has not been included in the model yet. The model assumes that each transaction is distinguished by a unique code (name) and only.

Pool of unrecorded transactions
Each participant maintains a pool of unrecorded (not yet included in the blockchain, unconfirmed) transactions. This pool is made up of transactions received by the participant over the network from the time t' of the last introduction of a block by the participant into his blockchain version up to the current time t, as well as of transactions that were in the pool by the time t' and were not included in the last block.

Events in the step t
The following events occur in the step t:

1. With a given probability $\mu(k)$, the node n accepts k newly created transactions for processing. (These are transactions either created by the participant n himself or not represented in the model by anonymous holders of simple wallets that are supposed to be switched with the node n). The node includes these transactions in its pool of unconfirmed transactions. It also sends them via communication channels that are commuting with it.
2. The node accepts from its correspondents the transactions transmitted over the network that have been received earlier by other nodes. The transmission time of the transaction packet from node to node is considered constant and denoted by τ_1. The participant checks to see if these transactions are contained in his pool of unconfirmed transactions. Those that are not yet contained, he adds to the pool and also sends them to his correspondents unless they were received from them.
3. The node n in the step t can receive from the correspondent nodes a new regular block created at one of the network nodes. After receiving the new block, the participant checks its number. If it is less than or equal to the number m that is last in the participant's current chain then such a block is rejected. As to the blocks with large numbers, they are checked for correctness. If the block is correct and its number is equal to or greater $m+1$ and it did not arrive at the node n earlier in the steps $\le t$, then the block is sent over the network to the correspondents of the participant n. (In the step t, the same block may come several times from different correspondents.) The data of the new block is distributed over the network at the speed $\tau_2 \ge \tau_1$, i.e. the correspondent nodes will receive this block in step $t+\tau_2$. (Due to the fact that simple nodes also participate in the real Bitcoin network, it is impossible to strictly indicate the links between the miner nodes as has been done in the model. The speed of information propagation over the network is modeled in a simplified form.)
4. If the number of the received block is equal to $m+1$ and the block refers to the block m in the chain of the participant n as the previous one, then the participant adds this block to his chain.
5. If the number of the received block $m' > m+1$, then the participant makes a (reverse) request to the node that sent the block. This is the request to send the chain of blocks with numbers $m+1 \le m'' < m'$. It will be received by the addressee in the step $t+\tau_1$ and the answer will come in the step $t+2\cdot\tau_1$. If the block $m+1$ in the response refers to the block m in the chain of the participant n as the previous one, then the participant adds this and subsequent blocks to his chain.
6. If the newly received block with the number $m+1$ refers to a block as the previous one that is other than the block m in the chain of the node n, then the participant n compares the names of the blocks in his chain with their names in the received block and finds such earliest number $\tilde{m}$ in which there is a discrepancy. Then in the same step t, he makes a (reverse) request to the node that sent the block with the intent that it sends to him all the other blocks starting with the number $\tilde{m}$.

The participant n needs some time for such a data exchange and check of the sent chain. Let us estimate this time as $q\tau_2 + 1$ where q is the number of the requested. After that, the node n replaces the blocks of its chain having the corresponding numbers by the received blocks and proceeds to use a new instance of the blockchain.

Comment. The number of the arrived block may be less than expected one in the case of attempt to falsify the chain. But it can also be smaller simply because of the longer path through the network in comparison with the path for some block almost simultaneously created at another node. Then, the participant n who have received a more "close" block goes to waiting for a block with a number greater by one.

If the network structure does not change and the node n in the step t is not in the state of waiting for a response to the request from the correspondent node, then it seems that the situation when the newly arrived block has a number greater than expected $m+1$ is impossible. When new nodes appear, this is generally possible. The new node can begin to function by loading the chain up to the number $m+1$. At the same time, information about the block $m+2$ which it immediately transmits to its correspondents can come to this node very quickly. But the paths along which the information about the block $m+1$ passes to these correspondents may turn out to be longer than the new path along which the information about the block $m+2$ has passed. However, this whole situation seems very casuistic in contrast to delays associated with a reverse request.

Fork situation

The case of a reverse request (Clause 6) describes within the framework of the model a fork situation when a participant receives several correct versions of the registry from the network. The method of its resolution presented above is somewhat different from the real one, but conveys the essence of the algorithm: the participant chooses the longest one from several versions of the chain he knows.

Validation

In practice, validation i.e. checking the correctness of a block and blockchain as a whole includes many actions. Block hashes, headers, block sizes, transactions themselves are checked. In this project, it is not necessary to model all of this. We will assume that all new blocks are correct. Then, the validation is reduced to checking the numbers and names of blocks (Clauses 3 to 6).

Mining

After validating and adding a new block to the blockchain (or blocks that came in response to a reverse request), the node n compares the composition (set of transactions) of the new block and its pool of unconfirmed transactions, leaving in the latter only those that are not still included in the blockchain. After that, the participant immediately, i.e. in the same step t, proceeds to mining and creating his own new block including in it all the unrecorded transactions remaining in the pool. Within the framework of the model, this means that for the new block its name and the list of transactions included in it are generated and a random process is started that simulates

the probability of finding by the participant the next target inequality. This probability is determined by the computing capacity of the node P_n and does not depend on history. In the model, it is calculated by the formula

$$\Omega_t^n = \frac{1}{T} \cdot \frac{P_n}{\sum_{k=1}^{N_t} P_k},$$

where T is the average time between the appearance of new blocks expressed in steps of the model.

Limitation of the block size

In practice, the block size is limited. Therefore, not all transactions from the pool are included in the new block, but according to the priorities. However, it is assumed in the original version of the model that the block size is unlimited and, therefore, the miner includes all unrecorded transactions by the beginning of the mining process.

Own new block

With some probability, the participant n in the step t solves the target inequality and thus completes the creation of his own new block. The important thing is that this should happen before the participant receives next following block from the outside. Otherwise, he has to start the mining process again. Having formed his own block, he does the same internal actions and spreads this block over the network in the same way as described above for the case when the correct block comes from the outside.

Comment. In practice, the Blockchain system is trying to maintain the periodicity of the appearance of new blocks approximately equal to 10 minutes. This is done by changing the difficulty of the problem being solved, and the correction is made not in every step, but also with a certain periodicity. In this model, it is not necessary to accept this complication.

Conclusions

The main goal of the subsequent computational experiments is to search for the organizational conditions under which the stable functioning of the Bitcoin system is ensured under various types of internal and external disturbances.

It is assumed, in the simulation mode by using the Monte Carlo method, to investigate the frequency of appearance of forks and their depth depending on a set of basic parameters: the number of the participants, the power of the miners, the signal passing time in the network, etc. It is also supposed to introduce into the model different scenarios of attacks on the system and observe its reaction. Then, by inventing defense mechanisms against these attacks, you can enter them into the model and study their effectiveness in a simulation mode.

References

1. Satoshi Nakamoto (2009). Bitcoin: A Peer-to-Peer Electronic Cash System, satoshin@gmx.com, www.bitcoin.org
2. Antonopoulos, Andreas M. (2014). Mastering Bitcoin. UNLOCKING DIGITAL CRYPTOCURRENCIES, O'Reilly Media, Inc., – 272 p.
3. Narayanan A., Bonneau J., Felten E., Miller A., Goldfeder S.. Bitcoin and cryptocurrency technologies : a comprehensive introduction. Princeton : Princeton University Press, 2016.
4. https://en.wikipedia.org/wiki/Decentralization
5. Raval S. Decentralized Applications. Blockchain Technologies in Use. St. Petersburg, 2017. 240 p. («O'Reilly Bestsellers»).
6. Ereshko F.I. Hierarchical games theory applied to lawmaking in the digital society. Business in law // Computational nanotechnolgy, 2017, No. 2, pp. 52–58.
7. Ereshko A.F., Vakhranev A.V. The project of the model building of technology of the distributed registers (description, the formal records and codes). IEEE Xplore Digital Library. Eleventh International Conference Management of Large-Scale System Development (MLSD), Moscow, Russia, 2018.