

# 以非齊次卜瓦松模型探討比特幣 區塊鏈雙重支付問題

呂易錦<sup>1</sup> 馬瀾嘉<sup>1\*</sup> 林億雄<sup>2</sup>

<sup>1</sup>國立成功大學統計學系暨數據科學研究所、金融科技商業創新中心

<sup>2</sup>台灣首府大學教育研究所

## 摘要

比特幣是世界上最昂貴的加密貨幣，它是在互聯網上的區塊鏈技術，比特幣系統的安全性是透過稱為挖礦的工作量證明演算法來解決。該技術利用網路對線上支付的全部交易，加上時間戳記和併入不斷延伸的工作量證明鍊條作為交易紀錄，除非更新全部的工作量證明，形成的交易紀錄將無法更改。雙重支付問題可以看作是攻擊者和誠實節點之間的競賽，當攻擊者的挖礦速度比誠實節點快時，攻擊就會成功。

在本研究中，使用兩個不同強度函數的非齊次卜瓦松過程：冪法則過程和對數線性過程，對累積事件進行建模。我們從比特幣礦池數據服務網站下載資料，根據以前未發佈的實際數據進行模擬，模擬參數是透過最大概似估計法估算的，我們分別模擬攻擊者和誠實節點的累積發生次數，以計算經驗機率。模擬顯示交易後 36 小時會成功攻擊的機率為 0.003，交易後 48 小時會成功攻擊的機率小於 0.001，這與比特幣協議提出交易後有 6 個區塊鏈加入時就能確認交易的結果不同。這項研究使我們對雙重支付攻擊的可能性有更進一步的了解，可做為比特幣社群未來制訂新政策的依據。

**關鍵詞：**強度函數；經驗機率；非齊次卜瓦松過程；工作量證明。

\*通訊作者聯絡方式：mcma@ncku.edu.tw

# A Nonhomogeneous Poisson Process Modeling for Double Spending Problems of Bitcoin Blockchain

Yi-Chin Lu<sup>1</sup>    Mi-Chia Ma<sup>1\*</sup>    Yi-Hsiung Lin<sup>2</sup>

<sup>1</sup>Department of Statistics, Institute of Data Science, and  
Center for Innovative FinTech Business Models of National Cheng Kung University,  
Tainan, Taiwan

<sup>2</sup> Institute of Education, Taiwan Shoufu University, Tainan, Taiwan

## Abstract

Bitcoin is the most expensive cryptocurrency in the world. It uses the technology of blockchain which functions on the Internet. This system is secured and works by solving the proof of work algorithm that is called mining. This technology uses the Internet to add time-stamp on all online payment transactions, and incorporates an ever-extending proof of work chain as the transaction record. Unless all the proofs of work are updated, the formed transaction records cannot be changed. The double spending problem can be regarded as a race between the attacker and honest nodes. A successful attack happens when the attacker mines blocks faster than the honest nodes do.

In this article, we model the cumulative occurrences using a nonhomogeneous Poisson process with two different intensity functions, a power law process and a log linear process. The parameters are estimated by a maximum likelihood method based on real-world data that have not been previously published. We collected observed data from <https://btc.com/block>, and simulated cumulative occurrences of the attacker and honest nodes respectively in order to find out the empirical probability. It shows that the probability of a successful attack is 0.003 after 36 hours. It is safer for a transaction to wait for 48 hours. It is against the bitcoin protocol that a transaction is not regarded as valid until the transaction is 6 blocks deep. This research provides us better understanding of the probability of a double spending attack, therefore the bitcoin community may institute new and better policies on the basis of our research.

**Keywords:** Intensity function, empirical probability, nonhomogeneous Poisson process, proof of work.

\* Corresponding author. E-mail address: [mcma@ncku.edu.tw](mailto:mcma@ncku.edu.tw)

## 1. Introduction

The global financial crisis of 2007–2008 was a disastrous event that brought tremendous changes to the whole world. Philippon (2016) provided some points of view of the current framework and its impact. He pointed out the current regulatory progression had made some improvements but still a long way to go. His research suggested applying alternative approaches that are likely to involve fintech; such techniques might constitute an answer to these financial problems.

Cryptocurrencies and the blockchain are currently the most pertinent fintech innovations. Bitcoin as the premier use of blockchain technology was first proposed in a white paper in 2008 by Satoshi Nakamoto. Nakamoto (2008) recommended a peer-to-peer version of electronic cash system allowing any two willing participants to make transactions directly with each other without the need for a trusted third party. The stated benefits of bitcoin blockchain technology include decentralization, transparency, immutability, cryptographically highly secured transactions. Notably, blockchains are still required to prevent double spending problems.

The bitcoin blockchain protocols are essentially a form of governance. The bitcoin blockchain technology is used to govern a global decentralized ledger on the Internet. Antonopoulos (2010) defined the participants as miners, who validate new transactions and record them in the global ledger. Mining is the process by which miners compete to solve a mathematical problem that consumes massive computational resources. This effort is called the “universal hashing” (Cormen et al., 2009) which guarantees a low number of collisions providing security to the propagated global ledger. The solution to the problem is referred to as the “proof of work”. The competition between miners is the basis for bitcoin security model and systematic maintenance. Miners compete against each other with CPU power in a struggle to provide "proof of work". The miners who win this contest earn rewards, including the right to record transactions on the blockchain. The system is secure as long as a majority of CPU power, or hash power, is controlled by honest nodes.

“Double spending” means transactions spending the same set of bitcoins more than once by definition (Antonopoulos, 2010). To make a successful double spending attack, the attacker would have to redo the proof of work of the block and all blocks after it; then catch up with and surpass the blockchain of the honest nodes (Nakamoto, 2008). One attack scenario against the consensus protocols is called “the 51% attack”. If the attacker is in control of the substantial hash power of the mining pool, he would be able to attack the system successfully at will. If the hash power in control by the attacker is less than 50%, a successful double spending attack is still possible. The original bitcoin white paper contains a discussion of the probability of double spending problems being analogous to a gambler’s ruin problem (Ross, 1996).

Rosenfeld (2014) made a clarification and expansion on Nakamoto (2008). The previous two proposed analyses of the double spending issue by Nakamoto (2008) and Rosenfeld (2014) which came up with similar conclusions were rational; however, these approaches have not been proven from real data and theoretical arguments. The double spending problem can not be easily modeled on the foundation of a gambler's ruin problem. Because the target hash values and the hash power distribution change, the catching-up probability and attacker's potential progress may not always be the same.

In this paper we analyze a series of data for a hash pool, or a bitcoin mining pool, which is composed of groups of cooperating miners who are willing to share mining rewards in proportion to their contributed mining hash power. The double spending race is a competition between attacker nodes and honest nodes; both sides try to mine blocks faster than the other side. In this competition, the hash power of mining pools plays a core role. We turn to real data for more information about the truth.

The main breakthrough of this study is to establish a nonhomogeneous Poisson process approach for the double spending issue. In this research, we consider how to model the data by a nonhomogeneous Poisson process (NHPP) with the intensity function modeled by two different processes. To maximize realism, we suggest using these models to analyze the data set of the bitcoin mining pool.

The rest of this article is arranged as follows: in section 2, we describe several approaches to the double spending issue; in section 3, we propose the NHPP model; in section 4, these parameters are estimated by maximum likelihood method using real data; in section 5, we compute the empirical probability based on simulated data; finally, some concluding remarks are presented in section 6.

## 2. Description of Statistical Modeling

### 2.1. The Original Approach for Double Spending Problems

The bitcoin white paper (Nakamoto, 2008) proposed an analysis of the probability of double spending problems. This analysis can be divided into three parts. First, to simplify the problem, Nakamoto (2008) characterized the race between the honest nodes and an attacker trying to generate a chain faster as a simple random walk. Assume that the probability of each step is identically and independently distributed, the probabilities which the honest node and an attacker find the next block can be expressed as “ $p$ ” and “ $q$ ”, respectively. Given these premises, the probability of a double spending problem can be analogous to a gambler's ruin problem. The catching-up probability which the attacker will catch up from  $z$  blocks behind, named “ $q_z$ ”, is shown as follows:

$$q_z = \begin{cases} 1 & \text{if } p \leq q \\ \left(\frac{q}{p}\right)^z & \text{if } p > q. \end{cases} \quad (2.1)$$

When  $p > q$ , the catching-up probability drops exponentially with the increase of the number of blocks “ $z$ ”. That is one of the reasons that some bitcoin communities suggested waiting until 6 blocks have been added into the blockchain before the confirmation of transactions.

After that, Nakamoto (2008) made a description of the attacker’s potential progress and its function. He defined the honest nodes that were ahead of “ $z$ ” blocks and the attacker’s potential progress as “ $k$ ”. A double spending attack is started by secretly mining a branch of the latest block. Afterwards, continue extending the secret branch until it is longer than the public branch, then broadcast the transaction records to the bitcoin blockchain. Under the bitcoin protocol, the longest proof of work chain is recognized as the valid chain. In this situation, “ $z$ ” blocks have been added into the blockchain while “ $k$ ” blocks have been made behind the scenes. Nakamoto (2008) assumed the potential progress of a secret chain follows a Poisson distribution with mean  $\lambda = z(\frac{q}{p})$ .

Finally, a conjunction of the consequences between the catching-up probability and the attacker’s potential progress was addressed. According to Nakamoto (2008), it can be concluded that the system can be secured by extending the chain. This produces pattern of exponential decay as the blockchain is extended. Hence, the more blocks are added into the blockchain, the less probable a double spending attack is. However, the assumptions of this model must be considered and the conclusions of this research must be verified.

## ***2.2. A Hashrate Based Approach for Double Spending Problems***

Another way to measure the probability of a double spending attack was provided by Rosenfeld (2014). As a pioneer, Nakamoto (2008) made a simple assumption that the attacker’s potential progress follows a Poisson distribution with mean  $\lambda = z(\frac{q}{p})$ . On the basis of Satoshi Nakamoto’s effort (Nakamoto, 2008), Rosenfeld (2014) made some additional calculations. If the assumption that the random variable  $k$  defined by Nakamoto (2008) follows a Poisson distribution is true, then the waiting time in a Poisson process is memoryless. After some derivations, the probability distribution of the random variable  $k$  can be proven to be a Negative binomial distribution.

In accordance with Rosenfeld (2014), the conclusion is similar to the Nakamoto (2008). The probability of a double spending attack vanishes as the blockchain becomes longer. But the probability in the model of Rosenfeld (2014) drops more slowly than the probability in the model of Nakamoto (2008). More confirmations are required to ensure the bitcoin systematic security. However, these conclusions of this research must be verified. First, there was neither evidence nor theoretical

background to show the attacker's potential progress is a Poisson process. Then, the probability of finding the next block quickly cannot be simply expressed as “ $p$ ” and “ $q$ ” which need the assumptions of independent and identical distribution (i.i.d.) because the environment of bitcoin blockchain system changes all the time. Finally, the catching-up probability didn't behave like the stochastic process of a gambler's ruin problem because of the violation of the i.i.d. assumption.

### 2.3. Theoretical Background

Bitcoin mining is the process of solving the “proof of work” algorithm. To be clear, mining is to guess the nonce value then transform the information of transactions into a hash value repeatedly. For a valid block, the hash value, also the block header, should be less than or equal to the target hash. Cormen et al. (2009) indicated that a good hash function approximately satisfies the assumption of simple uniform hashing. Bitcoin blockchain employs SHA-256 hash function to operate the system. The outcomes of SHA-256 hash function are uniformly distributed over the set  $\{0, 1, 2, \dots, 2^{256}-1\}$  composed of natural numbers. The number of all the possible outcomes is  $2^{256}$ . Let “ $D$ ” be the number of the set composed of valid hash values. The SHA-256 hash function is computationally indistinguishable (Coron et al., 2017), so miners must solve the “proof of work” algorithm by the “method of exhaustion”. Since SHA-256 approximately satisfies the assumption of simple uniform hashing, the probability of each guess is  $\frac{D}{2^{256}}$ .

With each guess being independent, repeatable, and having the same probability  $\frac{D}{2^{256}}$ , the process of bitcoin mining can be regarded as a Bernoulli trial. The total number of guesses needed to mine a block follows a geometric distribution with a very large mean  $\frac{2^{256}}{D}$ . It is trivial that the exponential distribution is a limiting case for the geometric distribution when the success probability, denoted as  $a$ , is small and the time is measured in units that are of size  $\frac{1}{a}$ .

As a result, the waiting time to mine a bitcoin block can be accurately described using some exponential distribution; that implies the cumulative occurrences can be modeled using Poisson distribution. With no doubt, the number of blocks mined in a specific time interval follows a Poisson distribution when the target hash value does not change. But the target hash value changes every 2016 blocks to make the expected waiting time to be ten minutes. Different probability of each guess,  $\frac{D}{2^{256}}$ , over different time intervals ought to have different means. Hence, in this paper we model the cumulative number of mined blocks, denoted as  $N(t)$ , using the model of nonhomogeneous Poisson process (NHPP).

## 2.4. Nonhomogeneous Poisson Process

The nonhomogeneous Poisson process, abbreviated as NHPP, can be applied in various fields. The modeling of the cumulative number of blocks mined can be regarded as a counting process. Under some conditions,  $N(t)$  is defined to be the cumulative number of blocks mined by the nodes during time interval  $(0, t]$  for  $t > 0$ , where  $N(t)$  is modeled by a nonhomogeneous Poisson process with mean function  $\Lambda(t)$ . One can also specify  $N(t)$  by the intensity function  $\lambda(t) = \frac{d\Lambda(t)}{dt}$ . The probability density function of this nonhomogeneous Poisson process is

$$P\{N(t) = n\} = \frac{\{\Lambda(t)\}^n}{n!} e^{-\Lambda(t)} \quad (2.2)$$

where  $n = 0, 1, 2, \dots$

If the intensity function,  $\lambda(t)$ , is constant, then  $N(t)$  is a homogeneous Poisson process, which is a special case of NHPP. Two types of models for the mean value function  $\Lambda(t)$  which should be non-decreasing and bounded above are considered in this research. The first type is the well-known power law process (Steven and Aist, 1989), which is given by the intensity function

$$\lambda_1(t) = \frac{\beta}{\theta} \left(\frac{t}{\theta}\right)^{\beta-1}, \quad t \geq 0. \quad (2.3)$$

The corresponding mean function is

$$\Lambda_1(t) = \left(\frac{t}{\theta}\right)^\beta, \quad t \geq 0, \quad (2.4)$$

where  $\beta$  is the shape parameter. This NHPP model can sometimes be referred to as the Weibull intensity. The second type is a NHPP process with a log linear intensity function (Soufiane, 2012)

$$\lambda_2(t) = e^{\gamma_0 + \gamma_1 t}, \quad t \geq 0. \quad (2.5)$$

The corresponding mean function is

$$\Lambda_2(t) = \frac{e^{\gamma_0(e^{\gamma_1 t} - 1)}}{\gamma_1}, \quad t \geq 0, \quad (2.6)$$

The main idea of this paper is to model the mining progress of the honest nodes and the attackers using these two aforementioned NHPP models based on the mining pool data, then find the probability of a successful double spending attack. To do this, we find out the intensity functions that can best fit the data set.

## 3. NHPP Modeling

### 3.1. Parametric Estimation of The Intensity

Let  $D_T = \{n; t_1, t_2, \dots, t_n; T\}$  be the data set, where  $n$  is the number of events observed such that

$0 \leq t_1 < t_2 < \dots < t_n \leq T$ , and  $t_i$  are the times of the events observed during the period of time  $(0, T]$ . We consider that parameters of these intensity functions,  $\theta$ ,  $\beta$ ,  $\gamma_0$  and  $\gamma_1$ , are unknown and must be estimated. In our NHPP modeling, for each parameter, we use maximum likelihood method to establish our proposed model.

Considering  $T$  as the truncation time of the model (see, for example, Cox and Lewis, 1966), for the power law process, the likelihood function for the parameter vector  $(\theta, \beta)$ , is given by

$$L(\theta, \beta | D_T) = (\prod_{i=1}^n \lambda_1(t_i)) \exp(-\Lambda_1(T)). \quad (3.1)$$

According to Bain and Enghardt (1991), the MLE estimators of  $\theta$  and  $\beta$  are

$$\hat{\beta} = \frac{n}{\sum_{i=1}^n \ln(\frac{T}{t_i})} \quad (3.2)$$

$$\hat{\theta} = \frac{T}{n^{\frac{1}{\beta}}} \quad (3.3)$$

For a log linear process, the likelihood function for the parameter vector  $(\gamma_0, \gamma_1)$ , considering  $T$  as the truncation time of the model (see, for example, Cox and Lewis, 1966), is given by

$$L(\gamma_0, \gamma_1 | D_T) = (\prod_{i=1}^n \lambda_2(t_i)) \exp(-\Lambda_2(T)). \quad (3.4)$$

In accordance with Kariuki et al. (2012), we must solve the following equations in order to find out the MLE estimators of  $\gamma_0$  and  $\gamma_1$ :

$$e^{\hat{\gamma}_0} = \frac{n\hat{\gamma}_1}{e^{\hat{\gamma}_1 T} - 1} \quad (3.5)$$

$$\sum_{i=1}^n t_i = \frac{nT}{1 - e^{-\hat{\gamma}_1 T}} - \frac{n}{\hat{\gamma}_1}. \quad (3.6)$$

Since the explicit solution for equation (3.6) does not exist, it may be solved numerically using the Broyden–Fletcher–Goldfarb–Shanno algorithm.

### 3.2. Double Spending Probability

Numerous theoretical arguments support the NHPP modeling for double spending problems on the bitcoin blockchain. Nevertheless, we usually have no idea about the parameter vectors of the mean functions and intensity functions. Different choices of statistical methods, such as parametric estimation, semiparametric estimation, and nonparametric estimation, can be applied for NHPP modeling. The MLE method, which is a special case of parametric estimation, plays a key role to complete our study.

With each parameter of all the intensity functions and mean functions well estimated, in this research we apply two NHPP models as just mentioned to calculate the double spending probability.



Denote the time “ $t_s$ ” as the time the attacker starts an attack and “ $t$ ” is over the interval  $(0, T]$ , where  $t > t_s$ . Then we define  $N(t)$  and  $N'(t)$  to be the block mined by the honest nodes and attacker, respectively. And the double spending probability would be

$$P(t) = \sum_{n=0}^{\infty} P([N'(t) - N'(t_s)) - (N(t) - N(t_s)] > 0 | N(t) - N(t_s) = n) P(N(t) - N(t_s) = n). \quad (3.7)$$

We have to use the programming skills to calculate the probability since there is no close form for the double spending probability.

#### 4. Mining Pool Real Data

Bitcoin is the most valuable and most widely used digital currency in the world. It provides a new path to transactions. But a problem must be prevented, which is called the double spending problem. Another digital currency, Bitcoin gold (BTG), faced a double spending attack, that caused a loss of 17.5 million U.S. dollars on May 24, 2018. The double spending problem is truly a portentous challenge for all the digital currencies.

It is almost impossible for a node to attack the bitcoin blockchain successfully on its own. But for an alliance of mining pools with the advantage of high CPU power, it is possible to attack successfully. In this section, we apply these two aforementioned NHPP models to complete our research. It is hard to ally two huge groups to launch an attack since it is not easy to reach a consensus. But many things are possible on the Internet. We have learned a painful lesson from the episode of the double spending attack of BTG on May 24, 2018. To be conservative, we consider the most rigorous situation, which is similar to the concept of a “stress test”.

Here, we assume the alliance of “BTC.com” and “Antpool”, which are the largest two bitcoin mining pools now, to be the attacker to perform a statistical analysis to calculate the probability of a successful double spending attack under the scale of CPU power. The main idea is to apply the current distribution of hash power to complete the analysis since the distribution of hash power is dynamic. We take these two NHPP processes mentioned above to fit the data corresponding to the mining pool data from November 4, 2018 to November 6, 2018. These data are collected from <https://btc.com/block>, and they account for about 72 hours of observations. We regard the moment of the last block mined on November 3, 2018 as time zero. This data set is reorganized into temporal form so that we can compute the interarrival time of all the blocks.

These estimates of the fitted NHPP models for the attacker and honest nodes are estimated by the MLE method. Then we construct the confidence intervals for each parameter under the significance

level of 0.05 to describe the uncertainty. For the attacker, the summaries of the power law process model and the log linear process model are listed in Table 1.

Table 1. Summaries of MLE for the attacker

Model	Parameter	Estimate	95% confidence interval	
			Lower bound	Upper bound
power law process	$\beta$	0.958	0.801	1.147
	$\theta$	0.488	0.195	1.218
log linear process	$\gamma_0$	0.530	0.173	0.888
	$\gamma_1$	-0.0005	-0.009	0.008

For the honest nodes, the summaries of the power law process model and the log linear process model are listed in Table 2.

Table 2. Summaries of MLE for the honest nodes

Model	Parameter	Estimate	95% confidence interval	
			Lower bound	Upper bound
power law process	$\beta$	0.995	0.887	1.115
	$\theta$	0.239	0.123	0.464
log linear process	$\gamma_0$	1.406	1.117	1.634
	$\gamma_1$	-0.0001	-0.006	0.005

Figure 1 and Figure 2 present the theoretical and empirical cumulative occurrences. We can observe that both the graphs of the cumulative occurrences estimated by power law processes and by log linear processes fit the data set very well. We also use the Akaike information criterion (AIC) proposed by Akaike (1974) to compare these NHPP models. This criterion is widely used in statistical model selection. The summaries of AIC values for different models and different parties are listed in Table 3.

Table 3. Akaike information criterion (AIC)

Model	attacker	honest nodes
power law process	120.162	-231.392
log linear process	120.365	-231.386

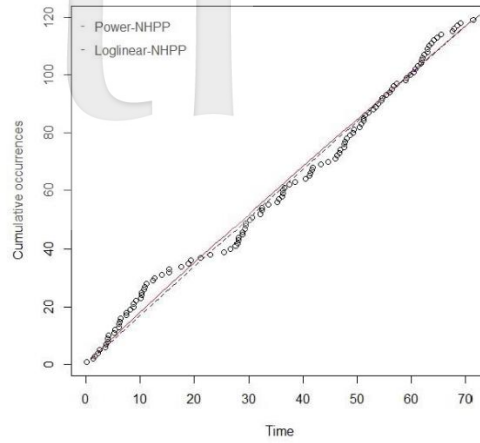


Fig.1 Fitted NHPP model for the attacker

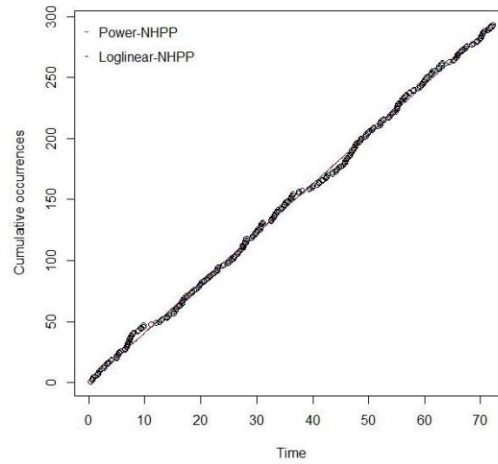


Fig.2 Fitted NHPP model for the honest nodes

## 5. Empirical Probability

In the simulation studies, we considered the NHPP models with the power law intensity and the log linear intensity. These parameters of the proposed models were initially estimated by MLE method based on the mining pool data of the bitcoin blockchain. These models with accurately estimated parameters were used to discover the empirical probabilities based on simulated data because no closed form to calculate the theoretical probability. According to law of large numbers, the empirical probability must converge to the theoretical probability when the sample size is large. Hence, we will simulate data with large sample size in this section.

For each NHPP model, we simulated 10000 observations with parameters  $\theta = 0.488$ ,  $\beta = 0.958$ ,  $\gamma_0 = 0.53$  and  $\gamma_1 = -0.0005$  for the attacker; with  $\theta = 0.239$ ,  $\beta = 0.995$ ,  $\gamma_0 = 1.406$  and  $\gamma_1 = -0.0001$

for the honest nodes, respectively. The empirical probability function of the proposed NHPP models would be

$$P(t) = \frac{I([N'(t)-N'(t_s)]-[N(t)-N(t_s)]>0)}{10000} \quad (5.1)$$

, where  $I(\cdot)$  represents the indicator function. We obtained the empirical probabilities shown in Table 4 by applying the simulated data to equation (5.1).

Table 4. Empirical probabilities

Hours	power law process	log linear process
1	0.461	0.468
2	0.428	0.418
3	0.392	0.403
6	0.306	0.308
9	0.228	0.231
12	0.165	0.159
15	0.109	0.111
18	0.070	0.076
21	0.049	0.048
24	0.028	0.030
36	0.003	0.003
48	<0.001	<0.001

It is obvious that the computing results of the two NHPP models did not show a large difference. Hence, we will use the consequences of the two process to state our conclusions.

## 6. Concluding Remarks

We establish a statistical model based on mining pool data to compute the probability of a successful double spending attack. We know the fact that the intensity function may increase or decrease, depending on the target hash value which changes every 2016 blocks in accordance with bitcoin blockchain protocols. Therefore, we suggest applying the NHPP process to model the data.

This analysis was completed in two stages. Initially, the parametric estimation was performed by MLE method. After that, we took the estimates of these parameters of the NHPP models into application for simulating the cumulative occurrences so that we could compute the empirical probabilities of the double spending problem. The two NHPP processes, the power law process and the log linear process, turned out to have similar consequences.

We assumed the most rigorous situation that the largest two bitcoin mining pools, “BTC.com” and “Antpool”, cooperate to attack the system. It was observed that the probability of a successful attack is 0.003 after 36 hours. Under that situation, it is safer for a transaction to wait for two days which is against the bitcoin protocol that a transaction is not regarded as valid until the transaction is 6 blocks deep. There are about 300 blocks in a three day period, thus, we must wait for more blocks to avoid the risk of double spending attacks.

Bitcoin, as the pioneer application and typical use of blockchain technology, is undoubtedly a great improvement of human civilization, but it is far from perfect. The low transaction volumes and high transaction fees are the problems of bitcoin. The previous two proposed analyses of the double spending issue by Nakamoto (2008) and Rosenfeld (2014) were rational; however, these approaches lacked support from mining pool data and theoretical arguments. The double spending problem cannot be easily modeled on the foundation of a gambler’s ruin problem. The target hash values and the hash power distribution change so that the catching-up probability and attacker’s potential progress may not always be the same.

In this paper, we suggest using the NHPP model with the power law process or loglinear process to provide a measurement by computing the empirical probability of a double spending attack. These probabilities based on the current three days of mining pool data are closer to the truth than the arguments of previous papers have been. More confirmation time would be safer. More dispersive of hash power would also be preferred. A low proportion of hash power implies lower probability to attack the system. These suggestions may bring us inconvenience. After all, bitcoin is not perfect. Furthermore, the bitcoin protocols may need to be revised according to a statistical model that is more reasonable. We do not make contribution to the bitcoin blockchain system directly. Instead, we hope the bitcoin community will institute new and better policies on the basis of our study.

## Acknowledgements

This research is supported by the “Higher Education SPROUT Project” and “Center for Innovative FinTech Business Models” of National Cheng Kung University (NCKU), sponsored by the Ministry of Education, Taiwan, R.O.C.

## References

- [1] Akaike, H. (1974). A new look at the statistical model identification, *IEEE Transactions on Automatic Control*, 19(6), 716–723.
- [2] Antonopoulos, A. M. (2010). *Mastering Bitcoin*, O'Reilly Media, Inc., California.
- [3] Bain, L. J. and Englund, M. (1991). *Statistical Analysis of Reliability and Life-testing Models*, 2<sup>nd</sup> edition, Marcel Dekker, New York.
- [4] Cox, D. R. and Lewis, P. A. W. (1966). *The Statistical Analysis of Series of Events*, Methuen, London.
- [5] Coron, J. S., Lee, M. S., Lepoint, T. and Tibouchi, M. (2017). Zeroizing Attacks on Indistinguishability Obfuscation over CLT13, *The 20<sup>th</sup> International Conference on Practice and Theory of Public-Key Cryptography*, International Association of Cryptologic Research, Amsterdam.
- [6] Cormen, T. H., Leiserson, C. E., Rivest, R. L., Stein, C. (2009). *Introduction to Algorithms*, 3rd edition, Massachusetts Institute of Technology, Massachusetts.
- [7] Kariuki V., Luke A. O. and Ali I. (2012). Likelihood based estimation of the parameters of a log-linear nonhomogeneous Poisson process, *International Journal of Science and Research*, 3(9), 200-204.
- [8] Nakamoto, S. (2008). Bitcoin: A peer-to-peer electronic cash system, <https://bitcoin.org/bitcoin.pdf>. [Online; accessed 23-April-2016].
- [9] Philippon, T. (2016). The fintech opportunity, *Working Paper of National Bureau of Economic Research*, No.22476, Massachusetts.
- [10] Rosenfeld, M. (2014). Analysis of hash rate based double spending, arXiv:1402.2009.
- [11] Ross, S. M. (1996). *Stochastic Process*, 2nd edition, John Wiley & Sons, Inc., New York.
- [12] Soufiane, G. (2012). Estimating parameters of a log-linear intensity for a repairable system, *Applied Mathematical Modelling*, 37(6), 4325-4336.
- [13] Steven E. R. and Asit P. B. (1989). The power law process: a model for the reliability of repairable systems, *Journal of Quality Technology*, 21(4), 251-260.