Blockchains Beyond Bitcoin: Towards Optimal Level of Decentralization in Storing Financial Data

Thach Ngoc Nguyen, Olga Kosheleva, Vladik Kreinovich, and
Hoang Phuong Nguyen

Abstract In most current financial transactions, the record of each transaction is stored in three places: with the seller, with the buyer, and with the bank. This currently used scheme is not always reliable. It is therefore desirable to introduce duplication to increase the reliability of financial records. A known absolutely reliable scheme is blockchain – originally invented to deal with bitcoin transactions – in which the record of each financial transaction is stored at every single node of the network. The problem with this scheme is that, due to the enormous duplication level, if we extend this scheme to all financial transactions, it would require too much computation time. So, instead of sticking to the current scheme or switching to the blockchain-based full duplication, it is desirable to come up with the optimal duplication scheme. Such a scheme is provided in this paper.

1 Formulation of the Problem

How financial information is currently stored. At present, usually, the information about each financial transaction is stored in three places:

- with the buyer,
- with the seller, and
- with the bank.

Thach Ngoc Nguyen
Banking University of Ho Chi Minh City, 56 Hoang Dieu 2, Quan Thu Duc, Thu Duc
Ho Chi Minh City, Vietnam, e-mail: Thachnn@buh.edu.vn

Olga Kosheleva and Vladik Kreinovich
University of Texas at El Paso, 500 W. University, El Paso, TX 79968, USA
e-mail: olgak@utep.edu, vladik@utep.edu

Hoang Phuong Nguyen
Division Informatics, Math-Informatics Faculty, Thang Long University, Nghiem Xuan Yem Road
Hoang Mai District, Hanoi, Vietnam, e-mail: nhphuong2008@gmail.com

This arrangement is not always reliable. In many real-life financial transactions, a problem later appears, so it becomes necessary to recover the information about the sale. From this viewpoint, the current system of storing information is not fully reliable: if a buyer has a problem, and his/her computer crashes and deletes the original record, the only neutral source of information is then the bank – but the bank may have gone bankrupt since then.

It is therefore desirable to incorporate more duplication, so as to increase the reliability of storing financial records.

Blockchain as an absolutely reliable – but somewhat wasteful – scheme for storing financial data. The known reliable alternative to the usual scheme of storing financial data is the *blockchain* scheme, originally designed to keep track of bitcoin transactions; see, e.g., [1, 2, 3, 4, 5, 6, 7, 8, 10, 11, 12].

In this scheme, the record of each transaction is stored at every single node, i.e., at the location of every single participant. This extreme duplication makes blockchains a very reliable way of storing financial data. On the other hand, in this scheme, every time anyone performs a financial transaction, this information needs to be transmitted to all the nodes. This takes a lot of computation time, so, from this viewpoint, this scheme – while absolutely reliable – is very wasteful.

Formulation of the problem. What scheme should we select to store the financial data?

It would be nice to have our data stored in an absolutely reliable way. Thus, it may seem reasonable to use blockchain for all financial transactions, not just for ones involving bitcoins. The problem is that:

- Already for bitcoins – which at present participate in a very small percentage of financial transactions – the world-wide update corresponding to each transaction takes about 10 seconds.
- If we apply the same technique to all financial transactions, this delay would increase drastically – and the resulting hours of delay will make the system completely impractical.

So, instead of using no duplication at all (as in the traditional scheme) or using absolute duplication (as in bitcoin), it is desirable to find the *optimal* level of duplication for each financial transaction.

This level may be different for different transactions:

- When a customer buys a relatively cheap product, too much duplication probably does not make sense, since the risk is small but the need for additional storage would increase the cost.
- On the other hand, for an expensive purchase, we may want to spend a little more to decrease the risk – just like we buy insurance when we buy a house or a car.

Good news is that the blockchain scheme itself – with its encryptions etc. – does not depend on whether we store each transaction at every single node or only in some selected nodes. In this sense, the technology is there, no matter what level of duplication we choose. The only problem is to find the optimal duplication level.

What we do in this paper. In this paper, we show how to find the optimal level of duplication for each type of financial transaction.

2 What Is the Optimal Level of Decentralization in Financial Transactions: Towards Solving the Problem

Notations. Let us start with some notations.

- Let d denote the level of duplication of a given transaction, i.e., the number of copies of the original transaction record that will be independently stored.
- Let p be the probability that each copy can be lost.

  This probability can be estimated based on experience.

- Let c denote the total cost of storing one copy of the transaction record.
- Finally, let L be the expected financial loss that will happen if a problem emerges related to the original sale, and all the copies of the corresponding record have disappeared.

  This expected financial loss L can estimated by multiplying the cost of the transaction by the probability that the bought item will turn out to be faulty.

Comments.

- The cost c of storing a copy is about the same for all the transactions, whether they are small or large.
- On the other hand, the potential loss L depends on the size of the transaction – and on the corresponding risk.

Analysis of the problem. Since the cost of storing one copy of the financial transaction is c, the cost of storing d copies is equal to $d \cdot c$.

To this cost, we need to add the expected loss in the situation in which all copies of the transaction are accidentally deleted. For each copy, the probability that it will be accidentally deleted is p. The copies are assumed to be independent. Since we have d copies, the probability that all d of them will be accidentally deleted is therefore equal to the product of the d probabilities p corresponding to each copy, i.e., is equal to p^d.

So, we have the loss L with probability p^d – and, correspondingly, zero loss with the remaining probability. Thus, the expected loss from losing all the copies of the record is equal to the product $p^d \cdot L$.

Hence, once we have selected the number d of copies, the overall expected loss E is equal to the sum of the above two values, i.e., to

$$E = d \cdot c + p^d \cdot L. \tag{1}$$

We need to find the value d for which this overall loss is the smallest possible.

Let us find the optimal level of duplication, i.e., the optimal d. To find the optimal value d, we can differentiate the expression (1) with respect to d and equate the derivative to 0. As a result, we get the following equation:

$$\frac{dE}{dd} = c + \ln(p) \cdot p^d \cdot L = 0, \tag{2}$$

hence

$$p^d = \frac{c}{L \cdot |\ln(p)|}.$$

By taking logarithms of both sides of this formula, we get

$$d \cdot \ln(p) = \ln\left(\frac{c}{L \cdot |\ln(p)|}\right).$$

Since $p < 1$, the logarithm $\ln(p)$ is negative, so it is convenient to change the sign of both sides of this formula. By taking into account that for all possible a and b, we have $-\ln\left(\frac{a}{b}\right) = \ln\left(\frac{b}{a}\right)$, we conclude that

$$d \cdot |\ln(p)| = \ln\left(\frac{L \cdot |\ln(p)|}{c}\right),$$

thus

$$d = \frac{\ln\left(\frac{L \cdot |\ln(p)|}{c}\right)}{|\ln(p)|}. \tag{3}$$

When p and c are fixed, then we transform this expression into an equivalent form in which we explicitly describe the dependence of the optimal duplication level on the expected loss L:

$$d = \frac{1}{|\ln(p)|} \cdot \ln(L) + \frac{\ln|\ln(p)| - \ln(c)}{|\ln(p)|}. \tag{4}$$

Comments.

- As one can easily see, the larger the expected loss L, the more duplications we need. In general, as we see from the formula (4), the number of duplications is proportional to the logarithm of the expected loss.
- The value d computed by using the formulas (3) and (4) may be not an integer. However, as we can see from the formula (2), the derivative of the overall loss E is first decreasing then increasing. Thus, to find the optimal integer value d, it is sufficient to consider and compare two integers which are on the two sides of the value (3)-(4): namely,
  - its floor $\lfloor d \rfloor$ and
  - its ceiling $\lceil d \rceil$.

Out of these two values, we need to find the one for which the overall loss E attains the smallest possible value.

Acknowledgments

This work was supported in part by the US National Science Foundation via grant HRD-1242122 (Cyber-ShARE Center of Excellence).

The authors are thankful to Professor Hung T. Nguyen for valuable discussions.

References

1. A. M. Antonopoulos, *Mastering Bitcoin: Programming the Open Blockchain*, O'Reilly, Sebastopol, California, 2017.
2. J. J. Bambara, P. R. Allen, K. Iyer, S. Lederer, R. Madsen, and M. Wuehler, *Blockchain: A Practical Guide to Developing Business, Law, and Technology Solutions*, McGraw Hill Education, New York, 2018.
3. I. Bashir, *Mastering Blockchain*, Packt Publishing, Birmingham, UK, 2017.
4. M. Connor and M. Collins, *Blockchain: Ultimate Beginner's Guide to Blockchain Technology – Cryptocurrency, Smart Contracts, Distributed Ledger, Fintech and Decentralized Applications*, CreateSpace Independent Publishing Platform, 2018.
5. D. Drescher, *Blockchain Basics: A Non-Technical Introduction in 25 Steps*, Apress, New York, 2017.
6. M. Gates, *Blockchain: Ultimate guide to understanding blockchain, bitcoin, cryptocurrencies, smart contracts and the future of money*, CreateSpace Independent Publishing Platform, 2017.
7. T. Laurence, *Blockchain For Dummies*, John Wiley, Hoboken, New Jersey, 2017.
8. A. T. Norman, *Blockchain Technology Explained: The Ultimate Beginners Guide About Blockchain Wallet, Mining, Bitcoin, Ethereum, Litecoin, Zcash, Monero, Ripple, Dash, IOTA And Smart Contracts*, CreateSpace Independent Publishing Platform, 2017.
9. M. Swan, *Blockchain: Blueprint for a New Economy*, O'Reilly, Sebastopol, California, 2015.
10. D. Tapscott and A. Tapscott, *Blockchain Revolution: How the Technology Behind Bitcoin Is Changing Money, Business, and the World Hardcover*, Penguin Random House, New York, 2016.
11. P. Vigna and M. J. Casey, *The Truth Machine: The Blockchain and the Future of Everything*, St. Martin's Press, New York, 2018.
12. A. K. White, *Blockchain: Discover the Technology behind Smart Contracts, Wallets, Mining and Cryptocurrency (including Bitcoin, Ethereum, Ripple, Digibyte and Others)*, CreateSpace Independent Publishing Platform, 2018.