Portrait of a Miner in a Landscape

Alex Biryukov
University of Luxembourg
alex.biryukov@uni.lu

Daniel Feher
University of Luxembourg
daniel.feher@uni.lu

***Abstract*—Mining is one of the core elements of the proof-of-work based cryptocurrency economy. In this paper we investigate the generic landscape and hierarchy of miners on the example of Ethereum and Zcash, two blockchains that are among the top 5 in terms of USD value of created coins. Both chains used ASIC resistant proofs-of-work which favors GPU mining in order to keep mining decentralized. This however has changed with recent introduction of ASIC miners for these chains. This transition allows us to develop methods that might detect hidden ASIC mining in a chain (if it exists), and to study how the introduction of ASICs effects the decentralization of mining power. Finally, we describe how an attacker might use public blockchain information to invalidate the privacy of miners, deducing the mining hardware of individual miners and their mining rewards.**

***Index Terms*—blockchain, mining, ASIC, Zcash, Ethereum**

I. Introduction

Since the inception of Bitcoin [1] the proof-of-work protocol is regarded as a revolutionary new consensus protocol that allows up to 50% malicious participants in an open P2P network (modulo Selfish-mining attacks). Its core building block is mining, which requires a participant to solve a cryptographic puzzle, where the most efficient way of solving is random guessing. Every time a user solves this puzzle, he becomes a leader in the consensus protocol and creates a new block of accepted transactions. In order to control the number of blocks, the difficulty of the puzzle is dependent on the overall hashing power of the network. The process of randomly guessing correct values to solve the puzzle is called mining.

After the launch of Bitcoin in 2009 only a few enthusiasts were mining with their desktop CPUs. As time went by and Bitcoin gained more traction, the users started experimenting how to gain higher and more stable profits from mining bitcoins. At first the more efficient GPU miners were introduced publicly in October 2010. On the other hand as the number of miners was quickly increasing, the expected time to mine a block became longer than a year.

This led to the birth of mining pools. At the end of November 2010 the first mining pool was launched. The idea was to join multiple miners into one large entity, where they would share the mining rewards based on the amount of work done towards the mining of the coin. The idea quickly took off, and in today's cryptocurrency world solo miners are a rarity.

Since then in Bitcoin and several other blockchains the main hardware to mine are application-specific integrated circuits (ASICs). On the other hand there are chains that aim to prevent ASICs and use ASIC resistant proof-of-works that favor GPU mining. One of the reasoning behind this decision is the attempt at making the chain more decentralized. Our paper mainly focuses on two such chains, namely Ethereum and Zcash. In these chains ASICs have been only recently introduced, and the community have not decided completely whether they want to defend against them or not.

Firstly in our paper we describe a general landscape of miners in these chains using both blockchain and mining pool information. This includes describing the most popular hardware per chain, what the distribution of mining power between miners is, and how GPU mined coins compare to each other in terms of profitability.

Following that, with the introduction of ASICs, questions arose whether hidden ASIC mining existed in the chain before the announcements of such hardware. We have developed methods that might be able to detect hidden ASIC mining based on mining software developer fees and the distribution of the mining power. In the observed chain (Zcash) we could give a bound on hidden ASIC mining (showing no significant hidden ASICs). The same metrics did change drastically after the public introduction of ASIC hardware to the network. We have also studied how the introduction of ASICs effects the decentralization of mining power.

Lastly, using the previously acquired knowledge on miners in these chains and the public blockchain information, we describe how an attacker might deduce the individual mining rewards and the mining hardware of a single miner, violating the privacy of miners. We also provide some countermeasures how a miner might mitigate the leakage of privacy by the choice of a mining pool and the usage of their rewards.

II. Background and Related Work

Mining pools have been receiving a larger attention recently, but mostly from a game-theoretic point of view. Eyal et al. [2] introduced the selfish miner attack, where colluding miners obtain a revenue larger than their fair share, which proves the bitcoin protocol is not incentive compatible. For the attack only 1/4 of the mining power is required, compared to the previous 1/2 bound.

Later Eyal in [3] provides a game theoretic analysis on competing mining pools attacking each other by seemingly joining the opposing pool, but not providing any actual proof-of-work. The paper calls the decision whether to attack the opponent pool or not the miner's dilemma. Later Tsavary and Eyal extended this work in [4].

978-1-7281-1878-9/19/$31.00 ©2019 IEEE

We have not seen an in-depth analysis of mining pools and miners in the literature and thus we provide it in this paper. Our analysis is mainly based on two mining pools which participate in both Ethereum and Zcash respectively. These are the pools ran by BitFly called Ethermine for Ethereum and Flypool for Zcash, and Nanopool for both Ethereum and Zcash. We have chosen them because the average and current power of each worker per miner is accessible through their API, if an attacker learns the exact Zcash or Ethereum address of the miner. In order to obtain the list of addresses in case of Ethereum an attacker can acquire this information by scanning the chain for transactions sent from the pools' main address in recent time. We have collected over a 100,000 addresses. In case of Zcash, we have implemented the techniques presented in the papers by Kappos et al. and Biryukov et al. [5], [6] to retrieve the mining payout transactions and build up a database of miner addresses, which consists of ∼25,000 addresses.

III. Mining Landscape

Ethermine is the largest, while Nanopool is the third largest Ethereum mining pool. Flypool was dominating the Zcash mining power with over 50% hash rate until the recent introduction of ASICs, and is currently the number four mining pool, while Nanopool was the second largest pool at a time, while currently it is eighth in rankings.

We were specifically interested in the workers, their momentary and average power, and their names, as we have noticed that miners often name them after the hardware itself. Overall we have acquired detailed worker information for 21,000 miners in Zcash and 52,000 in Ethereum. In these data sets we searched for the keywords of the card numbers or names. If we reduce the workers to ones that have a descriptive name, the dataset is reduced to roughly 10% of all the workers. Afterwards the histogram of the average hash rates of these workers is printable, and combining this information with the online reported hash rates for the different kinds of hardware, we can attach specific hardware to specific hash rates. Ethereum and Zcash use two very different ASIC resistant proof-of-work algorithms called Ethash and Equihash respectively. Their hash rates are measured in Mhashes per second for Ethash and solutions (Sol) per second for Equihash.

A. Ethereum

First, let us investigate Ethereum. Notice the periodic peaks in the histograms (Figure 1,2,3,4). These peaks represent the number of cards in the worker. For example in the case of the GTX 1050 we can distinguish 6 separate peaks at 14, 28, 42, 56, 70 and 84 Mhash/s respectively. Also notice that the peaks are getting wider, as the miners over- or underclock their cards by different amounts, which results in larger deviation as the number of cards increases.

Overall the most popular cards are the RX 580, RX 570 and GTX 1060 with around 6,000 workers for each. The exact data and the rest of cards that are worth mentioning are presented in Table I.

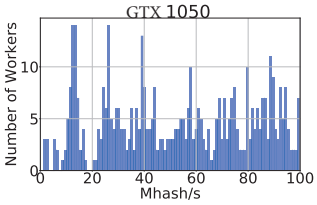

Fig. 1: Histogram of hash rates in Ethereum for the GTX 1050 GPU

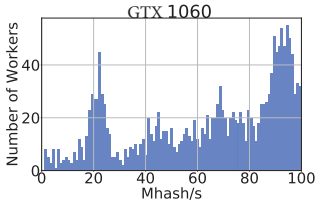

Fig. 2: Histogram of hash rates in Ethereum for the GTX 1060 GPU

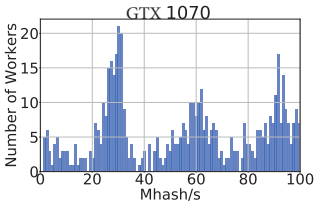

Fig. 3: Histogram of hash rates in Ethereum for the GTX 1070 GPU

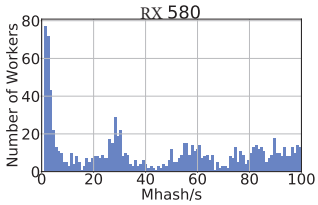

Fig. 4: Histogram of hash rate in Ethereum for the RX 580 GPU

B. Zcash

In Zcash the most popular card is the GTX 1060 and its different configurations, depending on how many of them are in a rig. The periodic peaks are observable here as well (Figure 5,6,7,8), although in this case the difference between the hash rate of cards is much larger than in Ethereum.

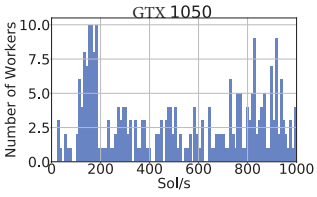

Fig. 5: Histogram of hash rates in Zcash for the GTX 1050 GPU

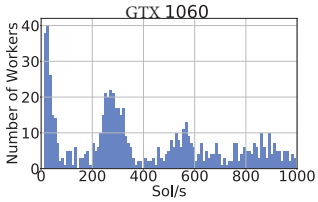

Fig. 6: Histogram of hash rates in Zcash for the GTX 1060 GPU

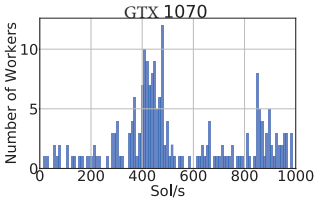

Fig. 7: Histogram of hash rates in Zcash for the GTX 1070 GPU

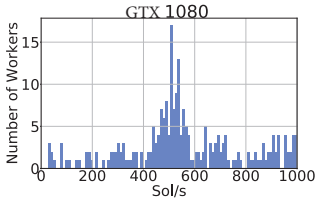

Fig. 8: Histogram of hash rates in Zcash for the GTX 1080 GPU

The larger difference in hash rate between cards provides a better distinction between the different rigs. In Table II we present the most popular rig configurations and their hash rates, while in Table III we present the most common cards and the number of times they appear.

Hardware	Num	Hardware	Num	Hardware	Num
GTX1050	631	GTX1050ti	856	GTX1060	5800
p106	1543	GTX1070	3327	GTX1070ti	616
GTX1080	855	GTX1080ti	1023		
RX460	656	RX470	3713	RX480	1650
RX560	1084	RX570	6068	RX580	6273

TABLE I: Number of times a card name has been recorded in our chosen Ethereum mining pools

Hash rate	Suspected Hardware	Hash rate	Suspected Hardware
150 Sol/s	GTX1050	870 Sol/s	3xGTX1060
270 Sol/s	GTX1060	1100 Sol/s	2xGTX1080 4xGTX1060
450 Sol/s	GTX1070	1400 Sol/s	2xGTX1080TI 5xGTX1060
550 Sol/s	GTX1080 2xGTX1060	1670 Sol/s	3xGTX1080 6xGTX1060
700 Sol/s	GTX1080TI		

TABLE II: Hash rates and their hardware counterparts

C. GPU Mining

In order to build a more complete picture we investigated Equihash and GPU mining in general, to have a better understanding of the dynamics of mining. First, we built an overview of the total Equihash-based mining ecosystem by adding up hash rates of every major blockchain using Equihash as its PoW. We show the total Equihash hash rate of these chains in Figure 9. We have identified Zcash (ZEC), Bitcoin Gold (BTG), Zencash (ZEN) and Zclassic (ZCL) as the main chains. If a chain doesn't appear on the graph until a certain point in time, it is either because it didn't exist before (Bitcoin Gold), or its hash rate was only marginal (less then 5 MSol/s[1]) compared to current Zcash. The huge increase in Zclassic's power in January and February 2018 is caused by a huge price increase during that time (prior to the BTCP-fork). Some of this power temporarily migrated from Ethereum or other GPU-mined coins.

The most notable feature of this graph is the exponential increase in mining power from June-December 2018 which is due to the introduction of ASIC miners in Zcash.[2]

One more observation regarding the topic of GPU mining is comparing Equihash mining to other ASIC-resistant PoWs. We have chosen Ethereum and Monero[3], as they are among the largest GPU mined coins. We compare the hash rate and the profitability of these Blockchains in Figures 10-11 by converting all rates into Sol/s. This is done by comparing the mining capabilities of the same GPUs on the different PoWs. The interesting observation in this graph is that even when there was a peak in Equihash mining (Oct 2017), compared to the sum there is no visible difference. This is caused by brief

[1]5 MSol/s is 5,000,000 hashes (Equihash solutions) per second.

[2]A small bump in the graph in Oct 2017 can not be explained directly, but is most likely a result of temporary ETH miner migration when ETH difficulty was rapidly increasing due to difficulty-bomb. It went back to normal when the difficulty bomb was defused by Ethereum hard-fork.

[3]By popular belief Monero's CryptoNight algorithm had hidden ASIC mining which was forked off. A large mining power drop after the fork is visible in the graph.

Hardware	Num	Hardware	Num	Hardware	Num
GTX1050	327	GTX1050ti	183	GTX1060	1783
GTX1070	791	GTX1070ti	207	GTX1080	918
GTX1080ti	617	GTX970	157		

TABLE III: Number of times a card name has been recorded in our chosen Zcash mining pools

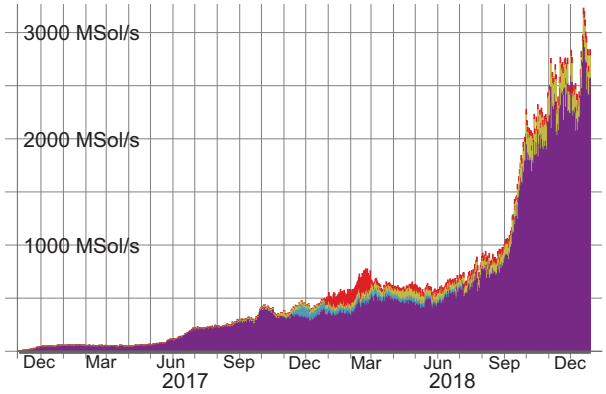

Fig. 9: Overall Equihash mining power over time (Purple: Zcash, Blue: Bitcoin Gold, Green: Zencash, Red: Zclassic)

miner migration from Ethereum mining to Zcash mining for better profitability.

Figure 11 also shows that after the introduction of ASICs in Zcash, the profitability curve has crossed the GPU-profitability line (green line "BASE"), calculated assuming an electricity price of 0.05 USD/kWh. Following graphs confirm that there are probably no GPU miners left in Zcash.

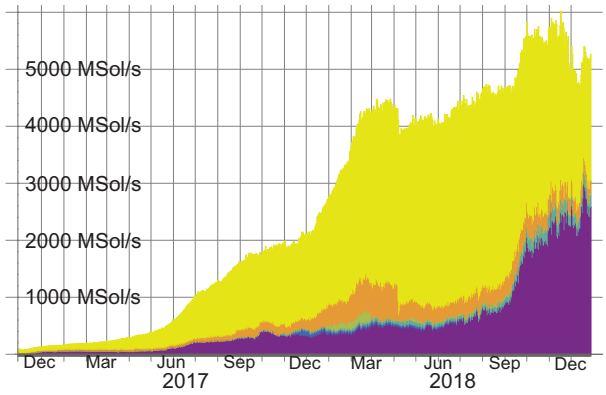

Fig. 10: Sum of total power of (formerly) GPU-mined blockchains (Purple: Zcash, Dark Blue: ZenCash, Light Blue: Bitcoin Gold, Green: Zclassic, Orange: Monero, Yellow: Ethereum)

D. GPU vs ASIC mining

Both chains (Zcash and Ethereum) use ASIC resistant proofs-of-work which favor GPU mining in order to keep mining decentralized. This however has changed with recent introduction of ASIC miners for these chains around May-June 2018. The ASIC over GPU efficiency improvement is currently around 2-5x for Ethereum's Ethash, and 10-30x efficiency improvement for Zcash's Equihash.

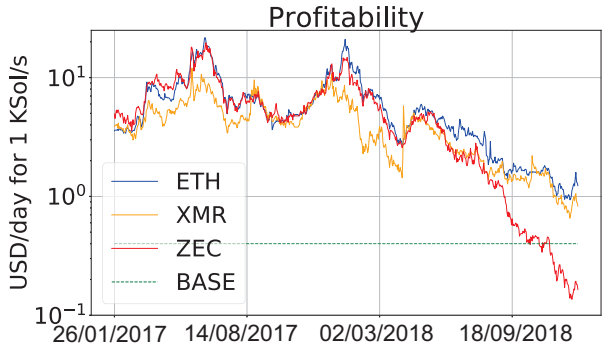

Fig. 11: Profitability in USD of different chains originally using GPU mining. (Red: Zcash, Blue: Ethereum, Orange: Monero, Green: GPU profitability line)

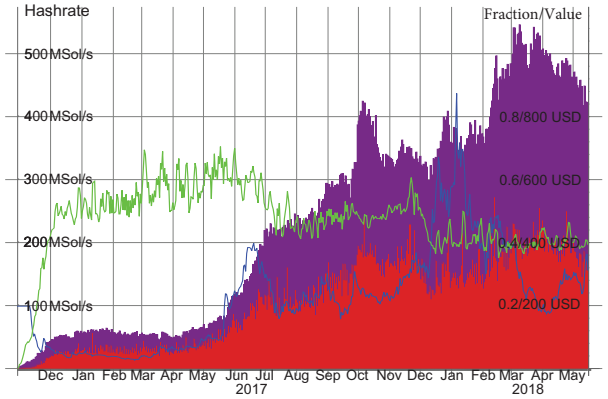

Fig. 12: Estimated portion of large miners with mining power of at least 8KSol/s, with the exact fraction in green and the ZEC/USD exchange rate in blue added

IV. Detecting ASIC miners

On May 3 the ASIC manufacturer Bitmain announced an ASIC miner for Equihash (and another company, Innosilicon later as well). Bitmain is one of the biggest ASIC provider companies in the world, while also having large shares in mining power in Bitcoin, and other large cryptocurrencies. These announcements raise the questions when did these companies develop an ASIC, did they launch and test it on Equihash-based cryptocurrencies before the announcement? Were ASIC miners present in the Zcash mining ecosystem before their official shipment dates, and if yes, to what extent? The following two techniques were aimed at trying to answer these questions. First we show our techniques only until June 2018, as the ASIC hardware only started shipping then. Later we show how the metrics we presented changed with ASICs.

A. Fraction of large miners in the mining power

Using the techniques from the papers on Zcash transaction linkability one can link most of the mining reward transactions. It is also relatively straightforward to approximate the mining power of an address based on the rewards it gets (how much value, in what time span and relative to the total Hash rate). This can be used to monitor the power of larger miners in the ecosystem (over 8KSol/s, Figure 12). We show this graph with some added information containing the exact fraction of large miners, and the daily exchange rate as well.

B. Mining Software Developer Fees

The most popular software for GPU mining is closed source, and has built-in developer fees. This is generally 2% of the mining rewards, and enforced by e.g. mining to the developers address for 72 seconds every hour (2% of an hour). Even though the developers usually try to obfuscate this address, we could find the fee addresses of all the major GPU software miners. By estimating the mining power of these addresses based on the previous approach, and then multiplying this power by 50 for the 2% rate[4], we can have an estimate on the mining power for GPU Equihash miners. The interesting and important point is that even if some other Equihash-based currency is mined with such software, the dev-fee is still sent to the Zcash mining pools. This is because the same software can be used to mine most of the Equihash-based coins.

We have produced this graph as well (Figure 13), where it shows about 80% rate for the first 6 months, and than reduces to about 60%, to later increase again to 80%. With this we can give an estimate that at least 80% of the Equihash mining power till May 2018 was provided by GPUs.

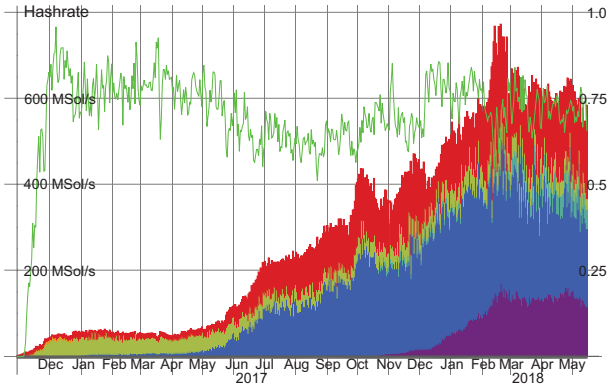

Fig. 13: Lower bound of GPU mining power based on the developer fees (Green: Claymore, Blue: EWBF, Purple: dstm, Light Blue: Bminer, Orange: Optiminer, Red: Remaining Hash rate). The green line presents the covered percentage.

From the month April 2017 there is an exponential difficulty increase (from about 50 MSol/s to 200 MSol/s) but only linear dev-fee fraction decrease from 0.8 to 0.6. ASIC or FPGA mining would have different effect – exponential dev-fee decrease. Explanation could be that large farms learned to disable dev-fee or (more likely) a good free miner has appeared or we did not find some extra dev-fee address. The exponential difficulty increase is most likely caused by the ZEC price hike.

We also see almost linear increase from July 2017 until March 2018 and in the last 3 months a slight linear decrease

[4]At the beginning various miners had different fees at the start from 2-15% but it seems they converged to 2% over time.

in Figure 9. However secret ASIC or FPGA linearly growing dominance should show as a linear decrease in the dev-fee ratio in Figure 13, which we do not observe.

CAVEAT: Dev-fee is an interesting metric, but if it is known to the adversary, it can be cheated by sending a fraction of the ASIC mining results as fees to dev addresses. Also we hope that developers of software miners do not run mining rigs pointed to the same addresses - this is unlikely but can not be completely ruled out.

C. Public Introduction of ASICs

At the start of June 2018 the first ASIC miners were shipped by the companies Innosilicon and Bitmain, which means the previous two methods can be used to inspect the data when it is known that there are ASICs in the network. As seen in Figure 15 the proportion of large miners have visibly increased and below 35KSol/s miners disappeared from the network.

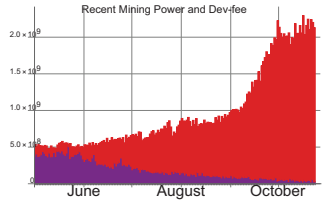

Fig. 14: The recent change in the projected mining power from dev-fees for the overall Equihash hash rate

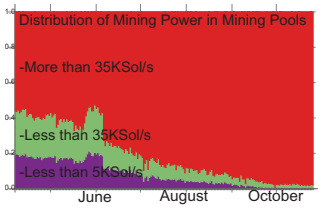

Fig. 15: The change in mining power distribution inside the mining pools.

The projected GPU mining power from developer fees have also exponentially decreased as seen in Figure 14, which is a predictable effect of ASIC dominance over the mining landscape and GPUs becoming unprofitable.

V. Mining Centralization

The decreasing projected power from dev-fees and the increasing overall hash rate lead to questions about the decentralization of mining power for Zcash as an effect of the introduction of ASICs. To have a more detailed view on the problem, we have investigated the attributes of mining power in mining pools in more detail in the Zcash and Equihash ecosystem.[5]

If we take a look at the graph of the proportion of large miners, the increase of large miners is obvious, while the actual number of average daily recorded miners that we have recorded went down from around 60,000 in the middle of April to around 13,000 by the end of November 2018 (Figure 16). It is visible that miners started leaving in the beginning of June 2018, and that tendency is still observable. This also shows how the introduction of ASICs might alienate GPU miners from a chain, resulting in a more centralized mining infrastructure. The picture has to be taken with a caveat, since theoretically a GPU miner with several mining rigs could point them to different mining addresses, though this is not very likely.

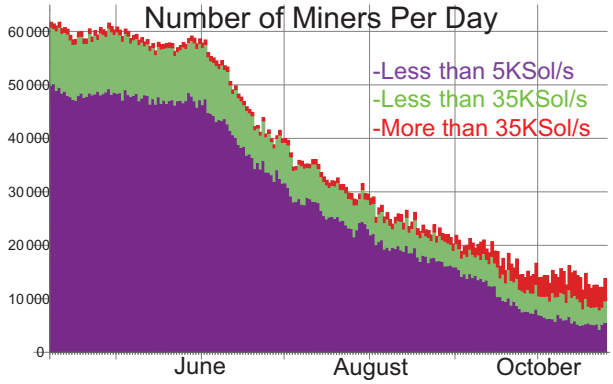

Fig. 16: Number of daily recorded miners in mining pools since April 2018. The number of small miners using probably only one rig has decreased from 50,000 to 5,000, while the middle portion has decreased from 10,000 to 4,000.

VI. Privacy of Miners

Miners might not want to reveal any information about themselves, especially if they are mining on a blockchain like Zcash, where privacy is the main feature. In that case mining on pools like Flypool or Nanopool is not advisable, as the information is publicly available for anybody who knows the miner's address. On the other hand we can still deduce most of the mining pool payouts from the public blockchain, which leads us to the question what can we learn about a miner if the only thing we see about them are their mining rewards on the blockchain?

A. Linkability of Mining rewards

Based only on blockchain information, we could record thousand of miners' approximate hash rates. This is based on three different metrics. The first one is the exact value of the reward, the second is the average chain hash rate since the last time that miner received a reward, and the third one is the number of blocks since that last reward payment. Combining this information we can estimate the mining power of that miner for that span of blocks. The more rewards the miner gets, the more precise our estimate can be. We do this analysis on example of GPU miners but it can be done for ASIC miners as well. First we have recorded all mining powers, and the ones under 3.5 KSol/s are displayed in the following histogram, where the width of a column is 5 Sol/s.

It is noticeable that the histogram follows a sum of independent Gaussian distributions, also called Gaussian mixture model (mining follows a Poisson distribution, but if the distribution's λ parameter is large enough, it can be approximated with a normal distribution). The obvious peaks are at 150, 270 and 550 Sol/s, etc. which correspond to popular mining cards or rigs. After this observation, one could manually or algorithmically fit a sum of Gaussians with different weights onto the histogram.

[5]The appearance of a close to 40% mining power in the beginning of June 2018 resulted in questions from the community about who could have such a large mining stake as an unidentified solo miner. Two weeks after its appearance its power was distributed into 2 separate addresses (later to even more addresses, while now it is under one address again), but probably it was controlled by the same entity all the time.

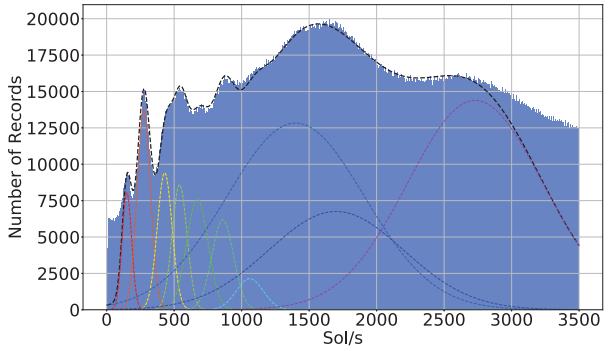

Fig. 17: Histogram of recorded hash rates in Zcash. The striped line is the Gaussian fitting. The separate colours are the single Gaussians.

Gaussian decomposition confirms results we observed earlier about popular cards/rigs in Table II. We consider this as a validation of our metrics, and it means that an attacker might be able to extract more information from the recorded hash rates, if all those rate are connected to the same miner. Based on these values an attacker can observe how stable a miner's hash rate is. If there is a significant increase of power, that stays the same for a long time, one can suspect that new hardware has been added to the miner's farm. As we have a general idea of the possible rigs and hardware, we can reduce the list of possible new hardware (Figure 18).

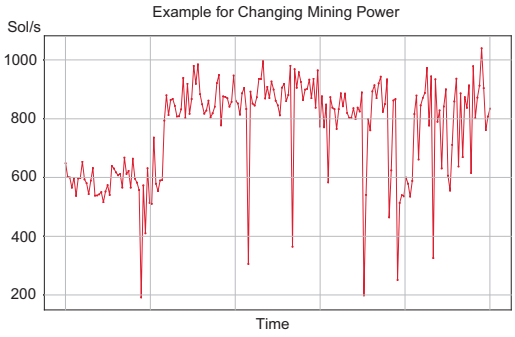

Fig. 18: Example for the case where a miner adds new hardware to its mining rig. In this case the power increases from around 600 Sol/s to around 850-900 Sol/s, which suggests that the miner added a GTX 1060 GPU.

From this information an attacker might deduce at what time a shipment of GPU or ASIC hardware was delivered to the miner. It might relate to geographical relations as well. For example if there is a GPU shortage, there might be information on when do different regions receive their new batches of specialized hardware. Correlating it with increases in mining power might reveal information on the geolocation of a miner. Vice-versa from increases in mining power one might deduce the batch schedules of hardware manufacturer.

B. Countermeasures

As the accuracy of an attacker's measurements depends on the regularity of rewards, we suggest using a large reward payment threshold, as the fluctuations in the global hash rate combined with the irregularity of payouts could generate enough noise such that the miner's power would be difficult to estimate correctly.

Next, even though instant payouts might sound as a user-friendly aspect for a mining pool, it results in even more accurate approximations. We also suggest frequent change of mining addresses, which would result in not enough data points for the attacker.

The drawback of constantly using new addresses is that when the miner wants to spend his/her coins, it should not do it in a single transaction from all of the addresses, because if all the addresses are inputs to the same transaction, then an attacker knows that all those addresses are controlled by the same entity. Instead, in the case of Zcash our suggestion is converting the coins to hidden z-addresses. In Ethereum one might consider mining to a smart contract instead of a direct account, and then withdrawing the coins from the contract. This approach would be still visible on the chain, but it is a bit harder to follow for an attacker, as it would involve analyzing the byte-code.

VII. Conclusions

In this paper we have presented an overview on decentralized GPU mining in cryptocurrencies. We have studied the most popular mining hardware, while also investigating the effect of introduction of ASICs in the mining ecosystem. We have provided methods that could be used to detect hidden ASIC farms in a network, and verified their effectiveness in practice. We have also shown how the overall effect of ASICs and reduced exchange rates damage the decentralization of mining power, leading to a disappearance of over 75% of the miners in Zcash.

Finally we have shown how using only blockchain information an attacker can learn the hash rate of a miner and might even deduce the structure of their mining rigs reducing their privacy. This study helps to expose privacy vulnerabilities in the current mining ecosystem which is crucial for privacy-preserving currencies and privacy-conscious users. It can also help to understand decentralization effects of GPU vs. ASIC hardware mining.

References

[1] S. Nakamoto, "Bitcoin: A peer-to-peer electronic cash system," 2008.
[2] I. Eyal and E. G. Sirer, "Majority is not enough: Bitcoin mining is vulnerable," *Communications of the ACM*, vol. 61, no. 7, pp. 95–102, 2018.
[3] I. Eyal, "The miner's dilemma," in *Security and Privacy (SP), 2015 IEEE Symposium on*. IEEE, 2015, pp. 89–103.
[4] I. Tsabary and I. Eyal, "The gap game," in *Proceedings of the 11th ACM International Systems and Storage Conference*, ser. SYSTOR '18. New York, NY, USA: ACM, 2018, pp. 132–132. [Online]. Available: http://doi.acm.org/10.1145/3211890.3211905
[5] G. Kappos, H. Yousaf, M. Maller, and S. Meiklejohn, "An empirical analysis of anonymity in zcash," in *27th USENIX Security Symposium (USENIX Security 18)*. USENIX Association, 2018.
[6] A. Biryukov and D. Feher, "Deanonymization of hidden transactions in zcash," 2018, https://cryptolux.org/images/d/d9/Zcash.pdf.