A POMDP-based Method for Analyzing Blockchain System Security Against Long Delay Attack*

(Extended Abstract)

Shuangfeng Zhang, Yuan Liu, Xingren Chen
Northeastern University
Shenyang, China
liuyuan@swc.neu.edu.cn

Xin Zhou
Nanyang Technological University
Singapore
xin.d.zhou@outlook.com

ABSTRACT

Blockchain bears the long-delay attack which is challenging to be analyzed. In this study, we propose a blockchain security analysis model based on Partially Observable Markov Decision Process (POMDP) against long delay attack by capturing the dynamic network delay. In our model, an observation function about the network delay is learned and updated based on a clustering algorithm about the timely network status. With the support of the observation function, a POMDP model is constructed for attackers to maximize their expected rewards. To analyze the security of a blockchain system against long delay attack, the utility of the attackers and normal miners with the same mining power are calculated and compared. The system is then regarded secured as the utility of the normal miners is no less than that of the attackers.

CCS CONCEPTS

• **Security and privacy** → *Formal security models*;

ACM Reference Format:
Shuangfeng Zhang, Yuan Liu, Xingren Chen and Xin Zhou. 2020. A POMDP-based Method for Analyzing Blockchain System Security Against Long Delay Attack. In *Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), Auckland, New Zealand, May 9–13, 2020,* IFAAMAS, 3 pages.

1 INTRODUCTION

A blockchain system is a distributed and asynchronized ledger to securely record data in an incremental manner [3]. With the exciting characteristics of blockchain technology, such as decentralization, nonrepudation, traceability, and transparency, it has attracted much attention in various fields, including finance [6], healthcare [2], e-governance [1] and so on. Due to the information transmission mechanism of the peer-to-peer network, the security of a consensus protocol is not only constrained by attackers' computing power, but also significantly impacted by the network latency [4, 7].JMM', L.] []The long delay attack was firstly formally proposed in a generalized asynchronous environment, where the attackers delaying the message transmission between honest miners to obtain a relatively longer mining time, thereby achieving a higher probability in successfully mining the next block. Wei al. further enriched the delay attack [5], by extending the network delay from a fixed boundary to a considerable extent. However, these studies do not consider the impact of dynamic network latency changes in analyzing blockchain system security against the long delay attack.

In this study, we aims to provide a new POMDP-based security analysis method for blockchain systems against the long delay attack. We construct a POMDP-based model for attackers in making their attacking decisions. The network latency is obtained through observing the time difference between information being transmitted and received, and a belief function is then built and updated based on a clustering method. With the support of the POMDP model, the attackers are able to adjust their attacking strategy to achieve their maximal expected rewards. The system is regarded as safe against the long delay attack if their rewards is no more than the honest miners, and vice versa.

2 THE PROPOSED ANALYSIS MODEL

Our analysis model consists of two main modules as shown in Figure 1. The first module is the POMDP model for the blockchain system, which is used for the attackers in deciding their attacking strategies. In the first module, six main components are specified, including a clustering method to extract the observation function about the network delay in real time. The second module is the security analysis of blockchain against the long delay attack by comparing the rewards of the attackers and honest miners. The evaluated system is regarded as safe against the long delay attack if the rewards of the attackers is no more than that of honest miners.

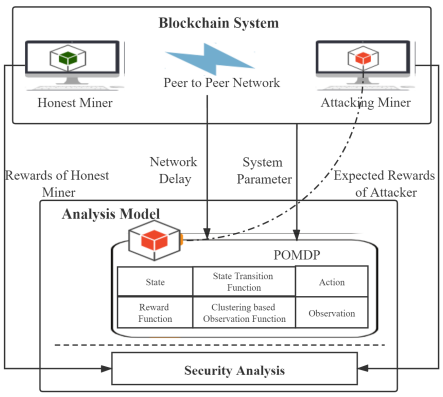

Figure 1: The Proposed Analysis Model

*Equal contribution by the first two authors. This work is supported in part by the National Natural Science Foundation for Young Scientists of China under Grant No.61702090 and No. 61702084.

Proc. of the 19th International Conference on Autonomous Agents and Multiagent Systems (AAMAS 2020), B. An, N. Yorke-Smith, A. El Fallah Seghrouchni, G. Sukthankar (eds.), May 9–13, 2020, Auckland, New Zealand.

2.1 The Proposed POMDP Model

The main objective of our POMDP model is assist an long delay attacker to strategically decide whether to launch an attack in a dynamic network condition.

2.1.1 Main Components. We specify the six main components in the POMDP model.

• **State.** The state in the POMDP is denoted by $S =< t, l >$ consist -ing of two parts, where $t \in$ {high,low} is the network delay state and $l \in \mathbb{N}$ is the relative block length.

• **Action**. Our model allows the attacker to perform three types of actions, denoted by $A = \{q, m, w\}$ including *query, malicious attack and wait* actions.

• **State Transition Function**. The general expression of the state transition function is described as $Pr(s' =< t', l' > | s =< t, l >, a \in A)$. For a query action, it will not cause a switch between different states. Thus,

$$Pr(< t', l > | < t, l >, q) = \delta(t', t) \tag{1}$$

where $\delta(t', t)$ is the Kronecker delta with value being 1 when $t' = t$, and 0 otherwise. For an attack action, the state transaction function can be expressed as follow.

$$Pr(< high, l' > | < high, l >, a) = \begin{cases} H & l' = l + 1 \\ 1 - H & l' = l - 1 \\ 0 & others \end{cases} \tag{2}$$

$$Pr(< low, l' > | < low, l >, a) = \begin{cases} L & l' = l + 1 \\ 1 - L & l' = l - 1 \\ 0 & others \end{cases} \tag{3}$$

For a wait action, the effect is similar to that of a query action.

$$Pr(< t'l' > | < t, l >, W) = \begin{cases} 1 - \alpha & l' = l \\ \alpha & l' = l - 1 \\ 0 & others \end{cases} \tag{4}$$

• **Reward function**. The reward function assigns a certain num -ber of rewards or punishment to the attackers in performing an action, denoted by $R(s', a)$. When $a = q$, the system should charge a certain relatively small cost in executing the pre-processes, resulting that $R(s', q) < 0$. When $a = m$, the reward value should be varied according to the network delay state and relative block length, as specified in Eq(5) to Eq(8).

$$R(< low, l - 1 > | < low, l >, m) = -c_m - c_d \tag{5}$$

$$R(< high, l - 1 > | < high, l >, m) = -c_m - c_d - \epsilon \tag{6}$$

$$R(< low, l + 1 > | < low, l >, m) = R_b - c_m - c_d \tag{7}$$

$$R(< high, l + 1 > | < high, l >, m) = R_b - c_m - c_d \tag{8}$$

For a wait action with $a = w$, the rewards is zero for all the cases except a special case where the network delay is low and $R(< low, l > | < low, l >, w) = -\epsilon$.

• **Observation**. When an attacker performs a query action, the current network delay states of each node are collected and calculated. Each calculated network delay is regarded as an observation. The observation value is denoted by o, which has two values in the set $\{high, low\}$.

• **Observation function**. It can be described by $Pr(o|s', q)$. In our POMDP model, only the network delay state is partially observable, and the relative length state is completely observable. Therefore, $Pr(o|s', q)$ is equivalent to $Pr(o|t', q)$, which is learned based on the current and history observation results.

2.1.2 Belief Function Updation. The POMDP model maintains a belief b over the network delay state, which is a probability distribution over all the possible states. Suppose $b(s)$ be the probability that the network state s happens, the updated belief state $b'(s)$ is calculated whenever action a is taken and observation o is received.

$$b'(s) = Pr(s|o, a, b) = \frac{Pr(o|s, a)}{Pr(o|b, a)} \sum_i Pr(s|i, a)b(i) \tag{9}$$

Given an observation o_t of the network delay at time t, current state s the belief of the current state $b(s)$ is then updated to $b'(s)$.Therefore the belief updation is then given by $b'(s) \propto Pr(o_t|s, q)b(s)$.

2.1.3 Learning Observation Function . A Dirichlet distribution $Dir^s(\phi^s_{o_t}, o_t \in \{high, low\})$ is initialized by using observation history set. The posterior probability of the occurrence of observations can be calculated according to the number of times $o_t = high$ and $o_t = low$ happening in the observation history. After a period of time and several rounds of observation, we get a set of posterior probability vectors and then process them based on a classical K-Means clustering method where K can be adjusted according to the observation time such that the overall clustering error is acceptable. After clustering, we select the nearest category C and obtain the observation function by getting the center of C. It can be used to initialize a new POMDP model which will be able to reflect current network delay state.

2.2 Blockchain Security Analysis

We assume that in time window W, the expected rewards of an attacker is $E(R_a)$ which is calculated based on strategy policy behaving in the blockchain system. At the same time, within the same time W, the expected rewards of an honest miner is$E(R_h)$. If the rewards of the attacker is greater than that of an honest miner in the same time, that is

$$E(R_a) > E(R_h) \tag{10}$$

then our analysis model will draw a conclusion that the analyzed blockchain system is risky against the long delay attack. On the contrary, if $E(R_a) \leq E(R_h)$ then the blockchain system is justified to be safe against the long delay attack.

3 CONCLUSION

In this paper, we build an analysis model based on the POMDP model to evaluate the security of a blockchian system against the long delay attack. In the POMDP model, the six main components are constructed so as to assistant an attacker to generate an optimal strategy policy, where a clustering based observation function is learned. Based on the strategy policy output by our POMDP model, the expected rewards of the attacker is achieved, which is then compared to the expected rewards of an honest miner. The analyzed system is then justified to be safe if the expected rewards of the attacker is no greater than that of the honest miner.

REFERENCES

[1] Charalampos Alexopoulos, Yannis Charalabidis, Aggeliki Androutsopoulou, Michalis Avgerinos Loutsaris, and Zoi Lachana. 2019. Benefits and Obstacles of Blockchain Applications in e-Government. In *52nd Hawaii International Conference on System Sciences, HICSS*. 1–10.

[2] T. Dey, S. Jaiswal, S. Sunderkrishnan, and N. Katre. 2017. HealthSense: A medical use case of Internet of Things and blockchain. In *International Conference on Intelligent Sustainable Systems (ICISS)*. 486–491.

[3] Satoshi Nakamoto. 2008. Bitcoin: A peer-to-peer electronic cash system. *Consulted* (2008).

[4] Rafael Pass, Lior Seeman, and Abhi Shelat. 2017. Analysis of the blockchain protocol in asynchronous networks. In *Annual International Conference on the Theory and Applications of Cryptographic Techniques*. Springer, 643–673.

[5] Puwen Wei, Quan Yuan, and Yuliang Zheng. 2018. Security of the Blockchain Against Long Delay Attack. In *24th International Conference on the Theory and Application of Cryptology and Information Security*. 250–275.

[6] Qi Xia, Emmanuel Boateng Sifah, Ke Huang, Ruidong Chen, Xiaojiang Du, and Jianbin Gao. 2018. Secure Payment Routing Protocol for Economic Systems Based on Blockchain. In *International Conference on Computing, Networking and Communications ICNC*. 177–181.

[7] Shuangfeng Zhang, Yuan Liu, and Xingren Chen. 2019. BIT Problem: Is There a Trade-off in the Performances of Blockchain Systems. In *International Conference on Blockchain and Trustworthy Systems*. 123–136.