Article

Application of Benford's law on cryptocurrencies

Jernej Vičič [1,3,†,‡] (ID), **Aleksandar Tošić** [2,4,†,‡] (ID)

[1]   University of Primorska Faculty of Mathematics, Natural Sciences and Information Technologies; jernej.vicic@upr.si

[2]   University of Primorska Faculty of Mathematics, Natural Sciences and Information Technologies; aleksandar.tosic@upr.si

[3]   Research Centre of the Slovenian Academy of Sciences and Arts, The Fran Ramovš Institute;

[4]   InnoRenew CoE

[*]   Correspondence: jernej.vicic@upr.si;

[†]   Current address: Glagoljaška 8, 6000 Koper, Slovenia

[‡]   These authors contributed equally to this work.

Abstract: Blockchain-based currencies or cryptocurrencies have become a global phenomenon known to most people as a disruptive technology, and a new investment vehicle. However, due to their decentralized nature, regulating these markets has presented regulators with difficulties in finding a balance between nurturing innovation, and protecting consumers. The growing concerns about illicit activity have forced regulators to seek new ways of detecting, analyzing, and ultimately policing public blockchain transactions. Extensive research on machine learning, and transaction graph analysis algorithms has been done to track suspicious behaviour. However, having a macro view of a public ledger is equally important before pursuing a more fine-grained analysis. Benford's law, the law of first digit, has been extensively used as a tool to discover accountant frauds (many other use cases exist). The basic motivation that drove our research presented in this paper was to test he applicability of the well established method to a new domain, in this case the identification of anomalous behavior using Benford's law conformity test to the cryptocurrency domain. The research focused on transaction values in all major cryptocurrencies. A suitable time-period was identified that was long enough to sport sufficiently large number of observations for Benford's law conformity tests and was also situated long enough in the past so that the anomalies were identified and well documented. The results show that most of the cryptocurrencies that did not conform to Benford's law had well documented anomalous incidents, the first digits of aggregated transaction values of all well known cryptocurrency projects were conforming to Benford's law. Thus the proposed method is applicable to the new domain.

Keywords: cryptocurrency, Benford's law, anomaly detection, method application

1. Introduction

Benford's law [1], also known as the first-digit law, has been widely used as a tool to discover anomalies in various data ranging from accounting fraud detection, stock prices, house prices to electricity bills, population numbers, natural phenomena, death rates and recently so popular COVID-19 cases reports. Cryptocurrencies, also referred as Blockchain-based currencies or cryptocurrencies or crypto coins, throughout the paper we will rely on the definition presented by [2] (cryptocurrency), have become a global phenomenon known to most people. A cryptocurrency is in fact quite narrow, albeit recognizable description of a subset of an umbrella class of cryptoassets. While still somehow geeky and not understood by most people, banks, governments and many companies are aware of its importance. Although there have past more than 8 years from the first transaction of the first cryptocurrency - Bitcoin (BTC) [3], only the last 5 years have seen a big enough number of transactions and a large time-frame that some statistical analysis can be carried out. Our research focused on empirical proof whether Benford law [1], a law of anomalous numbers, could be

used in a non-altered form for discovering fraudulent or at least suspicious activity on cryptocurrencies in the same way it is used in standard financial forensics.

Although we could observe the cryptocurrency transactions as just another financial tool that should comply to all the used mechanisms (among them also the Benford law conformity for identifying frauds and other anomalous behavior), there are some properties that must be addressed or at least be observed:

- mining transactions (mostly with mining pools) for all cryptocurrency assets that base on the Proof of work (PoW) [4] consensus mechanism by which the cryptocurrency blockchain network achieves distributed consensus. Mining pools, where most of the miners are concentrated, pay out rewards to miners based on the computing power contributed. The payouts are mostly scheduled to occur once the miner is owed more then the threshold to save up on transaction fees. As many miners keep the default threshold, many transactions are possibly of the same value.
- default transaction fees (GAS[1]) are the same.

Our research experiment started with gathering all transactions on the Ethereum (ETH) network. Ethereum was chosen for these properties: One of the biggest cryptocurrencies by market capitalization, and number of transactions processed, the network houses multiple cryptocurrencies (tokens) that could be compared directly (this part of the experiment is still open), well documented and accessible blockchain. The first preliminary results revealed that transaction values (non aggregated) of the whole Ethereum network do not conform to Benford's law [1] as is presented in Figure 1. Blue color depicts leading digits that conform to Benford's law, red color depicts non-conforming digits. The reasoning is further discussed in Section 4. Although this does not mean that there was any artificial manipulation or any other kind of anomaly, we investigated further. According to [5] the Benford's law metric can be used to achieve similar goals on aggregated data. We explored the same phenomenon on aggregated values (number of transactions in an observed period, aggregated transaction values, ...). Most of the aggregated values conform to Benford's law according to goodness of fit chi square ($\tilde{\chi}^2$) test [6] which in most literature like [7] is considered as a suitable tool to test Benford's law conformity. We extended our research to all major cryptocurrencies with enough transactions in the selected time-period.

The basic idea of the research was to test if Benford's law conformity can be used as a tool to detect anomalies in cryptocurrencies. The methodology is further presented in Section 4.
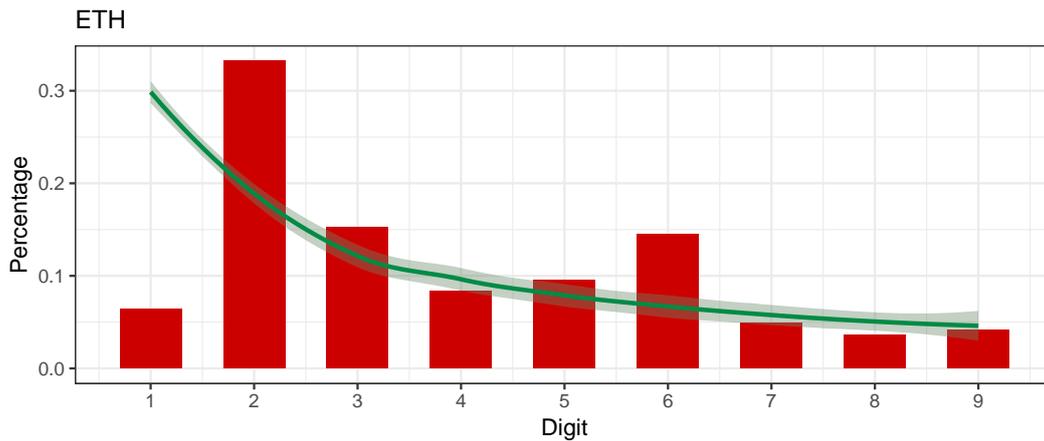
2. Benford's law

The Benford's law, also called the Newcomb–Benford law or the first-digit law is an observation about the frequency distribution of leading digits. The observation was first discovered by [8] and later rediscovered by [1]. The Benford's law defines a fixed probability distribution for leading digits of any kind of numeric data with the following properties [9]:

- Data with values from several distributions.
- Data that has a wide variety in the number of digits (e.g., data with plenty of values in the hundreds, thousands, tens of thousands, etc.)
- Data set is fairly large, as a rule of a thumb at least $50 - 100$ observations [10], although usually thousands of observations.
- Data is right skewed (i.e., the mean is greater than the median), and the distribution has a long right-tail rather than being symmetric.
- Data has no predefined maximum or minimum value (with the exception of a zero minimum).

---

[1]   Gas refers to the pricing value required to successfully conduct a transaction or execute a contract on the Ethereum blockchain platform.

Figure 1. The leading digits of all ETH transaction values do not conform to Benford's law. The daily aggregated values conform to the same metric (see Figure 2) leading to possible conclusion that there are too many automatic transactions in the network, but the aggregated values avoid this effect.
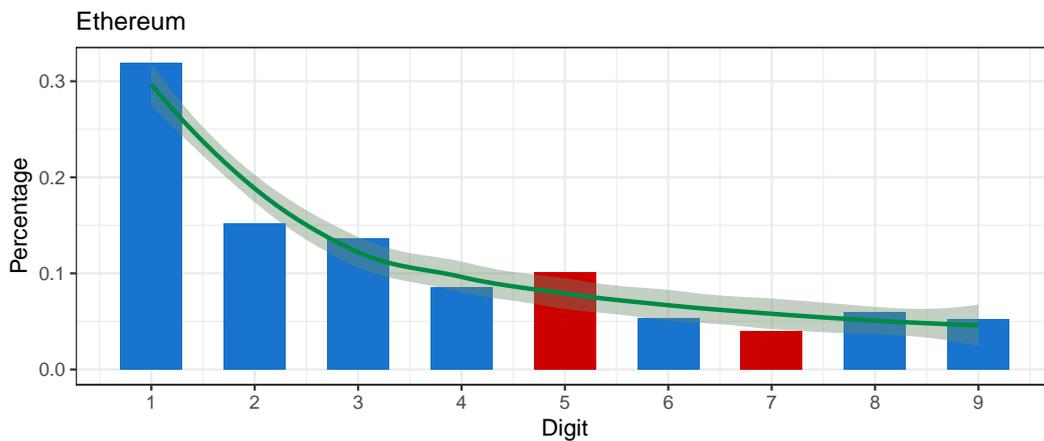
Figure 2. The leading digits of daily aggregated ETH transaction values in USD conform to Benford's law. Blue colored bars represent digits that conform and red colored bars represent digits that do not conform to Benford's law.
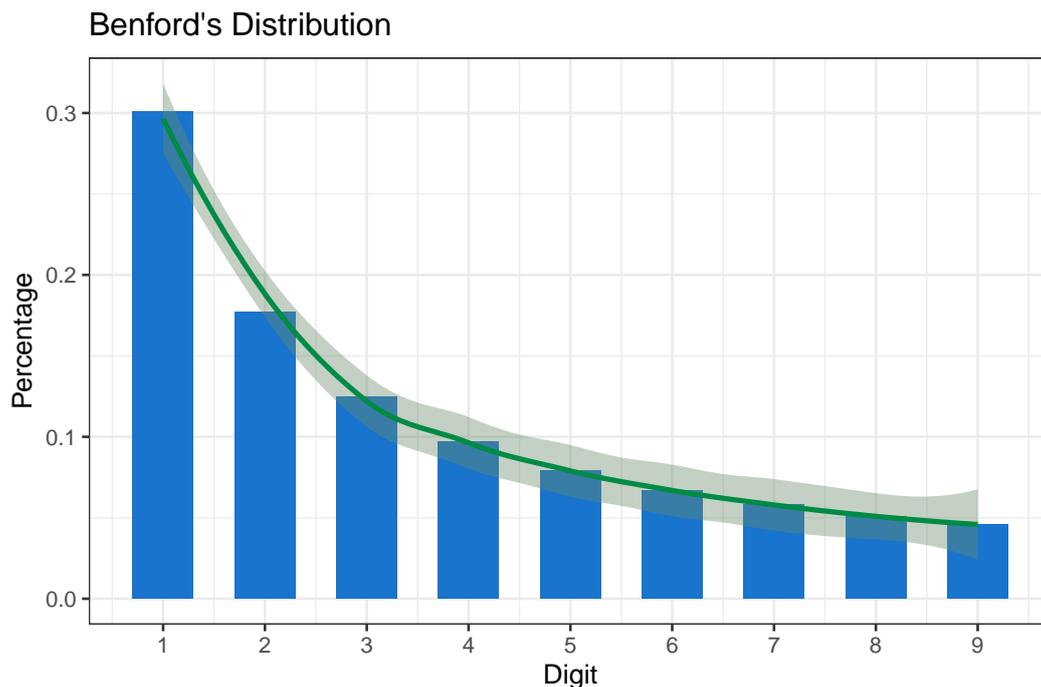
The distribution of digits is presented in Figure 3; the digit 1 occurs in roughly 30 % of the cases, and the other digits follow in a logarithmic curve. It has been shown that this result applies to a wide variety of data sets [9], some examples are presented in Section 3. The equation for the distribution of the first digits of observed data is presented in Equation 1.

$$P(d) = log_{10}(d+1) - log_{10}(d) = log_{10}(1 + \frac{1}{d}) \tag{1}$$

3. State of the art

Benford's law has been thoroughly researched and its theoretical grounds have been proved in many scientific papers. The methodology and basic mathematical grounds are discussed in greater detail by [11]. Many researchers have verified for themselves that the law is widely obeyed but have also noted that the popular explanations are not completely satisfying [12]. To the authors' knowledge, there has been no research in using Benford's law as a tool for detection of anomalies in cryptocurrency transactions.

Benford's law has been extensively used in the accountant fraud detection and prevention, there has been a lot of research in the area, such as [13]. [14] presents a literature overview of the area. [15]

Benford's Distribution

Figure 3. The distribution of digits in accordance to Benford's law [9]. Blue colored bars represent digits that conform to Benford's law.

introduces Benford's Law and Digital Analysis (analysis of digit and number patterns of a data set), which can be used as an analytical procedure and fraud detection tool. [7] presents Benford's law as a simple and effective tool for the detection of fraud. The purpose of the paper is to assist auditors in the most effective use of digital analysis based on Benford's law by identifying data sets which can be expected to follow Benford's distribution, and presenting types of frauds that would be "detected/not detected" by such analysis. However, there are some research findings that point out some inherent problems that potentially arise in the use of the Benford's law in the auditing process such as [16].

The simplicity of the Benford's law as a tool allows for a broad range of uses. [17] examined crime statistics at the USA National, State, and local level in order to test the conformity to the Benford's law distribution. [18] observe the distribution of initial digits of physical constants, their results are inconclusive, though.

One of the more recent researches involving Benford's law is [19]. Authors propose a test of the reported number of cases of coronavirus disease in 2019 in China with Benford's law and report that the reported numbers of affected people abide to Benford's law.

[20] presents and overview of identified frauds that can be committed in the cryptocurency paradigm, identified frauds include Ponzi schemes [21], fake initial coin offering schemes, pump and dump schemes as well as cryptocurrency theft. [22] identify the main reasons for frauds and manipulation in cryptocurrencies: lack of consistent regulation, relative anonymity, low barriers of entry, exchange standards and sophistication. [23] perform an end-to-end characterization of the counterfeit token in the Ethereum network, targeting Erc20 coins. [24] aims to demonstrate that Bitcoin, the most known cryptocurrency, constitutes a substantial danger in terms of criminal enterprise. [25] present an economic analysis of money laundering schemes utilising cryptocurrencies, it aims at providing an answer to the open question, whether cryptocurrencies constitute a driver for money laundering. [26] propose an approach to detect illicit accounts on the Ethereum blockchain using well proven machine learning techniques.

4. Methodology

As mentioned in Section 1, this paper proposes a methodology for identifying out-of-the-ordinary behavior and possibly detect frauds in blockchain based currency. As such, the purpose is to present scientific grounds that allow feasibility and usefulness of the method as well as to propose a set of usage guidelines and a use case where our hypotheses were confirmed. The observation sets need to conform to all the basic prerequisites for Benford's law as described in Section 2. This is the agenda for the executed research:

- Take all major cryptocurrencies into consideration.
- Express all aggregated daily transactions in one currency – we selected USD ($) as the most used FIAT currency in comparisons.
- Select a viable observation period.

  – Starting date for each currency was the date of the first successful transaction,
  – Ending date for the observation period was set long enough into the past so that the frauds or abnormal behavior were well documented (in the forms of lawsuits, scandals, vanished crypto currencies, well documented special properties of specific currencies). We selected the end of 2018 year, almost three years in the past.
  – A long enough observation period that makes Benford's law conformity observation feasible (as presented in Section 2). In the body of surveyed literature, the sample size varies from 200 [27] to a few hundred thousand. We opted for doubling the minimum sample size – selecting all crypto currencies with 400 or more transaction days.
  – Perform the MAD test [28] and classify all the cryptocurrencies according to [29] and visually observe all conformity graphs.
  – Perform a literature review for all the currencies that do not conform to Benford's and establish if there are any abnormalities documented for the selected time frame.

Testing conformity to Benford's law distribution has been done with many goodness of fit tests ranging from Pearson's Chi squared [30], Kolmogorov-Smirnov D statistics [31], Freedman's modification of Watson U^2 statistics [32], euclidean distance d statistics, and many others. However, no real data will ever follow the exact distribution; hence, most analysis supplements statistical testing with graphical representations that help in pointing out suspicious patterns in the data for further investigation. Additionally, different tests have different reactions on sample sizes. The Chi square test suffers from an excess power problem in that when the number of observations becomes large (above 5000 records estimated by [29]) it becomes more sensitive to insignificant spikes, leading to the conclusion that the data does not conform. [33] suggested some statistical tests can render misleading results when applied to large number of observations. On the other hand, [34] conclude that Mean Absolute Deviation MAD test [28] is reliable with as low as 200 observations (as additional safety measure, we opted doubling that value to 400 in our experiment). [35] proposed the Mantissa Arc test, which is a very interesting geometrical test. Unfortunately, it tolerates little deviation from Benford's distributions.

[29] concluded that the best test is Mean Absolute Deviation (MAD) and a lot of the state-of-the art literature agrees with this proposal. [29] also presents a list pf thresholds to classify the observed conformity:

- conformity *(0.000)*,
- acceptable conformity *(0.006)*,
- marginally acceptable conformity *(0.012)*,
- and nonconformity *(0.015 and above)*.

The adapted MAD is used to measure the average deviation between the heights of the bars and the Benford line. The higher the MAD, the lower the conformity. We opted to perform conformity tests using all three of the aforementioned tests as our sample sizes are well within the acceptable ranges.

All presented statistical tests are also supplemented with graphical representations; the results are presented in Section 5.

Basic recipe: Select a big enough set of aggregated data that conforms to Benford's law prerequisites described in Section 2. Count leading digits and perform Mean Absolute Deviation (MAD) conformity [29] on the gathered data. Plot simple bar charts with the numbers for each leading digit and visually and manually observe the distribution. If the data does not conform to Benford's law, investigate further.

4.1. Data

DataHub cryptocurrency datasets [2] hosts daily aggregated data about all transactions on all crypto coin networks from the first mined block on the Bitcoin network till the end of the 2018. As such it presents the perfect data source for our research. The problem that arises is how to get more recent data. The problem is further discussed in Section 7.

5. Results

This section presents the results of the experiment following the methodology from Section 4. All the figures in this section have the same format: a graph showing the distribution of leading digits. Red colored bars represent suspect values, which skew the distribution the most. Suspects are classified where mean absolute deviation is above the threshold of 4. The threshold can be adjusted to increase the sensitivity. Suspects are useful as a starting point for further investigation in case of nonconformity.

The time interval selected was between 2009 and 2018. Most of the cryptocurrencies were in early development phase without a use-case or product, consequently the amount of transactions recorded was negligible. Table 1 presents all cryptocurrencies that conformed to the prerequisites presented in Sections 2 and 4. Each cryptocurrency is presented by its name and the ticker, number of observations (equal to the number of days) and all the values from the Benford's law conformance test. The currencies were grouped into 4 groups according to [29] and also sorted according to this grouping from best to worst conformance.

Two "best conforming" cryptocurrencies, Ethereum classic (ETC) and Vertcoin (VTC), both still respectable projects, were classified as "Close conformity". Two biggest blockchain platforms by market capitalization, Bitcoin (BTC) and Ethereum (ETH), were classified as "Acceptable conformity" and "Marginally acceptable conformity" respectively. Figure 4 shows the Benford's law conformance chart for further visual examination for all four cryptocurrencies.

Six of the currencies from the Table 1 were classified as "non-conformant" to Benford's law: EOS (EOS), TENX token (TENX), Veritaseum (VERI), Basic Atention Token (BAT), PIVX (PIVX), Dogecoin (DOGE). Each of the cryptocurrencies from this list will be presented and discussed.

5.1. TENX token (TENX)

Figure 5 shows the TENX aggregated transactions and the conformance to Benford's law. The MAD value A well documented Wirecard scandal [3] shows a possible reason for non-conformity.

5.2. Veritaseum (VERI)

Figure 6 shows the VERI aggregated transactions and the conformance to Benford's law.

The U.S. Securities and Exchange Commission (SEC) said it has reached a settlement with Reggie Middleton, organizer of the fraught \$14.8 million Veritaseum (VERI) initial coin offering (ICO) [4]. The

---

[2]   DataHub cryptocurrency datasets: https://datahub.io/cryptocurrency
[3]   Crypto.com, TenX crypto debit cards frozen following Wirecard scandal: https://decrypt.co/33695/crypto-debit-cards-frozen-following-wirecard-scandal
[4]   Analysis of the Veritaseum Scam: https://steemit.com/money/@financialcritic/analysis-of-the-veritaseum-scam
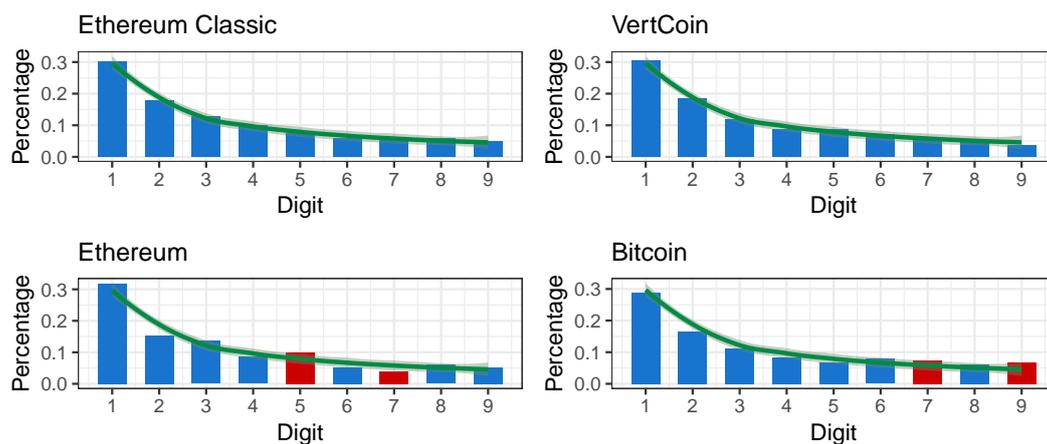
Figure 4. Two best conforming (ETC) and (VTC) currencies with "Close conformity" and two biggest crypto currencies (BTC) – "Acceptable conformity" and (ETH) – "Marginally acceptable conformity" for aggregated value in USD transaction history.
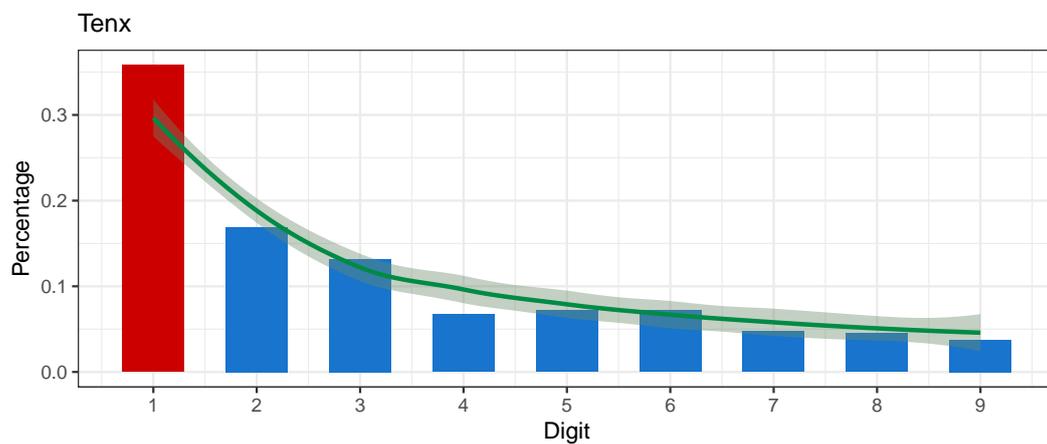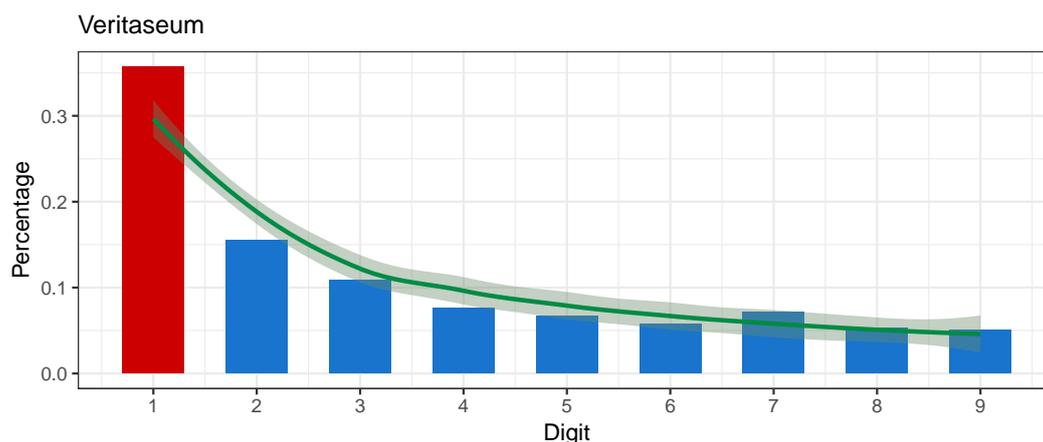

Figure 5. TENX aggregated transactions and the conformance to Benford's law. Digit 1 overflows, digit 4 (almost) underflows. Overall the daily aggregated transaction values do not conform.

case was closed on October 2019, but the frauds were committed well within the observation period of our research.

Currency	Obs.	Pearson's Chi-squared test		Mantissa Arc Test		MAD	MAD Conformity	Distortion Factor
		X-squared	P-value	L2	P-value			
Ethereum Classic (ETC)	750	1.766027	0.9873638	8.61E-05	0.9374726	0.00351481	Close	-0.1409321
Vertcoin (VTC)	1666	7.30948	0.5036398	0.000165673	0.7588044	0.005795195	Close	-1.621333
Metal (MTL)	400	5.15115	0.7413057	0.001522525	0.543889	0.01089584	Acceptable	-0.1257945
Status (SNT)	411	7.692396	0.4640798	0.001050824	0.6492816	0.01005221	Acceptable	-1.560401
Aragon (ANT)	452	5.696092	0.6812311	0.005388913	0.08752867	0.01078389	Acceptable	3.448495
Waves (WAVES)	603	5.14964	0.7414692	0.003501951	0.1210349	0.008216651	Acceptable	-1.721726
Iconomi (ICN)	658	10.17673	0.2528404	0.0008252317	0.5810012	0.0104604	Acceptable	0.8235436
NEO (NEO)	665	3.823118	0.8727192	0.0008927334	0.5522979	0.006478303	Acceptable	1.316035
Lisk (LSK)	811	11.45478	0.1772377	0.001803885	0.231552	0.009606102	Acceptable	3.072645
Stellar (XLM)	1009	9.622045	0.2925614	0.002200075	0.1086226	0.007992198	Acceptable	1.221268
Verge (XVG)	1387	8.300241	0.4047048	0.002115592	0.05316656	0.007575786	Acceptable	-2.84182
MaidSafeCoin (MAID)	1560	10.43771	0.2356377	0.003279288	0.006001835	0.007513696	Acceptable	3.73407
Dash (DASH)	1641	5.958045	0.6519316	0.001418531	0.09750916	0.00621291	Acceptable	0.8615983
DigiByte (DGB)	1649	2.59E+01	1.11E-03	3.21E-03	5.00E-03	0.01088511	Acceptable	-2.4136
Bitcoin (BTC)	1933	30.8193	0.0001512958	0.0006696828	0.2740357	0.01158613	Acceptable	5.881506
Gnosis (GNO)	468	8.754344	0.3634412	0.006937894	0.03889326	0.01312756	Marginally acc.	1.135551
Golem (GLM)	633	11.07461	0.1975074	0.003690431	0.0967096	0.0129236	Marginally acc.	6.131378
Zvash (ZEC)	653	20.82315	0.007632357	0.001029657	0.5104994	0.01293599	Marginally acc.	-0.9372237
Decred (DCR)	915	17.6832	0.02373108	0.0005975181	0.5788401	0.01375337	Marginally acc.	-1.586765
Ethereum (ETH)	1102	25.77399	1.15E-03	3.78E-04	0.658996	0.01482756	Marginally acc.	-0.08431323
NEM (XEM)	1230	27.13364	0.0006703807	0.008295528	3.70E-05	0.01417723	Marginally acc.	3.19854
Tether (USDT)	1258	34.91683	2.77E-05	1.38E-02	2.82E-08	0.01391653	Marginally acc.	-5.969747
EOS (EOS)	401	15.36398	0.05244271	0.003494984	0.2462301	0.0200535	Nonconformity	-2.819878
TENX token (TENX)	402	1.05E+01	2.34E-01	8.08E-03	3.89E-02	0.01539412	Nonconformity	-7.119347
Veritaseum (VERI)	431	11.32151	0.1841391	0.01211339	0.005402612	0.01726905	Nonconformity	-1.603899
Basic Atention T. (BAT)	438	19.05523	0.01456707	0.01293943	0.003456598	0.02196946	Nonconformity	0.2319942
PIVX (PIVX)	903	28.08438	0.0004584671	0.01199764	1.97E-05	0.01890993	Nonconformity	-7.031687
Dogecoin (DOGE)	1702	83.1755	1.12E-14	0.02422157	1.25E-18	0.0214206	Nonconformity	-9.527495

Table 1. Conformity tests for all major cryptocurrencies in the observed time-period with more than 400 days of transactions on the blockchain. The records are sorted according to MAD Conformity column, from close conforming to nonconforming.

Figure 6. VERI aggregated transactions and the conformance to Benford's law. Digit 1 overflows. Overall the daily aggregated transaction values do not conform.

5.3. Dogecoin (DOGE)

Figure 7 shows the DOGE aggregated transactions and the conformance to Benford's law. The coin was introduced as a satire initially in December 2013 and included an image of the Doge meme as its logo. The author of this coin/crypto currency revealed this motivation publicly. Some properties showing the soundness of our decision:

- On the 24.9.2018 (a randomly chosen date on a working day at the end of our observation period): last tweet from the official Tweeter account on 14. July 2018 (80 days)[5],
- fun and friendly internet currency, dogecoin logo is a dog from meme,
- 24 hour trading volume on all exchanges according to CoinCodex[6] was 42.51 US dollars (no millions).

Although in the last years Dogecoin gained a lot of positive reputation as being a "lost cause" founding platform and especially in the 2021 the value of the coin has seen a rapid increase in price with the help of celebrity exposure [36], these recent developments were excluded from our analysis.

5.4. Basic Attention Token (BAT)

Figure 8 shows the BAT aggregated transactions and the conformance to Benford's law. The transactions of the BAT coin are mostly automatically generated as this coin is the basis of digital marketing platform and as such break the Benford's law prerequisites.

5.5. PIVX (PIVX)

Figure 9 shows the PIVX aggregated transactions and the conformance to Benford's law. There was no scandal reported for the PIVX project in the observation period (actually the authors could not find any notable anomaly for this crypto currency). The only speculation that the authors could give is that PIVX network relies on anonymous transactions that could be used to hide anomalies.

5.6. EOS (EOS)

Figure 10 shows the EOS aggregated transactions and the conformance to Benford's law. EOS is regarded as a valid project and survived till 2021, the only drawback that comes into the picture is that

---

[5]   Dogecoin twitter account: https://twitter.com/Dodgecoin
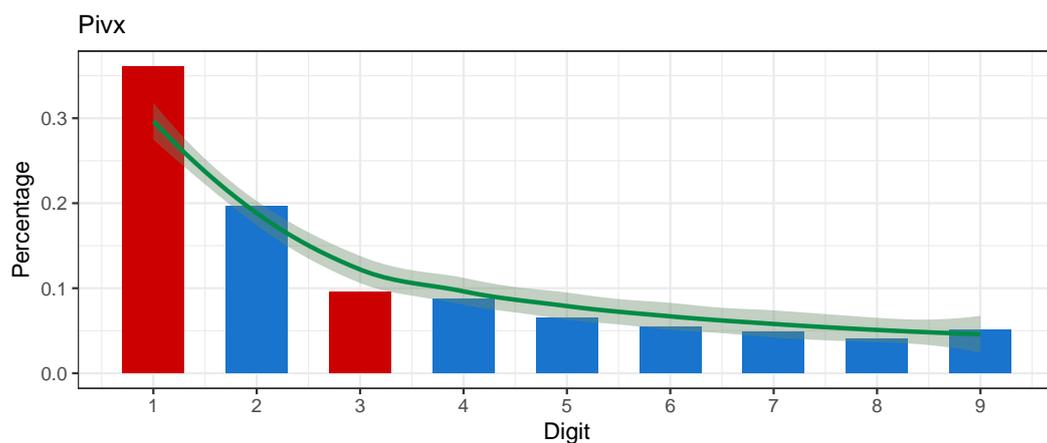[6]   Concodex: https://coincodex.com/crypto/XXX/exchanges/
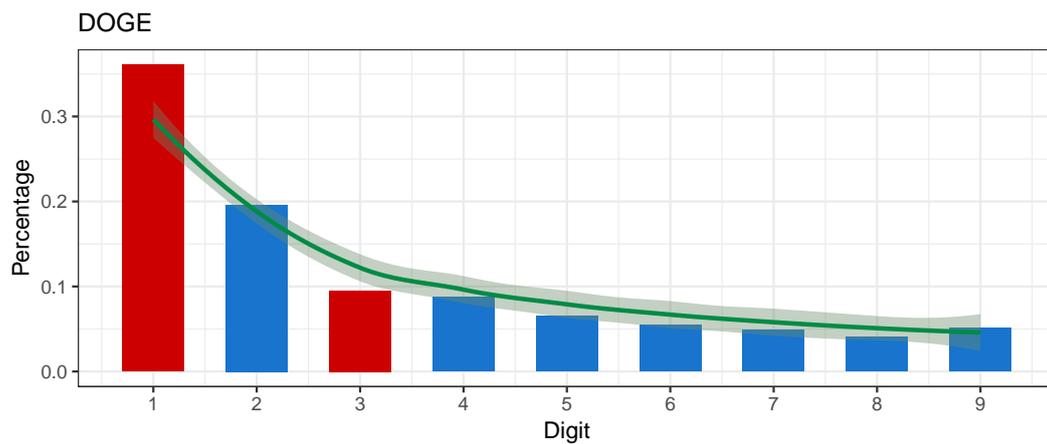
Figure 7. DOGE aggregated transactions and the conformance to Benford's law. Digit 1 overflows, digit 3 underflows. Overall the daily aggregated transaction values do not conform.
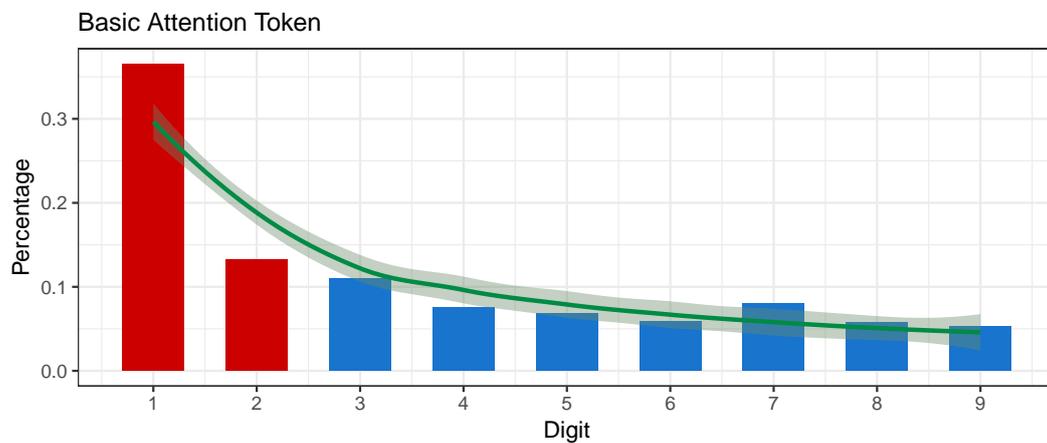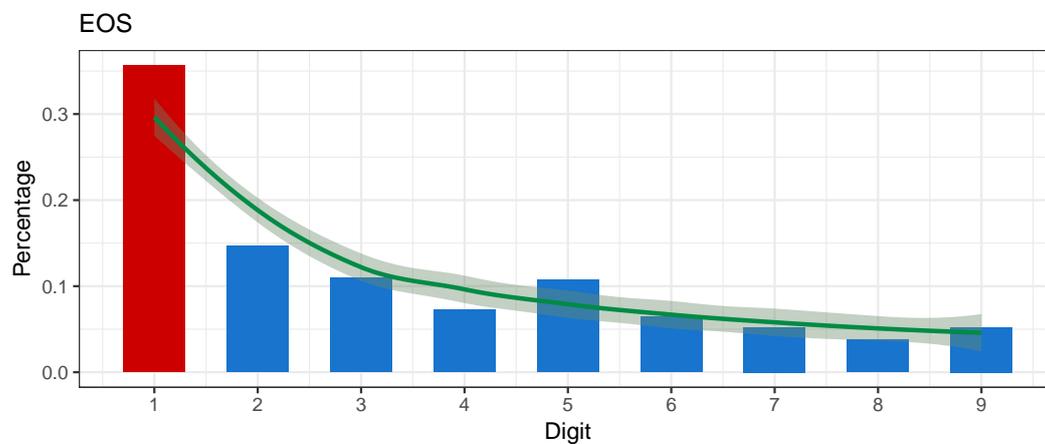

Figure 8. BAT aggregated transactions and the conformance to Benford's law. Digit 1 overflows, digit 2 underflows, digit 7 (almost) overflows. Overall the daily aggregated transaction values do not conform.

Figure 9. PIVX aggregated transactions and the conformance to Benford's law. Digit 1 overflows, digit 3 underflows. Overall the daily aggregated transaction values do not conform.

in the 2018 the project was in the starting phase and the backing capital risen by the backers of the project was an order of magnitude bigger than what the proposed project promised to accomplish [7].

Figure 10. TENX aggregated transactions and the conformance to Benford's law. Digit 1 overflows, digits 2 and 4 (almost) underflow. Overall the daily aggregated transaction values do not conform.

5.7. Additional currencies

An examination of all remaining cryptocurrencies that did not meet the criteria presented in Section 4, mainly due to the lack of data, show additional cases that support the validity of the presented method. By lowering the requirement for the minimum number of observations to 300 (days), we can observe additional cryptocurrencies that do not conform to Benford's law that have documented scams and scandals attributed to the observation period: Enigma (ENG) [8], SALT (SALT) [9], Waltonchain (WTC) [10].

6. Data availability

The data that support the findings of this study are openly available in Zenodo at https://zenodo.org/record/4682976, DOI: 10.5281/zenodo.4682976.

7. Discussion and future work

The main goal of the presented research was to test the applicability of Benford's law to the cryptocurrency domain, the research focused on some well documented anomalies and frauds from the past and compared the proposed metric on proven ecosystems that performed normally in the same time period. We focused on the time period between 2009 (time of the first transaction on Bitcoin network) and 2018 where there were already enough transactions to meet all Benford's law prerequisites, but also enough time has past so that the anomalies and frauds already emerged to public.

The results show that the proposed method is suitable for the proposed domain. All the big blockchain platforms by market capitalization that were not biased by any big scandal or lawsuit and

---

[7]   Why EOS Failed to Kill Ethereum: The Fatal Flaw of Centralization in a Decentralized Market: https://coincodex.com/article/10454/why-eos-failed-to-kill-ethereum-the-fatal-flaw-of-centralization-in-a-decentralized-market/

[8]   Enigma ethereum marketplace hijacked, investors duped by phishing scam: https://www.zdnet.com/article/enigma-ethereum-marketplace-hijacked-by-attackers/

[9]   SALT COIN EXIT SCAM! Massive selloff predicted by Morgan Stanley: https://www.youtube.com/watch?v=E2iNt3Z6qaY

[10]   Monumentally stupid tweet blows up in blockchain company's face: https://mashable.com/2018/02/28/waltonchain-twitter-scam-wtc/?europe=true

that are still functioning three years after the observation time-frame, such as Bitcoin (BTC), Ethereum (ETH) or OmiseGo (OMG), conform to Benford's law. The inspection of the six cryptocurrencies that were classified as non-conforming to Benford's law revealed 3 currencies with well documented anomalies: 2 (TENX and VERI) were tainted by scandals and lawsuits, one (DOGE) was invented as a joke and in the first years it was regarded so. As an additional observation, Dogecoin is now a respected cryptocurrency and in the last year grew to \$50B market capitalization. The method is obviously not suitable to predict the future of an observed cryptocurrency. The transactions of the BAT coin are mostly automatically generated as this coin is the basis of digital marketing platform, two remaining cryptocurrencies that were identified by the method as possible candidates for anomalous behaviour were EOS and PIVX, and although we could speculate to some extension why these two did not conform to Benford's law, the results are inconclusive.

All major cryptocurrencies that existed in the selected time-frame (2009 – 2018) were tested for the conformity to Bendord's law, the data availability statement is presented in Section 6.

Future work that is already undergone will focus on newer data, one such possible source has already been identified: Kaggle [11]. Another open issue that can be tackled with the same methodology is a comparison of all ERC20 tokens [37]. Ethereum based crypto currencies were selected to ensure a common (thus fair) technical basis – all these crypto currencies use the same technological platform, so all possible reasons for differences that arise from basic technology are eliminated.

References

1. Benford, F. The law of anomalous numbers. *Proceedings of the American philosophical society* **1938**, pp. 551–572.
2. Lansky, J. Possible state approaches to cryptocurrencies. *Journal of Systems Integration* **2018**, *9*, 19–31.
3. Nakamoto, S. Bitcoin whitepaper. Technical report, Bitcoin.org, 2008.
4. Jakobsson, M.; Juels, A. Proofs of work and bread pudding protocols. In *Secure information networks*; Jakobsson, M.; Juels, A., Eds.; Springer: Leuven, Belgium, 1999; pp. 258–272.
5. Shi, J.; Ausloos, M.; Zhu, T. Benford's law first significant digit and distribution distances for testing the reliability of financial reports in developing countries. *Physica A: Statistical Mechanics and its Applications* **2018**, *492*, 878–888.
6. Fischer, R.A. Statistical methods for research workers, 1925. *Edinburgh: Oliver and Boyd* **1925**.
7. Durtschi, C.; Hillison, W.; Pacini, C. The effective use of Benford's law to assist in detecting fraud in accounting data. *Journal of forensic accounting* **2004**, *5*, 17–34.
8. Newcomb, S. Note on the Frequency of Use of the Different Digits in Natural Numbers. *Amer. J. Math.* **1881**, *4*, 39–40. doi:10.2307/2369148.
9. Singleton, T.W. IT Audit Basics: Understanding and Applying Benford's Law. *Isaca Journal* **2011**, *3*, 6.
10. Kenny, D.A. Measuring model fit, 2015.
11. Berger, A.; Hill, T.P.; others. A basic theory of Benford's Law. *Probability Surveys* **2011**, *8*, 1–126.
12. Fewster, R.M. A Simple Explanation of Benford's Law. *The American Statistician* **2009**, *63*, 26–32. doi:10.1198/tast.2009.0005.

---

[11] Cryptocurrency Historical Prices: https://www.kaggle.com/sudalairajkumar/cryptocurrencypricehistory

13.   Kumar, K.; Bhattacharya, S. Detecting the dubious digits: Benford's law in forensic accounting. *Significance* **2007**, *4*, 81–83.
14.   Nigrini, M.J. Audit sampling using Benford's law: a review of the literature with some new perspectives. *Journal of emerging technologies in accounting* **2017**, *14*, 29–46.
15.   Drake, P.D.; Nigrini, M.J. Computer assisted analytical procedures using Benford's Law. *Journal of Accounting Education* **2000**, *18*, 127–146.
16.   Cleary, R.; Thibodeau, J.C. Applying Digital Analysis Using Benford's Law to Detect Fraud: The Dangers of Type I Errors. *AUDITING: A Journal of Practice & Theory* **2005**, *24*, 77–81. doi:10.2308/aud.2005.24.1.77.
17.   Hickman, M.J.; Rice, S.K. Digital Analysis of Crime Statistics: Does Crime Conform to Benford's Law? *Journal of Quantitative Criminology* **2010**, *26*, 333–349. doi:10.1007/s10940-010-9094-6.
18.   Burke, J.; Kincanon, E. Benford's law and physical constants: The distribution of initial digits. *American Journal of Physics* **1991**, *59*, 952–952. doi:10.1119/1.16838.
19.   Zhang, J. Testing Case Number of Coronavirus Disease 2019 in China with Newcomb-Benford Law, 2020, [arXiv:physics.soc-ph/2002.05695].
20.   Baum, S.C. Cryptocurrency fraud: A look into the frontier of fraud. PhD thesis, Georgia Southern University, 2018.
21.   Zuckoff, M. *Ponzi's scheme: The true story of a financial legend*; Random House Incorporated: Random House Tower, New York City, USA, 2006.
22.   Twomey, D.; Mann, A. Fraud and manipulation within cryptocurrency markets. *Corruption and Fraud in Financial Markets: Malpractice, Misconduct and Manipulation Edited by Alexander, C. and Cumming, D* **2020**, pp. 205–250.
23.   Gao, B.; Wang, H.; Xia, P.; Wu, S.; Zhou, Y.; Luo, X.; Tyson, G. Tracking Counterfeit Cryptocurrency End-to-end. *Proceedings of the ACM on Measurement and Analysis of Computing Systems* **2020**, *4*, 1–28.
24.   Brown, S.D. Cryptocurrency and criminality: The Bitcoin opportunity. *The Police Journal* **2016**, *89*, 327–339.
25.   Brenig, C.; Müller, G.; others. Economic analysis of cryptocurrency backed money laundering. ECIS 2015 Completed Research Papers; , 2015; pp. 1–18.
26.   Farrugia, S.; Ellul, J.; Azzopardi, G. Detection of illicit accounts over the Ethereum blockchain. *Expert Systems with Applications* **2020**, *150*, 113318.
27.   Carslaw, C.A. Anomalies in income numbers: Evidence of goal oriented behavior. *Accounting Review* **1988**, pp. 321–327.
28.   Gorard, S. Revisiting a 90-year-old debate: the advantages of the mean deviation. *British Journal of Educational Studies* **2005**, *53*, 417–430.
29.   Nigrini, M.J.M.J. *Benford's law : applications for forensic accounting, auditing, and fraud detection*; Wiley: 111 River Street, Hoboken, NJ, USA, 2012; p. 352.
30.   Pearson, K. X. On the criterion that a given system of deviations from the probable in the case of a correlated system of variables is such that it can be reasonably supposed to have arisen from random sampling. *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* **1900**, *50*, 157–175.
31.   Berger, V.W.; Zhou, Y. Kolmogorov–smirnov test: Overview. *Wiley statsref: Statistics reference online* **2014**.
32.   Freedman, L.S. Watson's UN2 statistic for a discrete distribution. *Biometrika* **1981**, *68*, 708–711, [https://academic.oup.com/biomet/article-pdf/68/3/708/582463/68-3-708.pdf]. doi:10.1093/biomet/68.3.708.
33.   Nigrini, M. Digital Analysis Using Benford's Law: Tests and Statistics for Auditors. *EDPACS* **2001**, *28*, 1–2, [https://doi.org/10.1201/1079/43266.28.9.20010301/30389.4]. doi:10.1201/1079/43266.28.9.20010301/30389.4.
34.   Druică, E.; Oancea, B.; Vâlsan, C. Benford's law and the limits of digit analysis. *International Journal of Accounting Information Systems* **2018**, *31*, 75 – 82. doi:https://doi.org/10.1016/j.accinf.2018.09.004.
35.   Alexander, J.C. Remarks on the use of Benford's Law. *Available at SSRN 1505147* **2009**.
36.   Livni, E. Serious money is flowing to the joke cryptocurrency Dogecoin. *New York Times* **2021**, *2.8.2021*, 1–2.
37.   Somin, S.; Gordon, G.; Altshuler, Y. Network analysis of erc20 tokens trading on ethereum blockchain. International Conference on Complex Systems; Springer, , 2018; pp. 439–450.