Blockchain VA

Simone Di Tanna 1765243, Francesco Colasante 1760608, and Giammarco D'Alessandro 1753102

Engineering in Computer Science, Sapienza University

June 9, 2021

Abstract

This paper aims to show a visual analytics system, in which the relations between the most important factors related to the PoW blockchains' energy consumption are presented: we talk about network hashrate and efficiency. In the following, mainly these two aspects of different blockchains will be explored; in particular, we will try to answer to some emblematic questions: what is the trend of the networks' hashrates? How is it increasing? How are partitioned the hashrates between the different blockchains? How are hashrate and efficiency related to energy consumption, and how much is the latter influenced by these two parameters? Are we going towards better or worse performances?

1 Introduction

Nowadays blockchain technology is widely known, for the larger part thanks to the cryptocurrencies' boom, and many discussions are carried on every day regarding the sustainability and the bonus/malus of this approach. However, not so often is carried on a deeper analysis about the factors underlying these concepts. Many people know, for example, that Bitcoin consumes a large amount of power. But why this happens? We can point out as the main guilties two attributes.

The first one is the hashrate: we can explain it as the total combined computational power that is being used to mine and process transactions on a Proof-of-Work blockchain. For those not familiar with this terminology, we can say that a "hash" is a fixed-length alphanumeric code that is used to represent words, messages and data of any length. Crypto projects use a variety of different hashing algorithms to create different types of hash code; think of them like random word generators where each algorithm is a different system for generating random words. To bring this concept into our context, it is fundamental to know that before new transactional data can be added to the next block in the chain, the so called "miners" must compete using their machines to solve a difficult mathematical problem. More specifically, miners are trying to produce a hash that is lower than or equal to the numeric value of the 'target' hash by changing a single value called a 'nonce'. Each time the nonce is changed, an entirely new hash is created. This is effectively like a lottery ticket system, where each new hash is a unique ticket with its own set of numbers.

Here is where our analysis fits: given the randomness of the process, it is straightforward to think that the more attempts in the shortest time we have, the higher is the probability to "win the lottery". This translates into the need for a very high number of operations executed, i.e., computational power, as we said previously. So we finally arrive to understand firstly the importance of hashrate, but efficiency nonetheless. In fact, the mining process also requires to minimize the costs derived from the energy consumption, exploiting the best hardware possible. This is expressed through the concept of efficiency, namely Joule divided hashes.

Application-specific integrated circuit (ASIC) mining hardware now dominates the crypto mining space and is solely designed to perform hashing functions. Some modern-day ASIC rigs are capable of achieving 110 tera-hashes per second, which equates to 110 trillion attempts at solving the hashing problem per second.

2 Dataset

The search for a dataset which would have worked for our analysis was hard, given that to make a sufficiently valid comparison we needed to consider different blockchains, different cryptocurrencies and different attributes. Finally we have found a way to build a dataset, consisting of a first part comprehensive of historical hashrates for several digital currencies, namely:

- Bitcoin (BTC)
- Ethereum (ETH)
- Litecoin (LTC)
- Monero (XMR)
- Bitcoin Cash (BCH)
- Bitcoin SV (BSV)
- Dash (DASH)
- Dogecoin (DOGE)
- Ethereum Classic (ETC)
- Vertcoin (VTC)
- Zcash (ZEC)

This part is based on data available on CoinMetrics, an open-source project to determine the economic significance of public blockchains, and it includes very different currencies, both by date of birth and by project. We focused our analysis in a specific amount of time, i.e. 2016-2020, and we aligned all the hashrates to the same unit of measure, Terahash/second (TH/s).
The second part of our dataset concerns more practically the mining for PoW blockchains, that we have seen involve the use of computationally performing hardware. We retrieved data about ASICs, in particular with the following attributes:

- *model*: the name of the hardware
- *release*: the release date of the hardware
- *hashRate*: the hashrate provided by the hardware (TH/s)
- *power*: the power consumption needed for the model (W)
- *algo*: the algorithm on which the model is based
- profitability: the estimated profitability reachable through the model ($/day)
- *efficiency*: the efficiency of the model (j/GH)

This data are composed by 159 different models, retrieved from ASIC miner value [1], a website providing an up-to-date estimate of the most profitable and efficient mining hardware over 27 algorithms, covering more than 200 crypto currencies. We focused on the data regarding the hardware suitable to compute the consensus challenges in each of the crypto we chose to analyze, i.e. the ASIC used for the following algorithms: SHA-256, ethash, equihash, cryptonight, eaglesong, scrypt and X11. Note that there are less algorithms (7) than crypto (11) in our analysis, and that is because different blockchain system may exploit the same mechanism to reach consensus (pretty common with crypto currencies), or different versions of it for which the same hardware is still valid or easy to adapt (e.g. both Litecoin (LTC) and Dogecoin (DOGE) rely on a Scrypt-based algorithm).

The last part of our dataset is about the electricity net consumption (measured in tera Watt per year, TWh/y) of 214 world countries, covering almost 40 years (from 1980 to 2019). We took this data from the federal U.S. institution for Energy Information Administration (EIA) that collects, analyzes, and disseminates independent and impartial energy information to promote public understanding of energy and its interaction with the economy and the environment [2]. We used this data to compare real countries energy consumption to those of the bolckchain systems we are considering. In particular in our analysis cryptos electricity consumption have been estimated starting from the ASIC efficiency and the hashrate data we have collected, not relying on assumptions on electricity cost, and its distribution, as in previous works [4].

3 Visualizations

The visualization tool fits in two different pages, namely *Asic* and *Hashrate*: the first one is composed by 4 graphs, while the second one is composed by 3 graphs, for a total of 7 different graphs.
Talking about the *Asic* page, the 4 graphs are coordinated and interactive. The top bar provides a selectable menu from which the user can choose different algorithms, differently color coded, corresponding to different cryptocurrencies; the page defaults to all of them selected. In the *Hashrate* page, just the graphs on the left are coordinated (fifth and sixth visualizations), while the last one is animated by definition. Now we will explore all the graphs more in detail, providing a description of what we want to represent and how.

3.1 Vis 1: Scatterplot

This scatterplot graph displays the hashrate of the ASIC with respect to the release date. In addition, the color shows the type of algorithm supported by the relative asic.

In addition, the blue line shows the growth trend of the hashrate with respect to the release date. The trend is calculated using a quadratic regression and is recalculated at runtime, based on the types of algorithms selected. Individually selected each algorithm, it is possible to analyze the trend of each type of algorithm, useful for predicting future hashrate values as well.

It is possible to dynamically calculate the average efficiency of a certain group of asics, by brushing on a certain area. The average efficiency (j/GH) will be shown dynamically to the right of the graph.

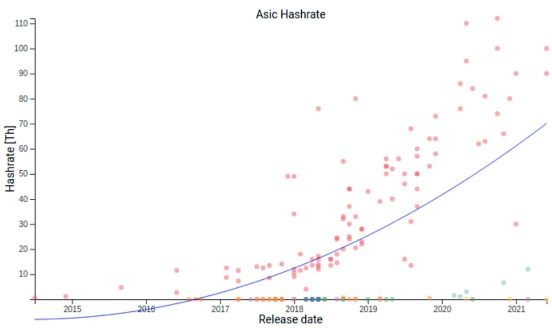

Figure 1: Asic scatterplot

3.2 Vis 2: Parallel coordinates

This parallel coordinates visualization of ASICs shows the characteristics of each hardware in a very explanatory way, since we can get at first glance the best values and understand if there is some particular distribution of the instances we are considering. In particular, the graph shows the attributes described in the previous chapter: *efficiency, hashrate, profitability, power, release, algorithm*. The graph will be filtered according to the user interactions on the first and third visualizations.

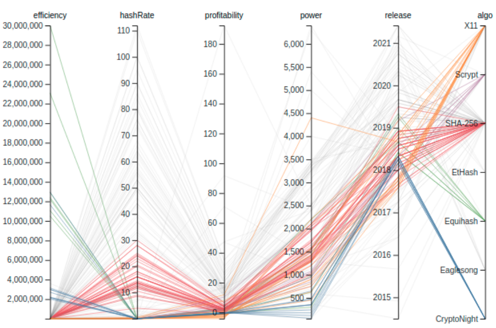

Figure 2: Parallel coordinates

3.3 Vis 3: PCA

We want to visualize data in a way that proximity reveals similarity between data objects.To do this, we used a multidimensionality reduction computing Principal Component Analysis technique and taking the first two components with greater variance.

The graph maintains the same color coding as the previous ones. In addition, this plot is coordinated in both ways and with the scatterplot. By brashing the scatterplot, the unselected asics are shown in this graph in gray, otherwise brushing on this graph, the selected asics are shown with a blue border on the scatterplot. This feature is convenient for understanding the similarities between different groups of asics, also observing their hashrate.

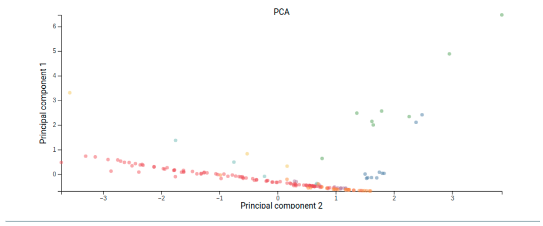

Figure 3: PCA

3.4 Vis 4: Table

This visualization provides a schematic representation of ASICs and their attributes, comprehensive of model name and release date. The table is sortable, in ascending or descending order, according to the desired attribute. The table rows are filtered according to the user interactions on the first visualization.

3.5 Vis 5: Time-series line chart

This graph shows us the amount of hashrate load (TH/s) for each cryptocurrency we analyzed, from 2016 to 2020. Data are represented with a logarithmic scale, given the huge impact Bitcoin has on the data. On the right, we have the cryptocurrencies, selectable by the user, to make comparisons on different lines. The graph is also brushable, to select a particular time frame. Lastly,T selection triggers the sixth visualization.

3.6 Vis 6: Stacked area

This chart shows the hashrate ratio of the different cryptocurrencies, using the same color coding as the previous chart. This chart is triggered and dynamically updated based on the selection in Vis 5.

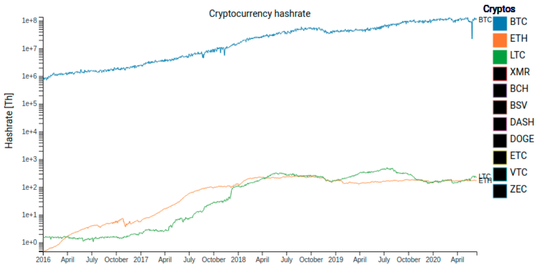

Figure 4: Time-series line chart

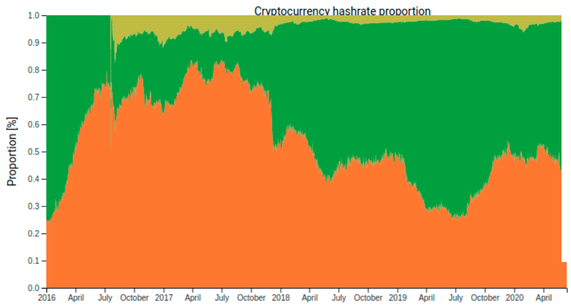

Figure 5: Stacked area

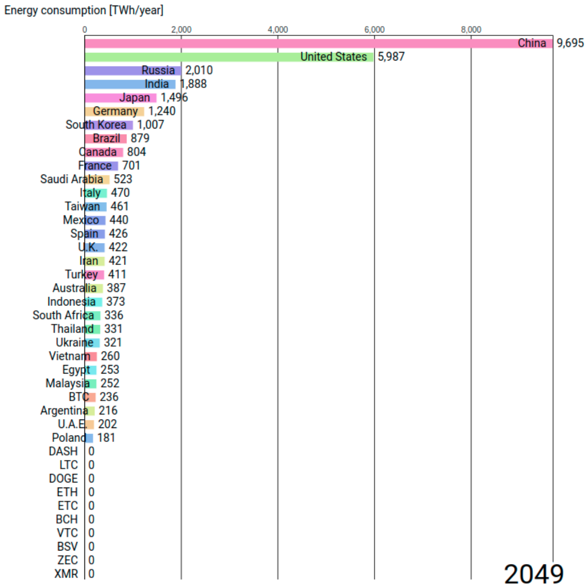

Figure 6: Bar race

3.7 Vis 7: Bar race

This graph shows the trend of the energy consumption of the countries of the world, compared with cryptocurrencies. The graph dynamically shows the energy consumption starting from the year 2016, and with a prediction, it reaches up to the year 2049.

4 Related works

About the positioning in the literature of our work, only a few others are visual analytics oriented. Anyway, given the popularity of the arguments nowadays, many papers exists with similar purposes. Since we started our analysis driven by the desire to better understand the dynamics underlying the energy consumption of various public blockchains, we must cite firstly the work of Vranken [7]. This makes us understand that "the ASICs that are currently being used by bitcoin miners, are most likely a mix of the newest available and some older ASICs. The actual mix used in practice is unknown. Bitcoin miners will not switch to newer hardware as long as mining with their current hardware is still profitable and the break-even point has not been reached yet at which revenues have covered the capital and operational expenditure of their current hardware", and also it confirms that, talking of Bitcoin, "[...] will be mined by those who can do it most cheaply, and others will be put out of business. It is therefore likely that surviving miners run the latest hardware at locations offering the lowest electricity costs to be competitive and to maximize profit". This alone would be enough to affirm the importance of energy during the mining process, and in particular the reachable profitability given the own efficiency.

This leads us to the sustainability problem, discussed in a Nature paper [5], where it is affirmed, through an estimation of the period 2016-2018, that mining of just Bitcoin, Ethereum, Litecoin and Monero "is responsible for at least 3–15 million tonnes CO_2 emitted, with Bitcoin being the largest contributor". In the same paper, it is also reported a comparison comprehensive of hashrates and prices time series of the discussed cryptocurrencies, similarly to our work.

About that, also Fantazzini and Kolodin [3] analyzed a possible relation between hashrate and price, concluding that "there was neither evidence of Granger-causality nor cointegration in the first examined sample (01/08/2016–04/12/2017), whereas there was evidence of unidirectional Granger-causality and cointegration in the second sample (11/12/2017–24/02/2020), going from the bitcoin price to the hashrate (or to the CPMs) but not vice versa".

Anyway, many discussions lead us to think that we should separate the hashrate/price relation from the hashrate/energy relation, since it seems that the latter has no reliability. As said in Song-Aste paper [10], "the ratio between mining cost and total transaction volume has not increased nor decreased over the last 10

years despite Bitcoin mining activity having increased by ten billion times during the same period".
Hence, we can say that the first part talking about hashrate is widely discussed in the literature, but a remarkable analysis merging also the hardware and energy arguments, i.e. also talking about ASICs and efficiency, has been conducted by few people. About it, we must cite the huge work of Gallersdörfer, Klaaßen and Stoll [6], which have done a very large analysis taking into account, for each cryptocurrency, the algorithm, the hashrate, the efficiency, the hardware involved and also the market cap. Thanks to their work they have drawn some important conclusions: firstly, "that currencies with ASIC-resistant algorithms consume an overproportionate amount of energy in relation to their market capitalization", but as previously said, a crucial problem for this type of analysis is that all estimates we can do about energy are subject to a large uncertainty, since "the selections and operation of the mining devices pose a significant challenge given that the mining industry operates secretively. Miners may shut down and ramp up certain devices temporarily as a response to variations in electricity prices and market prices (i.e., when electricity costs exceed mining revenues, as seen during coronavirus pandemic when market prices and hash rates tumbled)".
However, also taking into account this secrecy about mining devices, in [10] we can observe that "presently at least a billion USD per year is burned by the Bitcoin network for the proof of work. This amount corresponds to a one million times increase with respect to the costs in 2010. However, although large, this amount is $<0.5\%$ of the transaction volume over the network during the same period".
Finally, the best visual analytics work in this field were carried on by the Digiconomist [4], which provides the latest estimate of the total energy consumption of the Bitcoin network. Their work is really impressive, since they try to give a direct view on the Bitcoin energy consumption, using practical approaches and providing useful examples to a common user which would know about this phenom. The conclusions are really near to ours: Bitcoin energy consumption increased during years and probably will continue to do it. The difference is that we focused also on analyzing the basis of this phenomenon (namely hashrate and efficiency) and hence trying to go deeper through the causes.
Another great work which also gives us the idea of unstoppable growing of Bitcoin is the one conducted by De Vries [8], in which, in 2018, he stated "the average electricity consumed per transaction equals at least 300 kWh, and could exceed 900 kWh per transaction by the end of 2018", and nowadays we can relate this estimation with the actual data, which we found also during our work, in 2021: a single Bitcoin transaction has come to cost about 1500 kWh.
Suddenly, again De Vries [9] came back on the topic, updating his previous work: using time series the author shows us the trend of Bitcoin's network hashrate, highlighting the different pashes of the market. With a new estimation, he finally arrived to conclude that, until 2019, "the Bitcoin network to consume 87.1 TWh electrical energy annually", which "represents close to half of the current global data centre electricity use (200 TWh), while equaling the electricity use a country like Belgium (87.9 TWh)".

References

[1] *ASIC Miner Value*, live income estimation of all known asic miners. *https://www.asicminervalue.com/*.

[2] *U.S. EIA*, energy information administration, independent statistics & analysis. *https://www.eia.gov/*.

[3] N. K. Dean Fantazzini. Does the hashrate affect the bitcoin price? *Journal of Risk and Financial Management*, 2020.

[4] Digiconomist. Bitcoin energy consumption index. *https://digiconomist.net/bitcoin-energy-consumption*.

[5] M. J. Krause and T. Tolaymat. Quantification of energy and carbon costs for mining cryptocurrencies. *Nature*, 2018.

[6] C. S. Ulrich Gallersdörfer, Lena Klaaßen. Energy consumption of cryptocurrencies beyond bitcoin. *Joule*, 2020.

[7] H. Vranken. Sustainability of bitcoin and blockchains. *Science Direct*, 2017.

[8] A. D. Vries. Bitcoin's growing energy problem. *Joule*, 2018.

[9] A. D. Vries. Bitcoin's energy consumption is underestimated: A market dynamics approach. *Science Direct*, 2020.

[10] T. A. Yo-Der Song. The cost of bitcoin mining has never really increased. *Frontiers in Blockchain*, 2020.