Advances in Artificial Intelligence Research (AAIR)
ISSN 2757-7422
www.dergipark.com/aair/

Vol. 4 (No. 2), pp. 117-123, 2024
doi: 10.54569/aair.1608702
Published online: Dec 28, 2024

ALTO-assisted Peer Selection in Bitcoin P2P Network

Cihat ÇETİNKAYA [1,*]

[1] Department of Software Engineering, Faculty of Engineering, Muğla Sıtkı Koçman University, Turkiye;

Abstract

Blockchain-based applications rely on a decentralized structure wherein the transactions are recorded on a public ledger that is maintained by every node in the peer-to-peer (P2P) network. The transactions and blocks are propagated using a multi-hop broadcast and verified by every node in the network. Application Layer Traffic Optimization (ALTO), on the other hand, is a network protocol developed and maintained by the Internet Engineering Task Force (IETF) to provide network related information to the P2P applications to increase their performance. In this study, a novel peer selection method based on the network information provided by ALTO protocol is proposed to decrease the block propagation delay of the Bitcoin P2P network. The simulations show that the proposed peer selection method can effectively decrease the block propagation time and fork rate compared to Bitcoin's random peer selection and region-based peer selection methods.

Keywords*: P2P network, Blockchain, Traffic Optimization, Peer Selection, Bitcoin.*

1. Introduction

Blockchain technology emerged in 2008 with the development of the Bitcoin [1] cryptocurrency by a group of researchers using the nickname Satoshi and has since attracted attention from both academia and industry. Due to its distributed architecture, blockchains are used in smart contracts, internet of things (IoT), non-fungible tokens (NFTs), healthcare, logistics, and personnel digital security, as well as cryptocurrency.

In P2P applications, which also form the basis of blockchain-based systems, one of the factors affecting performance is the peer selection process in which the nodes in the P2P network select the peers with whom they will exchange data [2]. In addition to the fact that peer selection is usually done randomly, in some P2P applications, criteria such as the geographical distance between nodes, the upload/download bandwidth of the nodes, and the chunks of data held by the nodes also guide the peer selection process. However, both peer selection processes do not consider *(i)* the topology of the network on which the applications run and *(ii)* the localization of network traffic. While peer selection without network topology information usually reduces the performance of the P2P application, peer selection without considering network traffic information causes a cost for the economies of Internet Service Providers (ISPs). Therefore, it is crucial to consider both network topology and network traffic information when peer selection is performed in P2P applications.

In the literature, studies that provide such network-related information to P2P applications are classified into two different groups [3]. In the first group, network-related information is estimated and provided to the nodes by running a distributed application at the application layer [4, 5]. In the second group of studies, network-related information is provided by ISPs that own the network [6-8]. When the performances of the studies in both groups are analyzed on P2P applications, it is seen that the ISS-based approaches in the second group provide a more effective peer selection for P2P applications [3]. Therefore, the IETF (Internet Engineering Task Force) group introduced the ALTO protocol [9] to provide network-related information to the peers that run on P2P applications. One of the objectives of the ALTO protocol is to design and define the ALTO service that provides the necessary network-related information to the nodes running on P2P applications to perform better-than-random peer selection.

In this study, an ALTO-assisted peer selection method is proposed for blockchain-based systems. In the proposed method, peer selection is based on a multi-objective optimization model and aims to select peers that will reduce the block propagation delay by using network information obtained from the ALTO server. In this paper, the peer selection method is implemented on the Bitcoin P2P network since it is both a public blockchain, consists of hundreds of thousands of nodes deployed in many autonomous systems around the world, and has the highest commercial value. However, the peer selection method proposed in this study can also be applied to other alternative public blockchains such as Litecoin and Dogecoin with little or almost no modification.

The rest of this paper is as follows. Section 2. provides background information about Bitcoin, ALTO protocol and summarizes related works. Section 3. presents the proposed peer selection method, while Section 4. describes the simulation study and presents the results and comparative analysis. The conclusion and future work discussed in Section 5.

*Corresponding author
cihat.cetinkaya@mu.edu.tr

Received: 27/Dec/2024; Accepted: 28/Dec/2024.

2. Background and Related Works

According to the Bitcoin protocol, a new node joining the network first performs peer discovery mechanism since it does not yet have any information about the other nodes in the network. In the first phase of peer discovery, the node obtains information about nodes in the network by sending queries to DNS seeds that are hardcoded in the Bitcoin reference software. Then, the node tries to establish a connection by sending a *version* message to the nodes in the node list randomly obtained from DNS seeds (**Figure 1**). If the remote node sends *verack* message, the node adds the remote node as outgoing peer and the remote node adds the new node as incoming peer. The nodes also exchange information about other nodes they discover on the network by sending *addr* and *getaddr* messages to each other (**Figure 2**). By default, Bitcoin implements a total of 125 peer connections, 117 of which are incoming connections and 8 are outgoing connections. In our previous work [10], we investigated the optimum number of outgoing connections through simulations and found out that the optimum number for outgoing connections is 8-10 which is almost same with the default value of Bitcoin.

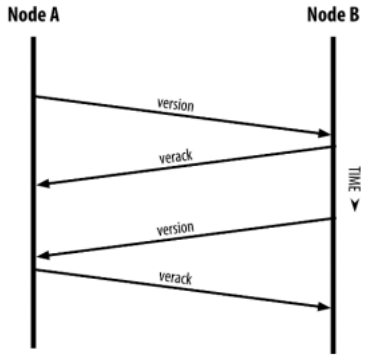

Figure 1. *Message timeline of connection establishment between nodes.*

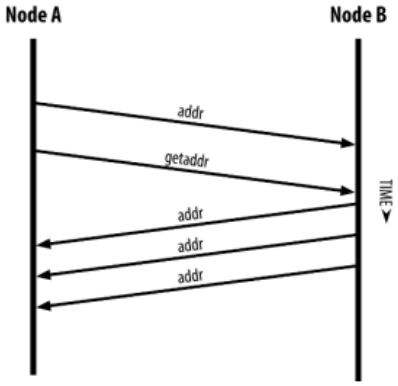

Figure 2. Message timeline of exchanging discovered nodes

The main purpose of the ALTO protocol, which was standardized by the IETF in 2014 as RFC 7285, is to define ALTO services that provide the network-related information to the applications running on the network so that the nodes in the applications can perform better than random peer selection. Upload and download bandwidth capacity of nodes, packet routing cost due to ISPs policy, topological hop count, end-to-end delay, traffic quota can be given as network-related information that is served by ALTO protocol. According to ALTO protocol, the ALTO server is responsible for delivering the ALTO service where ALTO client queries ALTO server with different ALTO queries. The multi-cost map service, which is one of the services defined in the ALTO protocol, enables multiple cost metric to be served by making a single query/response transaction. **Figure 3**. presents an example of multi-cost map that is received by an ALTO client. There are 2 different cost metrics in the map, *routingcost* and *hopcount*, both of which are numerical. According to the cost information given in the map, the *routingcost* between PID1 and PID2 is 5 while the *hopcount* is 23. In the proposed peer selection method, the node that joins the P2P network acts as an ALTO client and queries the ALTO server and receives multi-cost map for the candidate nodes in the network. The details are given in Section 3.

```
{
 "meta" : {
   "dependent-vtags" : [ ... ],
   "cost-type" : {},
   "multi-cost-types" : [
     {"cost-mode": "numerical", "cost-metric": "routingcost"},
     {"cost-mode": "numerical", "cost-metric": "hopcount"}
   ]
 }
 "cost-map" : {
   "PID1": { "PID1":[1,0],  "PID2":[5,23],  "PID3":[10,5] },
   ...
 }
}
```

Figure 3. *Example of a multi-cost map received by ALTO client after querying ALTO server* [11]

Several peer selection methods have been proposed for blockchain-based systems. In [12], the nodes in the network are clustered with respect to physical distance between nodes and the peer selection is performed within cluster. In [13-15] the closest nodes based on the geographical distance are considered in the peer selection process. In [16], the peers are selected based on the ping latencies between nodes. In [17, 18], similarly aimed to select peers with low delay according to the protocol messages that nodes received from peers. In [19], the authors propose a region-based peer which is based on regional information of nodes. Previous studies performed peer selection without having any information about the real network on which they run. The novelty of the proposed study is that it uses up-to-date fine-grained network information obtained from ALTO services.

.

3. Proposed Peer Selection Method

In the proposed peer selection method, when a new peer joins the Bitcoin P2P network and after getting the initial node list from the DNS seeds starts the peer discovery process. Differing from the default peer discovery process given in Section-II, in the proposed peer selection method the node does not send *verack* messages to the other peers in the network. Instead, the peer keeps discovering the nodes in the network by sending *getaddr* messages to the other nodes in the network. After completion of this step, the node has obtained information (e.g. IP address and protocol version) of several nodes in the network and keeps their information in a node list. In the second step, the node queries the ALTO server using ALTO multi-cost service with the IP addresses of the nodes in the node list and receives the end-to-end delay, upload and download bandwidth capacities as the cost variables of the nodes. The node applies a certain threshold value to each of the delay, upload and download bandwidth values of the nodes in the node list and creates a candidate peer list from nodes that do not exceed the threshold value for all three values.

In the third step, the node applies a multi-objective optimization model aiming to find a peer that minimizes the delay while maximizing the upload and bandwidth capacities. Let N represents the nodes in the candidate peer list. For each $n \in N$; dl_n, $up_n$and dw_n denotes the delay, upload and bandwidth values of the node n, respectively. A row vector for each node is constructed and given in Eq.(1) as follows:

$$\overrightarrow{N_{(n)}} = [dl_n, 1/up_n, 1/dw_n] \tag{1}$$

In Eq.(1), reciprocal values of upload and download bandwidth are used. Thus, the optimization model given in Eq.(3) aims to minimize all cost variables. Since each cost variable may have different effects on the performance of the peer, a weight vector of $\vec{w}$ is assigned to objective variables and weighted objective variables are calculated in Eq.(2) by inner product of $\vec{w}$ and $\overrightarrow{N_{(n)}}$.

$$\Psi_n = \langle \vec{w}, \overrightarrow{N_{(n)}} \rangle \tag{2}$$

Since the aim of the optimization model is to find a node that minimizes all objective variables, an utopia point v is defined that represents the optimal solution. The optimization model is given as follows:

$$\begin{aligned} &minimize \quad \|\Psi_n, v\| \\ &\text{n} \\ &subject\ to \quad n \in N \end{aligned} \tag{3}$$

The optimization model given in Eq.(3) is solved using an exhaustive search and returns a node n_i whose objective variables are closest to utopia point v among other nodes in $\overrightarrow{N_{(n)}}$. Then, n_i is removed from the candidate peer list and the node tries to establish an outgoing connection with n_i by sending a version message as discussed in Section 2. The third step of the peer selection process is repeated until the node successfully connects to at most 6 outgoing peers. The rest of the peers are randomly selected from the node list populated in the first step of the peer selection process. By selecting a subset of peers randomly makes the proposed peer

selection method resilient to eclipse attacks [20].

4. Simulation Study

First, the simulation environment based on Simblock Bitcoin P2P simulator [21] was set up. Then the proposed work was tested with different weights given in Section 3. Last, the performance of the proposed work was compared with Bitcoin's default peer selection method and region-based peer selection like method presented in [19] from the literature.

4.1. Simulation Setup

Simblock is a discrete-event simulator that can simulate Bitcoin P2P network with the exact same parameters that Bitcoin P2P network had in 2015 and 2019. Since its release, Simblock has been extensively used [22] by Blockchain researchers and developers. Since the default parameters of Simblock are out of date and do not reflect the current characteristics of the Bitcoin P2P network, the up-to-date parameters were gathered for Bitcoin P2P network and these parameters were passed to Simblock.

Geographical distribution of the nodes used in the simulations are given in **Table 1**. The values presented in **Table 1** were obtained by averaging the unique nodes discovered in Bitcoin P2P network using Bitnodes API [23] from July 1, 2024, to July 31, 2024. In the simulations, the number of nodes varied as 500, 1000, 2000 and 4000, and the nodes were randomly located in the network regions according to the rates given in **Table 1**. The download/upload bandwidths of the nodes were calculated using country-based values obtained from the testmy.net [24] website. The latencies between the network regions were retrieved from Verizon [25].

The other Bitcoin-related parameters used in the simulations are presented in **Table 2**. The end block height parameter, which indicates the number of blocks to be mined during the simulations, is calculated as the total number of blocks mined in the Bitcoin network between July 1, 2024, and July 31, 2024, and the average block size parameter is calculated as the arithmetic average of the total sizes of the blocks mined in the same period.

Table 1. *Geographical distribution of nodes*

Region	Rate
North America	18.9 %
Europe	59.7 %
South America	4.3 %
Asia Pacific	13.7 %
Japan	1.6 %
Australia	1.8 %

Table 2. *Simblock parameters used in simulations*

Parameter	**Value**
# of nodes	[500, 1000, 2000, 4000]
Average block size	1.69 MB
Compact block size	13 KB
End block height	4713
Block generation interval	10 mins

4.2. Performance Evaluation

After setting up the simulation environment, the proposed peer selection method was tested using different weights $\vec{w}$ given in the Section 3. Each test was conducted 30 times and average block propagation delay and fork rate values were reported. It was found that the best result obtained for the proposed selection method was when the weight vector $\vec{w}$ was assigned as $\langle 0.5,0.25,0.25 \rangle$ where the elements of the vector denote delay, upload bandwidth and download bandwidth, respectively.

To evaluate the performance of the proposed peer selection method, the results were compared with Bitcoin's default peer selection method and region-based like method. In Bitcoin's default peer selection method, the peers were selected randomly. In region-based like method, 6 out of 8 peers were selected within the same region as node's where remaining 2 peers were selected randomly from outside of the node's region.

Both peer selection methods were tested 30 times with the same simulation parameters used to test the proposed method and the average results were reported.

Figure 4, **Figure 5** and **Figure 6** present the average block propagation delay, block propagation delay to reach 50% of nodes and, block propagation delay to reach 90% of nodes, respectively. It can be seen that the proposed peer selection method outperforms the default peer selection of Bitcoin and region-based like peer selection by reducing the block propagation time in general. This indicates that using fine-grained network-related data in peer selection process plays an important role in Bitcoin P2P network.

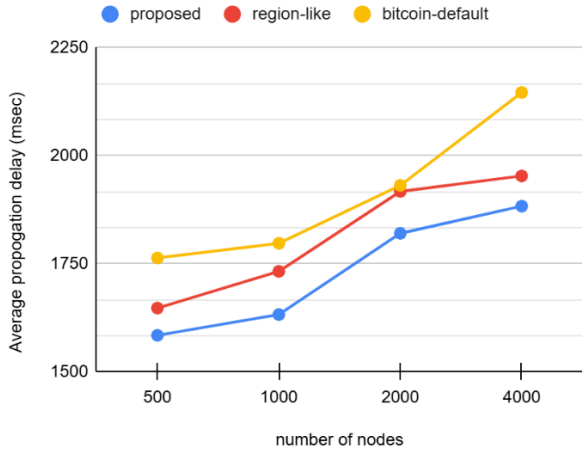

Figure 4. *Average block propagation delay*

Table 3 presents the block propagation delay where the block reaches every node in the network. It is observed from **Table 3**. that the time required for the block to reach all nodes in the network is the lowest in the proposed method. However, it is seen that the values obtained are close to each other in all peer selection methods. This is because there are nodes in the network with low bandwidth capacity and high end-to-end delay.

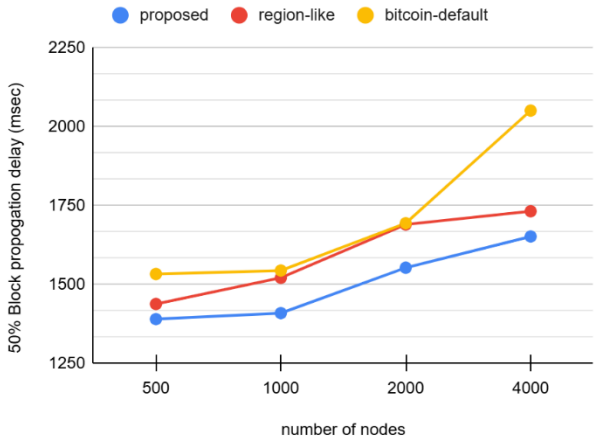

Figure 5. *Block propagation delay to reach 50% of nodes*

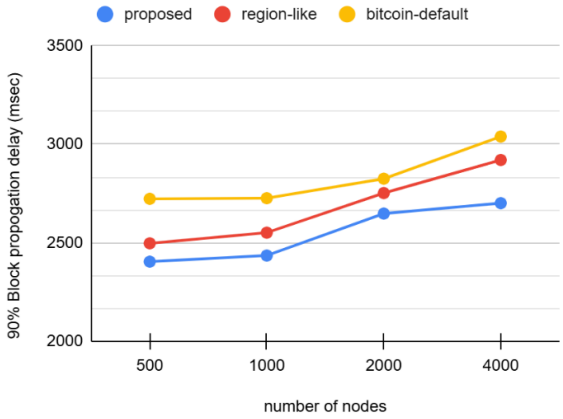

Figure 6. *Block propagation delay to reach 90% of nodes*

Table 3. *Block propagation delays reaching all nodes in the network. (msecs)*

	Number of nodes			
Method	500	1000	2000	4000
proposed	7873	8232	9951	10808
region-like	8118	8559	10182	10980
bitcoin-default	8284	8749	10331	11222

The average fork number of the peer selection methods are given in **Table 4**. As in block propagation delay, the proposed method has a lower fork number than the other peer selection methods for all the tests that were conducted with different number of nodes. In addition, as the number of nodes in the network increases, the chances of nodes finding good peers also increases, so the number of forks decreases in all methods. All the results obtained with the simulations show that the peer selection made by obtaining detailed information about the network thanks to the ALTO protocol plays an important role in reducing the block propagation delay and thus increasing the security of the blockchain by minimizing the possibility of forks.

Table 4. *Average number of forks in the blockchain.*

	Number of nodes			
Method	500	1000	2000	4000
proposed	13.9	13.6	13.5	12.8
region-like	14.9	14.1	14.0	13.3
bitcoin-default	16.6	14.9	14.1	13.8

5. Conclusion

In this paper, a novel peer selection method for Bitcoin P2P network was proposed under the guidance of ALTO protocol. The proposed method aims to select the optimum peers that minimize the block propagation delay by considering the upload/download bandwidth capacity of the candidate nodes and the network delay between the node and the candidate nodes, which are the network-related information obtained by the node from the ALTO server. The simulation results show that ALTO-assisted peer selection outperforms default random peer selection of Bitcoin and region-based peer selection. It is also shown that both the upload and download bandwidth capacities of nodes affect the block propagation delay. As a future work, we plan to propose a reinforcement learning-based peer selection method.

Declaration of interest

The authors declare that there is no conflict of interest.

Acknowledgements

This work is funded by the Scientific and Technological Research Council of Turkey (TUBITAK) Electric, Electronic and Informatics Research Group (EEEAG) under grant 121E401.

References

[1] Nakamoto, S. “Bitcoin: A Peer-to-Peer Electronic Cash System”, 2008, [Online] Available: https://bitcoin.org/bitcoin.pdf (accessed: December 01, 2024).

[2] Shen X, Yu H, Buford J, Akon M. “Handbook of Peer-to-Peer Networking”, New York, Springer, 2010.
[3] Gurbani VK, Hilt V, Rimac I, Tomsu M, Marocco E. "A survey of research on the application-layer traffic optimization problem and the need for layer cooperation”, IEEE Communications Magazine, 47, 107-112, 2009.
[4] Costa M, Castro M, Rowstron A, Key P. “PIC: Practical Internet coordinates for distance estimation”, in Proceedings of International Conference on Distributed Systems, 2003.
[5] Dabek F, Cox R, Kaashoek F, Morris R. “ Vivaldi: A Decentralized Network Coordinate System”, in Proceedings of ACM SIGCOMM, 2003, 15-26.
[6] Saucez D, Donnet B, Bonaventure O. “Implementation and Preliminary Evaluation of an ISP-Driven Informed Path Selection”, in Proceedings of. ACM CoNEXT, 2007,1-2.
[7] Aggarwal V, Feldmann A, Scheideler C. “Can ISPs and P2P systems co-operate for improved performance?”, ACM SIGCOMM Computer Communications Review (CCR), 37(3), 29-40, 2007.
[8] Xie H, Yang YR, Krishnamurthy A, Liu Y, Silberschatz A. “P4P: Provider Portal for (P2P) Applications”, in Proceedings of ACM SIGCOMM, 2008, 351-362.
[9] Alimi R, Penno R, Yang Y, Kiesel S, Previdi S, Roome W, Shalunov S, Woundy R. “Application-Layer Traffic Optimization (ALTO) Protocol”, 2014, [Online], Available: https://datatracker.ietf.org/doc/rfc7285/ (accessed: December 01, 2024).
[10] Cetinkaya C. “A Study on the Impact of Connection Number Parameter of Nodes on the Performance of Bitcoin Peer-to-Peer Network”, 5th International Conference on Data Science and Applications, 2022, 131-134.
[11] Randriamasy S, Wendy R, Schwan N. “Multi-Cost Application-Layer Traffic Optimization (ALTO)”, 2017, [Online], Available: https://datatracker.ietf.org/doc/rfc8189/ (accessed: December 01, 2024).
[12] Fadhil M, Owenson G, Adda M.“A Bitcoin Model for Evaluation of Clustering to Improve Propagation Delay in Bitcoin Network”, in Proceedings of IEEE Intl Conference on Computational Science and Engineering, 2016.
[13] Fadhil M, Owenson G, Adda M. “Locality based approach to improve propagation delay on the Bitcoin peer-to-peer network”, in Proceedings of the IFIP/IEEE International Symposium on Integrated Network and Service Management, 2017, 556-559.
[14] Park S, Im S, Seol Y, Paek J. “Nodes in the Bitcoin Network: Comparative Measurement Study and Survey”, IEEE Access, 7, 57009-57022, 2019.
[15] Sudhan A, Nene M. “Peer Selection Techniques for Enhanced Transaction Propagation in Bitcoin Peer-to-Peer Network”, in Proceedings of the 2nd International Conference on Intelligent Computing and Control Systems, 2019, 679-684.
[16] Sallal M, Owenson G, Adda M. “Proximity Awareness Approach to Enhance Propagation Delay on the Bitcoin Peer-to-Peer Network”, in Proceedings of the International Conference on Distributed Computing Systems, 2017, 2411-2416.
[17] Wang K, Kim H. “FastChain: Scaling blockchain system with informed neighbor selection”, in Proceedings of the 2nd EEE International Conference on Blockchain, 2019, 376-383.
[18] Aoki Y, Shudo K. “Proximity neighbor selection in blockchain networks”, in Proceedings of the 2nd IEEE International Conference on Blockchain, 2019, 52-58.
[19] Matsuura H, Goto Y, Sao H. “Region-based Neighbor Selection in Blockchain Networks”, in Proceeding of the IEEE International Conference on Blockchain, 2021, 21-28.
[20] Heilman E, Kendler A, Zohar A, Goldberg S. “Eclipse attacks on Bitcoin’s peer-to-peer network”, USENIX Security Symposium, 2015, 129–144.
[21] Aoki Y, Otsuki K, Kaneko T, Banno R, Shudo K. “Simblock: A Blockchain Network Simulator”, in Proceedings of IEEE Conference on Computer Communications Workshops, 2019, 325-329.
[22] Shudo K, Hasegawa T, Sakurai A, Banno R. "Blockchain Network Studies Enabled by SimBlock," 2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Dubai, United Arab Emirates, 2023, pp. 1-2.
[23] Global Bitcoin nodes distribution, [Online] Available: https://bitnodes.io/api/ (accessed: December 01, 2024).
[24] Internet Speed Test, [Online], Available: https://testmy.net (accessed: December 01, 2024).
[25] Verizon Network Performance, [Online] Available: https://verizon.com (accessed: December 01, 2024).