The first glance at the simulation of the Tangle: discrete model

B. Kuśmierz*

IOTA Foundation

November 6, 2017

Abstract

In this paper we present preliminary results obtained with the computer simulation of the Tangle - directed acyclic graph adapted for decentralized information storage. The first scalable, permissionless distributed ledger to use this technology is IOTA. IOTA protocol is designed for Internet of Things, Web 3.0 and other applicable sectors where the standard blockchain architecture comes up short. We examine basic properties of the Tangle, this include analysis of cumulative weight and stability of tips number, for different tip selection mechanisms.

1 Introduction

Today's constant internet connection is a fact and smart devices are becoming more and more popular. This brings a need for development of protocol for Internet of Things (IOT). In order to make such network function properly it needs convenient payment system and rapidly growing cryptocurrencies seem to be a natural candidate. No doubt cryptocurrencies are one of the most innovative and disruptive inventions of the last decade. However, the vast majority of them are based on blockchain technology which has some flaws, especially noticeable within the IOT industry. What makes blockchain based payment system hard to implement in all smart devices are scalability issues. Constant growth of transaction fees is major obstacle in establishing true micropayments. One of the solutions to this problem could be an increase of block size or decrease of time between blocks. However such move brings need for more storage space and that might threaten the very idea of trustless peer-to-peer cryptocurrency. Satoshi Nakamoto [1] wanted a radical decentralization of the proposed payment system. Requirement of large storage space might concentrate full nodes in the hands of the few richest people. Another characteristic of standard blockchain not suited for IOT industry is the division of users into two groups: those who issue transactions, and those who approve transactions (miners). Both groups have fundamentally different goals and roles. Blockchain scalability issues are intensively studied and there are variety of promising solutions like the lightning networks [2], proof of stake protocols [3] and novel architecture of DAG (directed acyclic graphs) also known as "the Tangle" [4]. The first cryptocurrency-platform based on this idea is IOTA, developed by the IOTA foundation.

Currently IOTA foundation is developing a computer simulation of the Tangle network. We hope this tool will improve our understanding of the system and allow us to find optimal parameters for the smoothest performance. Moreover, we want to study possible attacks on the Tangle. It is our priority to understand them and find the best methods of protecting against them. As systematic

*author's contact information: bartosz.kusmierz@iota.org

research progresses we will publish results to the IOTA community. In this paper we present preliminary results and validate certain results from the Whitepaper [4]. This includes analysis of cumulative weight and evolution of tips number. In the future we want to share results of more advanced studies.

2 Methodology

The computer simulation had been based on formalism and algorithms proposed in the Whitepaper [4], however certain adaptations had to be made. In particular we use the discrete model instead of the continuous one, explored in Whitepaper. In the future we want to examine simulations of continuous model as well.

Terminology we use is analogical as in Whitepaper. In order to issue a transaction, users must directly approve two other transactions. If transaction A directly approves B it is denoted $A \rightsquigarrow B$. One says A indirectly approves Z if there is a sequence of length at least three of transactions satisfying: $A \rightsquigarrow B \rightsquigarrow ... \rightsquigarrow Z$. The weight of a transaction is proportional to the amount of work that the issuing node invested into it. In the simulation all of the transactions are assumed to have the same, constant weight. Very important notion introduced in the Whitepeper is the cumulative weight of transaction denoted $\mathcal{H}_x$. The cumulative weight is a sum of own weight of a particular transaction plus the sum of own weights of all transactions that directly or indirectly approve this transaction.

The fundamental unit of time in the simulation is a *time step*. During each time step there is on average λ transactions issued, exact number is chosen from Poisson distribution $\mathbf{Pois}(\lambda)$. Each new transaction approves two older ones. We assume that all of the calculations required for device to issue a transaction are done within one time step. Two tips selection mechanism are considered:

• Tips selection "at random" - tips are chosen from the list of available tips randomly (uniform distribution).

• MCMC algorithm (Markov Chain Monte Carlo) - random walk of particles towards the tips. Particles are released in the tangle, first pair of particles at different tips determines transactions to approve.

In the case of MCMC algorithm we use random walk of 10 particles, starting positions of particles are chosen randomly (uniform distribution) from transactions issued between 100 and 200 time steps ago. This mimic release of particles deep in the Tangle. The transition probability P_{xy} of particle moving from transaction x to y (y approves x directly, $y \rightsquigarrow x$) is analogical as in the Whitepaper[4]:

$$P_{xy} = \exp\big(-\alpha(\mathcal{H}_x - \mathcal{H}_y)\big)\left(\sum_{z:z\rightsquigarrow x}\exp\left(-\alpha(\mathcal{H}_x - \mathcal{H}_z)\right)\right)^{-1}. \tag{1}$$

Where one sums over all transactions z that approve x.

3 Results

Our analysis revolve around recreation of results obtained in the Whitepaper [4]. We focused on cumulative weight growth and stability of number of tips $L(t)$. Results are presented in a form of graphs.

3.1 Cumulative weight

Performed simulations confirm findings of the Whitepaper in the matter of cumulative weight growth. Obtained data strongly suggest existence of two phases of growth: "adaptation period" (exponential growth) and linear growth (see figs. 1, 2 and 3). This is the case for both examined tip selection mechanisms. Behavior analogical to what is observed in figs. 1-3 is typical for almost all other transactions. The only exceptions are permanent tips. Permanent tips are transactions that are either approved by negligible number of transactions or not approved whatsoever.

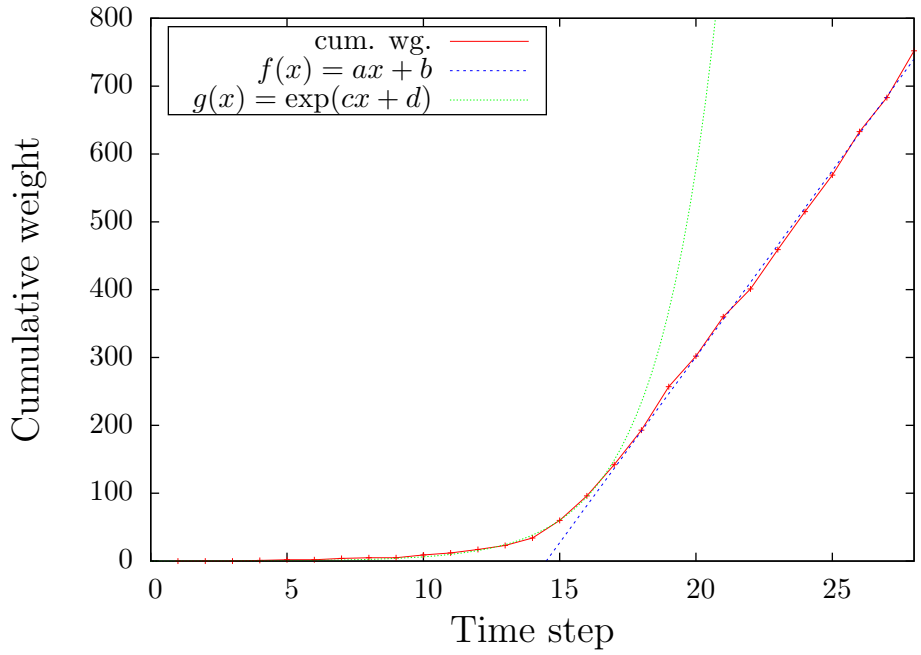

Figure 1: Cumulative weight of 200th transaction issued during the simulation and the best fitted: linear and exponent functions ($f(x) = ax + b$, $g(x) = \exp(cx + d)$ respectively). Flow rate of new transactions: $\lambda = 50$; tip selection algorithm: "at random". Linear trend of cumulative weight continues for large time steps. $f(x)$ had been fitted on an interval $[19,\ 30]$, $g(x)$ on an interval $[4,\ 16]$. Values of parameters: $a = 54.83 \pm 0.87$, $b = -795 \pm 21$; $c = 0.452 \pm 0.016$, $d = -2.69 \pm 0.25$.

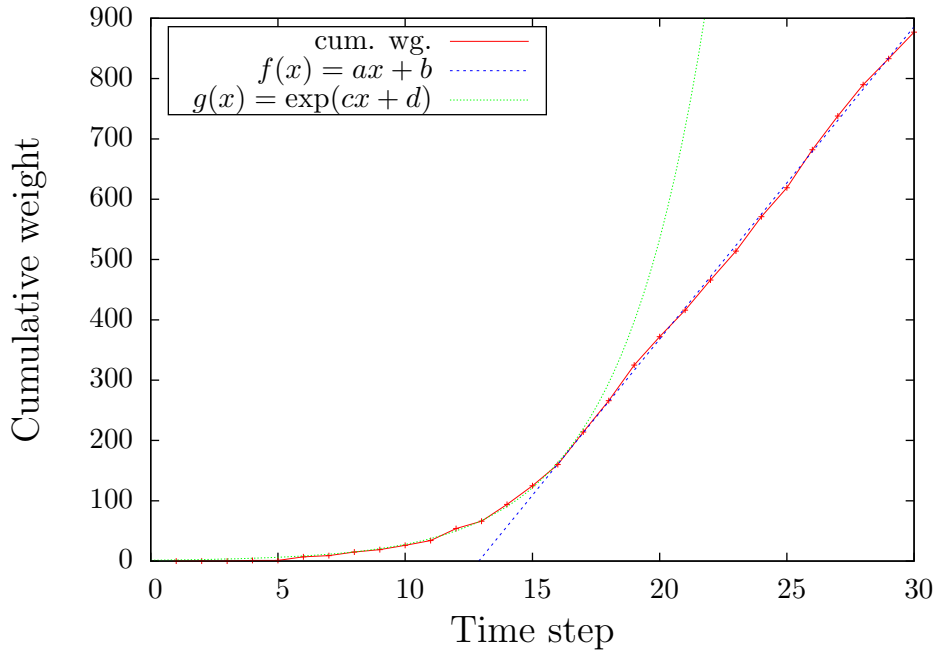

Figure 2: Cumulative weight of 200th transaction issued during the simulation and the best fitted: linear and exponent functions ($f(x) = ax + b$, $g(x) = \exp(cx + d)$ respectively). Flow rate of new transactions: $\lambda = 50$; tip selection algorithm: MCMC random walk of 10 particles towards the tips for $\alpha = 0.001$. Linear trend of cumulative weight continues for large time steps. $f(x)$ had been fitted on an interval $[22,\ 30]$, $g(x)$ on an interval $[6,\ 21]$. Values of parameters: $a = 49.94 \pm 0.40$, $b = -981 \pm 12$; $c = 0.3117 \pm 0.015$, $d = -1.96 \pm 0.30$.

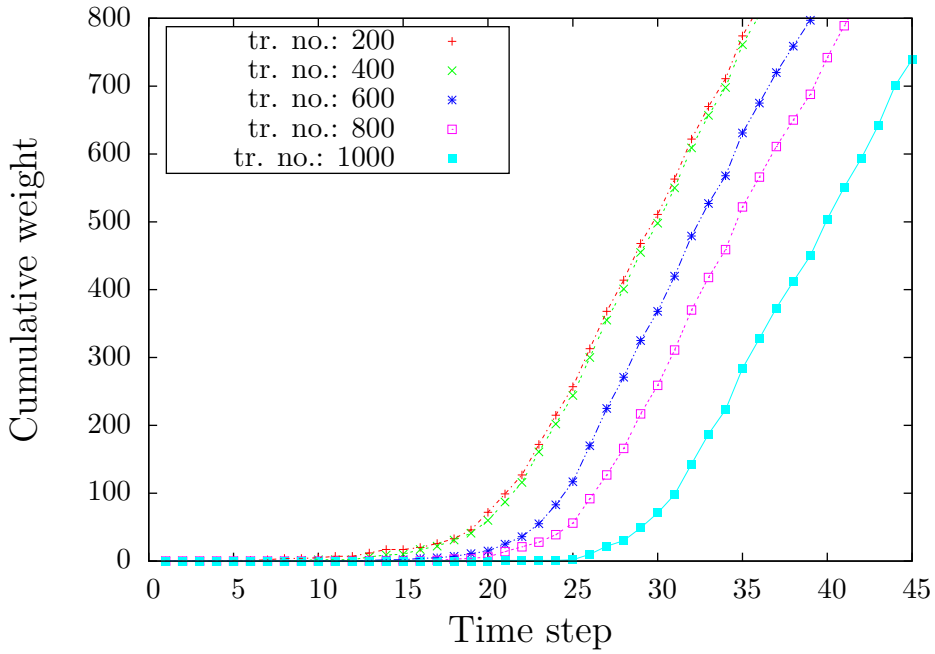

Figure 3: Cumulative weight of 200th, 400th, 600th, 800th and 1000th transaction issued during the simulation. Flow rate of new transactions: $\lambda = 50$; tip selection algorithm: MCMC random walk of 10 particles towards the tips for $\alpha = 0.001$. Linear trend of cumulative weight continues for large time steps.

3.2 Number of tips

We analyzed number of tips $L(t)$ as a function of the time step. In the case of tip selection "at random", $L(t)$ remains stable on the examined time interval (see fig. 4). Although the number of tips varies greatly, it fluctuates around an average L_0. What is worth stressing fluctuations seem to be bounded. The Whitepaper discusses this fact and makes remarks that stochastic process $L(t)$ should be positive recurrent, in particular probabilities $\mathbb{P}[L(t) = k]$ should stabilize as time goes to infinity. This fact is confirmed by a "well-behaved" histogram given in a fig 5.

The Whitepaper predicts an average value of tips number L_0. We want to emphasize that data obtained in the simulation differs from predictions of Whitepaper. The average value of $L(t)$ for data presented in fig. 4 equals $63.037 = \lambda \cdot 1.2607$. We find this discrepancy to be a result of differences between continuous and discrete model (used in the Whitepaper and the simulation respectively). We will explore differences between two models in the future [6].

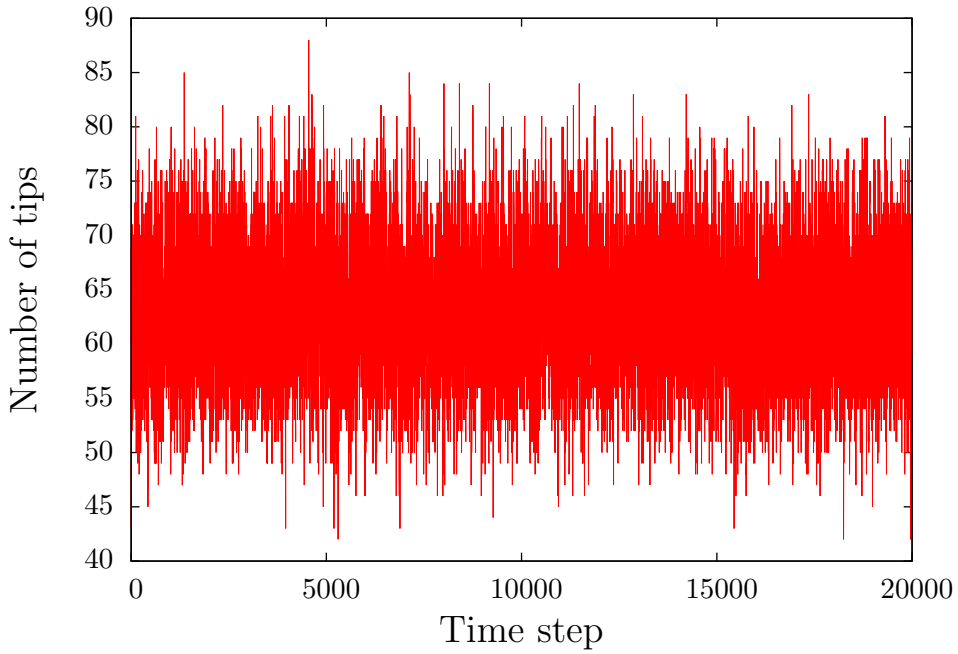

Figure 4: Values of $L(t)$ - number of the tips (unconfirmed transactions) at any given time step of the simulation. Flow rate of new transactions: $\lambda = 50$; tip selection algorithm: "at random"; simulation involves 10^6 transactions what corresponds to roughly 20000 time steps ($10^6/\lambda$). $L(t)$ varies grately, but fluctuates around an average.

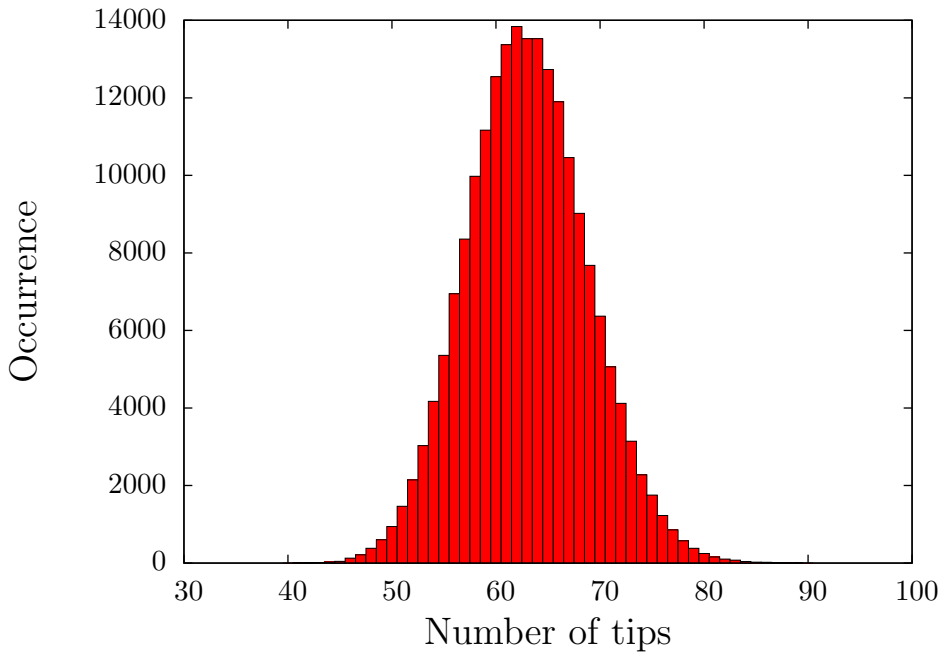

Figure 5: Histogram for values of $L(t)$. Flow rate of new transactions: $\lambda = 50$; tip selection algorithm: "at random"; simulation involves 10^7 transactions.

In the case of MCMC, when parameter α is small, $L(t)$ is expected to exhibit similar behavior as for the tip selection "at random". Indeed, fig. 6 shows that on examined time interval, $L(t)$ seems to be stable. There are theoretical arguments, based on classical results from probability theory that for any positive α, number of tips actually diverges. One can notice that each tip has nonzero probability of being unapproved for any arbitrarily large, but finite time. However after certain time, starting positions of particles are placed further than the considered tip and there is no chance of approving it. Thus each transaction has nonzero probability of becoming permanent tip and by Borel-Cantelli lemma number of tips will become infinite in the limit of large time[5]. Nonetheless obtained data shows that growth of number of tips is very small, actually unobservable on examined time interval. On the other hand random walk of particles, when α is large should be concentrated near fixed paths, determined by the highest cumulative weight. Then MCMC tip selection mechanism no longer resembles tip selection "at random" and $L(t)$ undergo different evolution. As simulation suggests in this regime number of tips grows linearly and slope of line can be significant. Example of such behaviour is given in the fig. 7 ($\alpha = 5$).

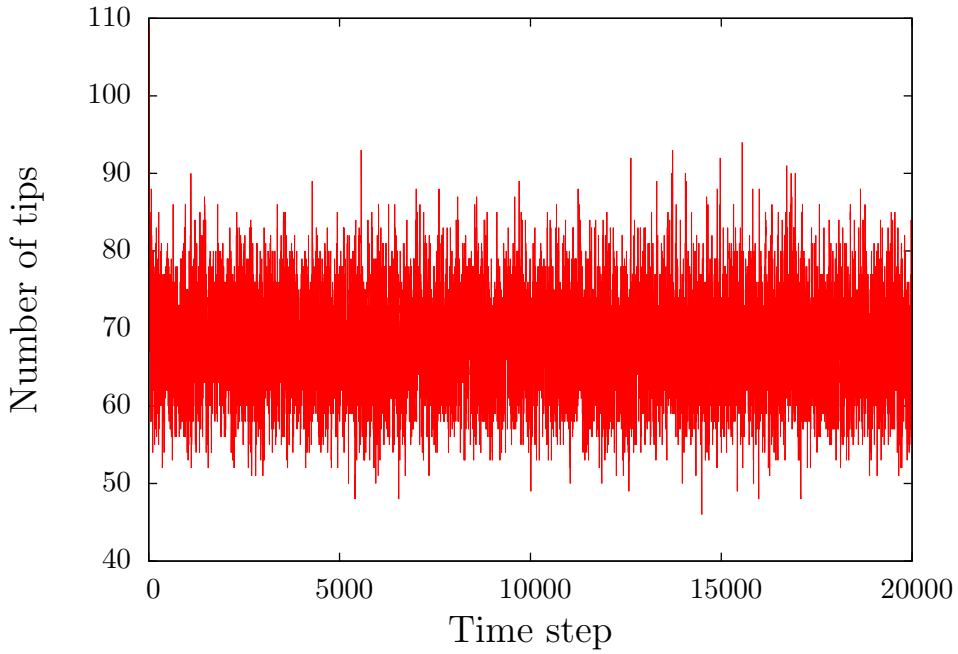

Figure 6: Values of $L(t)$ - number of the tips (unconfirmed transactions) at any given time step of the simulation. Flow rate of new transactions: $\lambda = 50$; tip selection algorithm: MCMC random walk of 10 particles towards the tips for $\alpha = 0.001$; simulation involves 10^6 transactions. $L(t)$ varies greatly, but fluctuates around an average.

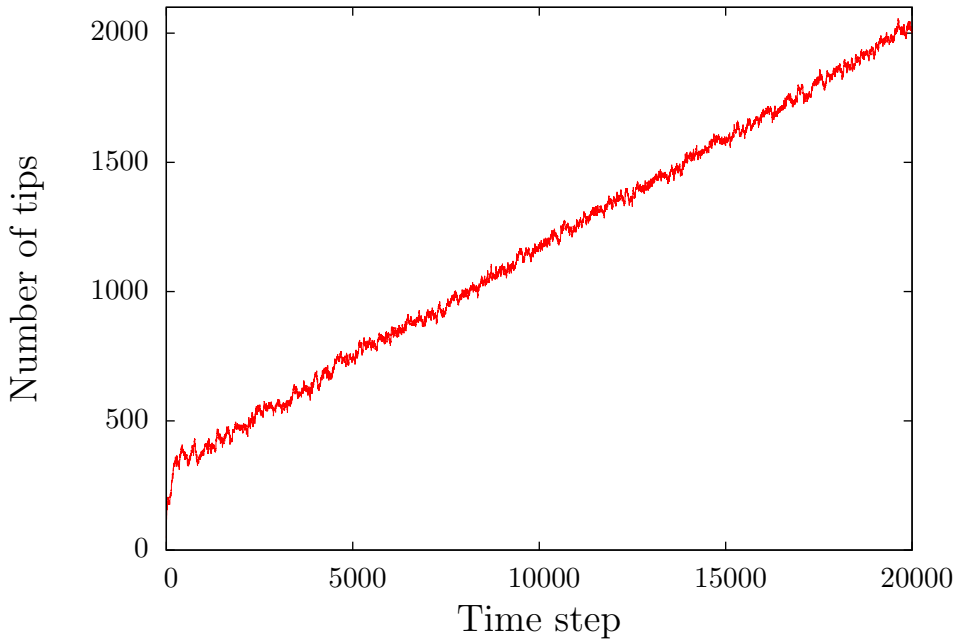

Figure 7: Values of $L(t)$ - number of the tips (unconfirmed transactions) at any given time step of the simulation. Flow rate of new transactions: $\lambda = 50$; tip selection algorithm: MCMC random walk of 10 particles towards the tips for $\alpha = 5$; simulation involves 10^6 transactions. $L(t)$ grows approximately like linear function of time.

4 Conclusions and further study

The presented data confirm some of the results obtained analytically in the Whitepaper [4]. Analysis of the cumulative weight revealed two phases of growth - exponential and linear. Using the simulation we were able to examine numbers of tips $L(t)$ as a function of time. The presented plot confirm that in the case of tip selection "at random" $L(t)$ remains stable. Moreover, we examined values of $L(t)$ when tip selection is guided by MCMC. For small α, $L(t)$ remains stable in examined time interval. Even though analytical evidence indicate $L(t) \xrightarrow{t\to\infty} \infty$ (in the absence of additional efforts by the users to establish this transaction), we find rate of divergence to be insignificant (see subsection **3.2 Number of tips**). For large values of α growth rate of $L(t)$ is significant and follows linear fashion. We can not confirm findings of the Whitepaper in the matter of average number of tips L_0. However discrepancies are most probably due to the differences in used models. We are working on the simulation of the continous model and hope to examine differences between two models [6].

So far we used the simulations to recreate already known results, however we see wider range of applications. Simulations can be especially useful when analytical calculations are extremely hard and can not be performed without strong and unrealistic assumptions/simplifications. What is important the simulation allows for quantitative analysis of stability of the Tangle for different parameters and algorithms. We will be able to compare different tip selection mechanisms, probability transition functions in MCMC and other features [7]. This tool allows us for examination of different types of attacks on the network like "parasite chain attack" or influence of lazy tips. In the future we want to publish our results in more complete form.

References

[1] S. Nakamoto, "Bitcoin: A Peer-to-Peer Electronic Cash System", https://bitcoin.org/bitcoin.pdf (2008)

[2] J. Poon, T. Dryja, "The Bitcoin Lightning Network: Scalable Off-Chain Instant Payments" https://lightning.network/lightning-network-paper.pdf (2016)

[3] BitFury Group, "Proof of Stake versus Proof of Work: White Paper" http://bitfury.com/content/5-white-papers-research/pos-vs-pow-1.0.2.pdf (2015)

[4] S. Popov, "The Tangle", https://iota.org/IOTA_Whitepaper.pdf (2016)

[5] J. Jakubowski, R. Sztencel, "Wstep do teorii prawdopodobieństwa" (Introduction to the probability theory), Script, Warsaw, sec. ed. (2001)

[6] B. Kuśmierz and S. Popov, "Simulation of the tangle", in preparation.

[7] J. Propp and D. Wilson, "Exact sampling with random Markov sampling with applications to statistical mechanics", Random Structures and Algorithms, vol. 9, pp. 232-252 (1996)