Empirical Study of Silent ASICs Mining in CryptoNight Blockchain System

Edis Mekic*, Safet Purkovic*, Kristijan Kuk**
* State University of Novi Pazar/Department of Technical Sciences, Novi Pazar, Republic of Serbia
** The Academy of Criminalistic and Police Studies /Department of Informatics, Belgrade, Republic of Serbia
emekic@np.ac.rs; spurkovic@np.ac.rs;kristijankuk.ftn@gmail.com;

***Abstract*—Hash rate analysis can provide important insight on the processes within blockchain based systems. Since Hash functions can be solved just by usage of brute force calculation, block chain need to utilize calculating capabilities of different calculating machines ranging from CPU, GPU or ASICs. Hash rate determine level of the overall calculating power and manipulation with hash functions delivered to blockchain system. Manipulation with this values can compromise blockchain system. This work researched ASICs development and silent mining action on CryptoNight based block chain system. Empirical proofs of this action are derived and success level of mitigation action by core development team is analyzed.**

I. INTRODUCTION

Client-server based systems for storing of data use massive server and data storage systems. Client data are protected and controlled by server maintenance teams, and those teams have full control on data. Emergence of blockchain systems introduced new possibilities for data sharing, storing and protection. [1].

Blockchain systems are based on usage of cryptographic hash function and Peer to peer networks (P2P). Hashing is a mathematical algorithm that maps data of arbitrary size to a bit string of a fixed size (a hash). Hash is designed to be a one-way function, that is, a function which is infeasible to invert.

The only way to recreate the input data from an ideal cryptographic hash function's output is to perform a brute-force search of possible inputs. After calculation of this inputs they are compared with original hash value to see do they produce a match, or to use a rainbow table of matched hashes. Only after successful match new block can be added to system.

As direct result of this any attempt to receive solution for hash function, which is needed to create new block into block chain, we need to use extensive calculation power. This power is provided by calculating machines interconnected in P2P network [2].

Besides calculation of hash rate, block chain system require consensus need for validation of transaction and change in data within block chain systems. Attacks for compromising blockchain are based on providing majority of hashing power to system in order to change or compromise data [3].

Calculating power can be acquired by purchase of the calculating machine (CPU or GPU based) or by development of machines based on application-specific integrated circuits (ASICs).

Second type of attack can be mitigated by changing of the hashing routine by development team. This change can make those machines obsolete [4].

This research will analyze hash rate of the block chain system based on CryptoNight algorithm to empirically prove existing of the ASICs machines and their misuse. Analysis of hash rate will also provide empirical proof of efficiency of the measures taken by the algorithm development team in order to mitigate effects of ASICs machines in block chain system.

During research we analyzed changes of hash rate of network between different releases of CryptoNight algorithm. Every change required change of the routine for calculation of hash. For the CPU and GPU based system it

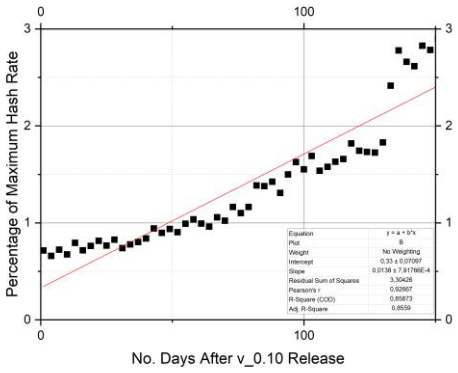

Figure 1. Hash rate trend between v.010 and v.011 CryptoNight releases

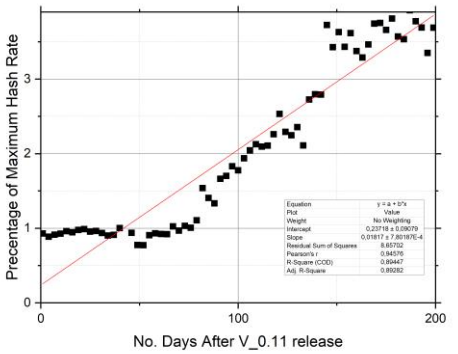

Figure 2. Hash rate trend between v.011 and v.012 CryptoNight releases

was enough to change and upgrade software for calculation. Application-specific integrated circuits (ASICs) could not follow this procedure since every change would make those machines unusable.

Since data of the total the hash rates of the network are given in any needed time, we will use those data to empirically provide proof of successful unnoticed mining action based o specialized ASICS machines..

II. Revisions of CryptoNight Algorithm and Linear Regression Trend

Representative network of CryptoNight algorithm based block chain systems is Monero network. Main idea of CryptoNight based blockchain systems are to be mineable only by GPU or CPU based machines. Developing this type of system they can protect small mining entities, which support decentralization of system. Involvement of high calculating power ASICs machines can centralize calculating power in smaller number of mining entities. Long term effect would be centralization of system.

CryptoNight white paper proposed solution of this problem based on following principles:

CryptoNight relies on random access to the slow memory and emphasizes latency dependence. Each new block depends on all the previous blocks (unlike, for example, script). The algorithm requires about 2 Mb per instance:

It fits in the L3 cache (per core) of modern processors.

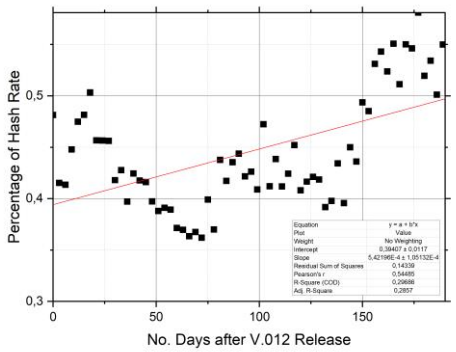

Figure 3. Hash rate trend between v.012 and v.013 CryptoNight releases

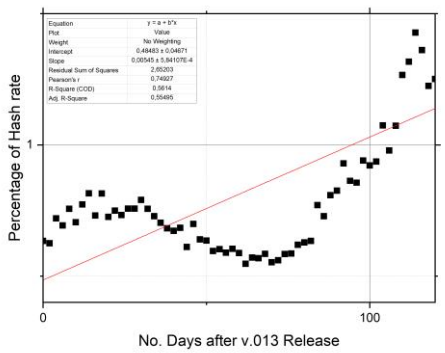

Figure 4. Hash rate trend between v.013 and v.014 CryptoNight releases

Megabyte of internal memory is almost unacceptable for the modern ASICs.

GPUs may run hundreds of concurrent instances, but they are limited in other ways. GDDR5 memory is slower than the CPU L3 cache and remarkable for its bandwidth, not random access speed [5].

CryptoNight community was stormed and surprised by announcement of Bitmain that they developed ASICs machine for successful mining of CryptoNote based blockchain systems [6].

Introducing of the ASICs machine would have immediate and obvious effect on overall has rate of the system. ASICs machines will ignite trend of sudden hash rate change and increase. This change should be statistically different when compared with trend of hash rate increase after regular upgrades of the CryptNight protocol. Also question arises did Bitmain used those machines for silent Monero mining before official delivery date.

Background mining activities will be analyzed by following trend of hash rate change on overall network between two revisions of the block chain system.

Analysis will be concluded by taking highest amount of hash rate before new release as maximum value of Network hash rate. Regarding this value we will calculate percentage drop of hash rate and recovery rate of hash rate after change of protocol.

Starting point of analysis was release of Monero v0.10 *Wolfram Warptangent revision*, This was a mandatory update due to the planed hard fork activities. This hard fork was implemented because unexpectedly high adoption rate of RingCT transactions necessitating a modification of the dynamic block size limiter algorithm.

Second point of analysis was Helium Hydra v0.11.0.0 major release of the Monero software. This was a mandatory update which increased the minimum ring signature size to 5 across the network, banned duplicate ring members in a ring signature, and enforced use of ringCT for all transaction outputs. This release of the software presented a number of major improvements to Monero network system.

Third point of analysis was Lithium Luna v0.12.0.0 point release of the Monero software, was part of the v0.12 network wide update. This major release increased the minimum ring signature size, sorted inputs to disable leak wallet choice by inference, and slightly changed the proof-of-work algorithm to prevent DoS attacks by ASICs.

This release of the software presented a number of major improvements to Monero, as well as a large set of bug fixes. Also this is release which is part of mitigating effort for suspected silent mining activity with ASICs machines.

Fourth point of analysis was Berillium Bullet release. This is the v0.13.0 release of the Monero software. This major release enabled Bulletproofs for reduced transaction sizes, set the ring size globally to 11 for uniformity of transactions, updated the PoW algorithm to CNv2, and finally set the max transaction size at half of the penalty free block size.

Final point of analysis was Boron Butterfly v0.14.0 release of the CryptoNight routine. This major release added a new PoW based on CryptoNight-R, added a new block weight algorithm, and introduced a slightly more efficient RingCT format. This is an intermediary, stable

release specifically for the network update, and does not represent the bulk of the effort on Monero. That effort will be completed in the 0.14.1 release, which will followed during March after the network update.

Trends for the Hash rate levels are calculated using following linear regression model.

Linear regression model is based on (d_i, h_i) the data set i=1,2,..n. Where d_i is i-1 day after new release, and h_i is ratio of maximal value of Hash rate before hard fork and average Hash rate on i-1 days after hard fork. Developed linear regression model fits data in form

$$h_i = \beta_0 + \beta_1 d_i + \varepsilon_i \quad (1)$$

Parameters are calculated using

$$\hat{\beta}_1 = \frac{SDH}{SDD} \quad (2)$$

$$\hat{\beta}_0 = \bar{h} - \beta_1 \mathrm{h} \quad (3)$$

Quality of regression is measured using following parameters.

First parameter of the applied fitting method is Reduced Chi-square χ value. This parameter is equal to the residual sum of square (RSS) divided by the degree of freedom. Value of a $\chi>>1$ indicates a poor model fit. Value $\chi>1$ indicates that the fit has not fully captured the data (or that the error variance has been underestimated). In principle, a value χ is equal to 1 indicates that the extent of the match between observations and estimates is in accord with the error variance. A $\chi<1$ indicates that the model is "over-fitting" the data: either the model is improperly fitting noise, or the error variance has been overestimated [7].

Second parameter is Pearson's correlation coefficient Pearson's r. Pearson's r denotes the strength of linear relationship between paired data. The value of Pearson's r can be between -1 to 1. Positive value of Pearson's r indicates that there is positive linear correlation between predictor variable and response variable. The value of zero indicates that there is no linear correlation between data. What's more the closer the value is to -1 or 1, the stronger linear correlation is [8].

Final parameter is R-square, also known as the coefficient of determination (COD). It is a percentage of the response variable variation that explained by the fitted regression line.

III. Empirical Analisys of Hash Rate Trends Beteween Protocol Releases

New releases and interconnected hard fork inevitable bring down overall hash rate of the network. Reason for this is time which miner invest in updating their software and setting up GPU or CPU based calculating machines.

Hash rate trend between v.010 and v.011 CryptoNight releases have stable trend. Linear regression data are straightforward we can see that during the time with increase of the time more calculating power is invested in the system and Pearson's r value of 0,92667 prove this conclusion. Almost 86% of calculated data fit in this model(R-Square (COD)=0,85873). While we omitted some data which are result of noise or other non-covered issues in analysis Adj. R-Square=0,8559). Hash rate recovered after 50 days and after that continued to follow linear trend.

Hash rate trend between v.011 and v.011 are interesting. Still this model fits in linear regression model which is proved with high Pearson's r of 0,94576. Adj. R-Square of 0,89282 showed that trend lines can cover more data and almost 90% of calculated data fits model (R-Square (COD)=0,89447). There is no substantial decrease of Hash rate level after new release, while in the period of official announcement of ASICs machines we have unusually high increase of Hash Rate on almost 400% in comparison with Hash rate before release.

Release v.012 change protocol and results are interesting. Drop rate of the hash rate is dramatic, and during all period this hash rate do not reach level of the prerelease hash rates.

Other interesting occurrence that in this case linear regression model almost completely fail in analysis. Pearson's r showed positive correlation of value 0,54485 while R-Square (COD) of 0,29686 showed that we can explain only 30% of presented data, all other data consist noise of the system. This is also expected since in this period blockchain market is hit with high decrease of value so there are no positive trends in investing in calculation power.

Release v,013 showed similar trend for first 100 days after new release. Value of Hash rate is similar to the levels before release. After that we have new spike in Hash rate which will be mitigated in release v.014. This spike give us better fitting with Pearson's r=0.74927, R-Square (COD)=0.5614 and Adj. R-Square=0.55495.

Since there are different occurrence which can influence increase of Hash rate (increased value of coin, DoS attacks, Bot mining activities etc.) We conducted One-Way ANOVA to try to prove is there statistical differences between V.010, v.011 and v.013 releases. Analysis showed that f-ratio value is 0.02389. The p-value is .976395. So there are not significant differences in those cases for $p < .05$.

Statistical difference between v.010, v.011 and v.012 is statistical significant. The f-ratio value is 3.26202. The p-value is .039437. The result is significant at $p < .05$.

Statistical difference between v.012 and v.013 is also statistical significant the f-ratio value is 7.39353. The p-value is .006973. The result is significant at $p < .05$.

Those differences showed that indeed between v.011 and v.012 releases was dramatically change of hash rate, which correspond to official announcement of successful development of ASICs machines. Dramatically drop of hash rate level after implementation of v.012 release showed us that huge number of calculating power is removed from system. This is indication that mitigating effort of Monero core development team was successful, and that ASICs machines were successfully removed from system.

IV. Conclusion

Statistical analysis showed that there was unusual Hash rate anomaly between v.012 and v.013 release.

This anomaly increased Hash rate to unprecedented levels in short period of time. Announcement of the implementation of ASICS machines, and Monero team mitigating effects showed that this statistical difference can be appointed to silent mining activity conducted by the company which developed ASICs machines.

While Hash rate in period when there was no mitigating efforts of the Monero core development team, linearly progress and can be statistically depicted with linear

regression model. This is not case in period of time when mitigating efforts were not implemented.

We can establish linear correlation between Hash rate and abnormal misuse of network, we can also track successful implementation of mitigation of the attack on the network.

Further research must be concluded to develop model of Hash rate change when there is no linear correlation between Hash rate increases during recovery of the network after new releases of CryptoNight hashing routine.

REFERENCES

[1] Zyskind, G., & Nathan, O. (2015, May). Decentralizing privacy: Using blockchain to protect personal data. In Security and Privacy

[2] Naor, M., & Yung, M. (1989, February). Universal one-way hash functions and their cryptographic applications. In Proceedings of the twenty-first annual ACM symposium on Theory of computing(pp. 33-43). ACM.Workshops (SPW), 2015 IEEE (pp. 180-184). IEEE.

[3] Aitzhan, N. Z., & Svetinovic, D. (2018). Security and privacy in decentralized energy trading through multi-signatures, blockchain and anonymous messaging streams. IEEE Transactions on Dependable and Secure Computing, 15(5), 840-852.

[4] Kurzweil, R., Richter, R., Kurzweil, R., & Schneider, M. L. (1990). The age of intelligent machines (Vol. 579). Cambridge, MA: MIT press.

[5] Noether, S. (2014). Review of CryptoNote white paper. HYPERLINK" http://monero. cc/downloads/whitepaper_review. pdf" http://monero. cc/downloads/whitepaper_review. pdf.

[6] https://twitter.com/BITMAINtech/status/974180147166261248

[7] Bevington, P. R., & Robinson, D. (1969). Data Analysis and Error Analysis for the Physical Sciences.

[8] Cohen, J. (1988). Statistical power analysis for the behavioral sciences (2nd ed.)