



The Rich Still Get Richer: Empirical Comparison of Preferential Attachment via Linking Statistics in Bitcoin and Ethereum

Dániel Kondor^{1*}, Nikola Bulatovic², József Stéger², István Csabai² and Gábor Vattay²

¹Singapore-MIT Alliance for Research and Technology, Singapore, ²Department of Physics of Complex Systems, Eötvös Loránd University, Budapest, Hungary

Bitcoin and Ethereum transactions present some of the largest real-world complex networks that are publicly available for study, including a detailed picture of their time evolution. As such, they have received a considerable amount of attention from the network science community along with analyses from economic and cryptographic perspectives. Among these studies, in an analysis on the early instance of the Bitcoin network, we have shown the clear presence of the preferential attachment, or the “rich-get-richer” phenomenon. Now, we revisit this question, using a recent version of the Bitcoin network that has grown almost 100-fold since our original analysis. Furthermore, we additionally carry out a comparison with Ethereum, the second most important cryptocurrency. Our results show that preferential attachment continues to be a key factor in the evolution of both the Bitcoin and Ethereum transaction networks. To facilitate further analysis, we publish a recent version of both transaction networks, and an efficient software implementation that is able to evaluate linking statistics necessary for learn about preferential attachment on networks with several hundred million edges.

Keywords: bitcoin, ethereum, preferential attachment, network science, degree distribution

OPEN ACCESS

Edited by:

Radosław Michalski,
Wrocław University of Science and
Technology, Poland

Reviewed by:

Cuneyt Gurcan Akcora,
University of Manitoba, Canada
Gianluigi Viscusi,
Imperial College Business School,
United Kingdom

*Correspondence:

Dániel Kondor
kondor@csh.ac.at

Specialty section:

This article was submitted to
Blockchain Technologies,
a section of the journal
Frontiers in Blockchain

Received: 16 February 2021

Accepted: 06 August 2021

Published: 23 August 2021

Citation:

Kondor D, Bulatovic N, Stéger J,
Csabai I and Vattay G (2021) The Rich
Still Get Richer: Empirical Comparison
of Preferential Attachment via Linking
Statistics in Bitcoin and Ethereum.
Front. Blockchain 4:668510.
doi: 10.3389/fbloc.2021.668510

1 INTRODUCTION

Cryptocurrencies have presented a disruptive change for both economics and computer science. Over the past years, interest in cryptocurrencies resulted in a huge amount of money invested in them (Baur et al., 2018; Begušić et al., 2018) and a growing amount of research carried out on diverse application possibilities of the underlying technologies, e.g., blockchain and decentralized trust (Bonneau et al., 2015; Yli-Huumo et al., 2016; Zheng et al., 2016; Seres et al., 2020; Liu et al., 2021). At the same time, cryptocurrencies provide a unique opportunity as financial systems where the whole list of transactions is exposed, making possible to study the dynamic interactions taking place in them (Kondor et al., 2014a; Phetsouvanh et al., 2019; Oggier et al., 2020; Wu et al., 2020); this allows the study of the complete history of how novel, alternative financial systems evolve from their inception (Seebacher and Maleshkova, 2018; Dixon et al., 2019).

Furthermore, the appearance of cryptocurrencies has helped research connecting network information with economical analysis to gain momentum due to the availability of high volume data (Gurcan et al., 2018). With several booms and busts in price dynamics, there have been a significant amount of interest in understanding and predicting price fluctuations (Kondor et al., 2014b; Akcora et al., 2018; Kurbucz, 2019), and trying to understand cryptocurrency markets based on a comparison with traditional financial instruments (Baur et al., 2018; Begušić et al., 2018).

Considering the list of transactions as an evolving network, cryptocurrencies present one of the largest real-world networks that can be analyzed by the scientific community, with several hundred million total edges. This can be of interest in itself, as it allows to test theories about evolving and time-varying networks on large scales with better statistical confidence. While there is significant interest in how cryptocurrencies work from a network science perspective (Di Francesco Maesa et al., 2018; Liang et al., 2018; Motamed and Bahrak, 2019; Wu et al., 2020; Fischer et al., 2021), we still do not have a comprehensive understanding of which are the relevant processes that shape their network structure.

In the current study, we evaluate key network characteristics on the transaction networks of Bitcoin and Ethereum, the two most popular cryptocurrencies. We specifically look at network evolution and the dynamics of how nodes gain new transaction partners and gain or lose balance. We build on our previous work that focused only on the initial phase of Bitcoin and found that preferential attachment drives the evolution of the transaction network and concentration of wealth (Kondor et al., 2014a). Considering the scale of Bitcoin and the many factors influencing transaction dynamics, it is remarkable how well power-law degree distributions and preferential attachment describe its evolution. In the current work, we extend our previous analysis to a significantly longer period of trading with multiple up- and downturns in the market for both Bitcoin and Ethereum; in the case of Bitcoin, this means an almost 100-fold growth in total network size. This allows us to test if the main transaction dynamics found previously stay significant during a timeframe when cryptocurrencies gained several orders of magnitude in total investment and became a main market component instead of just a niche. We show that a process of preferential attachment continues to be determinant for both cryptocurrencies and is robust with regard of the time period analyzed and the method used to reconstruct the transaction network.

We download and process the transaction history of both Bitcoin and Ethereum and reconstruct the temporally evolving transaction network. Since the main components of the network are the transactions which are instantaneous events, there are multiple possible choices for defining a network among the addresses (Kiffer et al., 2018; Motamed and Bahrak, 2019; Phetsouvanh et al., 2019; Wu et al., 2020). We show that the activity of addresses is characterized by fat-tailed distributions both in terms of temporal extent, number of transactions they participate in and addresses they come in contact with. Most addresses are short lived according to the practice of users of frequently generating new addresses to obtain increased privacy, while some addresses participate in an especially large number of transactions over an extended time range, giving rise to power-law degree distributions in the aggregated network (Di Francesco Maesa et al., 2018; Fischer et al., 2021).

We perform a more in-depth analysis of transaction dynamics, testing how preferential attachment can explain the broad degree distributions seen in the aggregated transaction networks. We evaluate statistics of new edge formation using the rank function methodology developed in our previous work (Kondor et al., 2014a) using different levels of temporal aggregation, testing also

the robustness of results. During our analysis, we perform an in-depth comparison among Bitcoin and Ethereum, focusing on comparing the transaction dynamics of regular addresses in the two systems and between addresses and smart contracts in Ethereum.

2 RELATED WORK

2.1 Preferential Attachment

Preferential attachment is a model of network evolution originally suggested by Barabási and Albert (1999) and Barabási et al. (1999), based on the models studied originally in different contexts by Yule (1925) and Simon (1955). The original model predicts a power-law degree distribution with an exponent of $\gamma = 2$; it was later generalized to yield networks with power-law degree distributions of arbitrary exponents (Dorogovtsev and Mendes, 2000a). Preferential attachment was observed either directly or indirectly in many real-world complex networks in the past decades (Jeong et al., 2003; Kunegis et al., 2013; Perc, 2014), including an early phase of Bitcoin (Kondor et al., 2014a).

The original model of Barabási and Albert (1999) and Barabási et al. (1999) assumes a continually growing network, where only newly joined nodes initiate edges, and connection probabilities depend linearly on the degree of existing nodes. While this captures key aspects of growing networks, many questions naturally arise about the importance of the underlying assumptions and the extent they are expected to be present in real-world networks. In accordance with this, researchers have focused on investigating potential generalizations in multiple directions, gaining insights into a more generalized class of dynamical processes that involve a form of preferential attachment (Albert and Barabási, 2002).

Models of preferential attachment typically assume that the dependence between node degrees and the probability of connecting to them follows a functional form such as: $p(k) \sim k^a$. A key early result was that for growing networks, an asymptotically linear form, i.e., $a = 1$ is required to result in a power-law degree distribution (Krapivsky et al., 2000). In the sublinear case ($a < 1$), the result in the degree distribution is a stretched exponential, while the $a > 1$ case yields highly concentrated networks, where only a finite number of nodes will have degrees larger than a threshold value k_a (dependent on the a exponent) even in the infinite limit (Krapivsky et al., 2000; Albert and Barabási, 2002).

In the case of asymptotically linear preferential attachment, a series of generalizations show that differences in the exact form for small k (including a finite probability to connect to zero degree nodes), and in edge dynamics can result in power-law degree distributions with a wide range of exponents beyond the $\gamma = 2$ case of the original model of Barabási and Albert (Dorogovtsev et al., 2000a; Dorogovtsev and Mendes, 2000a; Dorogovtsev and Mendes, 2000b; Dorogovtsev and Mendes, 2001a; Dorogovtsev and Mendes, 2001b). An especially interesting case is networks with accelerated growth, a case where not only the number of nodes, but average node degrees are growing (Dorogovtsev and Mendes, 2001b): in this case, internal edges are assumed to appear

with a time-dependent rate. Again, linear preferential attachment leads to power-law degree distributions, with exponents that are dependent on the growth rate of the average node degree.

Recently, mechanisms that go beyond looking at node degrees have also been investigated. A motivation for this is social networks, where agents are not assumed to make conscious decisions based on existing node degrees (a global information that might not be readily available); on the other hand, node betweenness can be a strong determinant on possible opportunities for establishing new connections (Topirceanu et al., 2018). More generalized measures of node fitness have also been suggested such as based on recency (Nsour and Sayama, 2021) or a combination of node degrees with an aging factor (Dorogovtsev and Mendes, 2000a). In real networks, it has been shown that a combination of node degree and fitness can explain network growth statistic well (Pham et al., 2015; Pham et al., 2016; Aspembitova et al., 2019).

2.2 Cryptocurrency Analysis

Cryptocurrencies present one of the largest complex network datasets available for study. Beside theoretical interest, there are many practical implications of intelligence extracted from analyzing the transaction dynamics.

Early works typically focused on identifying main features, and characterization of cryptocurrency networks via established metrics in network science (Ron and Shamir, 2013; Kondor et al., 2014a); in the case of Bitcoin, a distinct early phase was identified where the system functioned more as an experiment in its initial 2 years, before wider adoption (Kondor et al., 2014a; Liang et al., 2018). Networks defined based on the transactions, and dynamical properties show fat-tailed distributions typical of complex systems; beside degree distributions, this was observed in metrics of address activity, such as inter-event times as well (Kondor et al., 2014a; Guo et al., 2019; Wu et al., 2020). Further aspects of complex networks identified in cryptocurrencies include the small-world property, network densification over time (Di Francesco Maesa et al., 2018; Ferretti and D'Angelo, 2020; Wu et al., 2020) and a presence of disassortative mixing, suggesting that a significant number of transactions happen between ordinary users and large players providing services (Kondor et al., 2014a; Guo et al., 2019).

Preferential attachment was demonstrated as a mechanism generating fat-tailed degree and wealth distribution in Bitcoin in our previous work (Kondor et al., 2014a). In contrast, now we are looking at a network that is a result of an almost 100-fold growth since our initial analysis, thus it is an exciting question whether the initially identified dynamics have continued to hold over this intense expansion of activities. Furthermore, we present a comparison with Ethereum, whose network is different both on a technical level (it uses an account based model instead of an UTXO model; see the next section for an explanation of this difference) and on a conceptual level by the presence of smart contracts. While several previous works focused on the comparison of network structure in Bitcoin, Ethereum and other cryptocurrencies, these works did not include the analysis of preferential attachment (Ron and Shamir, 2013; Liang et al., 2018; Guo et al., 2019; Ferretti and D'Angelo,

2020; Wu et al., 2020). The recent work of Di Francesco Maesa et al. (2018) analyzed measures that correspond to a concentration of wealth and found that it shows an increase over time, consistent with a “rich-get-richer” phenomenon; we note that these results were obtained by investigating the time evolution of aggregated measures and not the dynamics of individual transactions, i.e., a vastly different methodology to our previous (Kondor et al., 2014a) and current work, yet are compatible with our main findings.

Recently, Aspembitova et al. (2019) focused specifically on preferential attachment in Bitcoin and suggested a fitness-based model to explain the power-law degree distributions in Bitcoin that is consistent with the short-lived nature of addresses. While their analysis provides an interesting conceptual framework similar to the “hot-get-richer” and fitness-based models (Pham et al., 2016; Nsour and Sayama, 2021), we believe their reasoning against degree preferential attachment to be problematic, as they are considering empirical connection probabilities as a function of node degree without taking into account the underlying evolving degree distribution of the network. In our work, we explicitly update the degree distribution over the course of the network evolution, allowing us to evaluate a true *preference* toward nodes with higher degrees among all available ones at any point in time. This way, our investigation of preferential attachment is directly comparable to previous works where it was empirically found in networks in different contexts (Jeong et al., 2003; Kunegis et al., 2013; Perc, 2014).

Going beyond network structure, an important research direction utilizing cryptocurrency network information focuses on anonymity and the traceability of transactions. While anonymity was not among the original design goals of Bitcoin, cryptocurrency transactions are often regarded pseudo-anonymous despite the public record of them, since linking addresses to actual users is only possible using externally available information. Building on this, multiple mechanisms were later proposed to enhance anonymity in Bitcoin and several alternative cryptocurrencies were implemented with a stronger focus on anonymity (Bonneau et al., 2015; Anderson et al., 2016; Heilman et al., 2017; Conti et al., 2018). Several heuristics were proposed for address clustering based on transaction patterns, i.e., for identifying groups of addresses in cryptocurrency networks that are controlled by the same entity (Meiklejohn et al., 2013; Nick, 2015; Phetsouvanh et al., 2019; Fischer et al., 2021). In line with this, there exist solutions and services with the goal of “mixing” bitcoins or other cryptocurrencies, with the goal of making the flow of money less traceable; formal and practical analysis of such possibilities has attracted a significant research interest as well (Chen et al., 2017; Heilman et al., 2017; Miller et al., 2017; Conti et al., 2018). Practical applications typically focus on tracing the movement of money linked to illegal and illicit activities such as extortion or the sale of prohibited items (Portnoff et al., 2017; Paquet-Clouston et al., 2019; Oggier et al., 2020).

Finally, we note that there is a significant research interest in modeling price fluctuations in cryptocurrency markets and uncovering connections to network dynamics. Results include the characterization of price fluctuations and comparison with

traditional financial instruments (Baur et al., 2018; Begušić et al., 2018); characterization of risk based on network structure and motifs (Gurcan et al., 2018; Dixon et al., 2019); establishing connections between network activity and price dynamics (Kondor et al., 2014b; Alabi, 2017); and developing price predictions by exploiting inherent information in the transaction network (Akcora et al., 2018; Kurbusz, 2019). In the current work however, we only focus on the evolving network structure, and not consider market fluctuations.

3 METHODS

3.1 Data Collection

3.1.1 Bitcoin

We adapted the Bitcoin Core client program (version 0.19) by adding functionality to write out data about transactions and blocks in a CSV format¹. We used this client to download and extract the blockchain on February 7, 2020. Our data includes 616,345 blocks with 500,663,153 transactions among 609,963,452 unique addresses in total.

We note that Bitcoin uses an UTXO (unspent transaction output) graph model: a transaction lists a number of outputs as bitcoin values together with cryptographic challenges required to spend them in the future. The network itself has no notion of addresses or balances. In practice, almost all transaction outputs follow a standard pattern, requiring a signature with a given private key for spending it. A standard representation of the hash of the corresponding public key is then referred to as the “Bitcoin address” that received the amount associated with that output. There is a small fraction of transactions where such an association cannot be easily made (Caprolu et al., 2021); while the flow of Bitcoins can still be followed in these cases, we did not associate such transaction outputs with any Bitcoin address.

A key feature of Bitcoin is that any user can have an unlimited number of addresses. While balances are not kept track of by the network explicitly, it is possible to calculate them by summing up either all incoming and outgoing transaction values associated with an address, or alternatively, all unspent transaction outputs corresponding to it.

We construct a network among addresses by creating a directed edge between each input and output address for each transaction, excluding self-edges. The resulting network has 3,648,627,182 unique edges, that appear 4,834,306,446 times in total. Note that in Bitcoin, a transaction can have multiple input and output addresses and thus can result in the addition of multiple edges (Phetsouvanh et al., 2019); e.g., a transaction with 10 distinct input and output addresses will result in 100 edges. Also, transaction inputs must always include the full amount received by a previous transaction output; when spending less than this amount, the remainder (or “change”) is directed to one of the addresses of the spending user in a separate transaction

output. This results in a large number of self-edges and chains in practice.

As it is common practice to create new addresses regularly (it is often advised not to reuse addresses), there have been a significant research interest in trying to identify groups of addresses that belong to the same user, or are controlled by the same entity (Wu et al., 2020; Fischer et al., 2021; Liu et al., 2021). At the same time, there are several known methodologies and services that aim to “mix” bitcoins in a way that limits the possibility of such grouping and tracking the flow of money as well (Bonneau et al., 2015; Heilman et al., 2017; Phetsouvanh et al., 2019). In the current work, we do not attempt such grouping, instead we look at network structure and dynamics at the level of individual addresses.

3.1.2 Ethereum

We use the OpenEthereum client to synchronize with the blockchain and then use the Ethereum-ETL client to output the transaction history in CSV format. We extracted data on February 2, 2020; this includes the first 9.4 million blocks in the chain, with a total of 628,810,973 transactions among 68,429,208 unique addresses. Ethereum transactions are one-to-one: each transaction has only one input and output address and thus can be directly mapped to a directed edge in a network among addresses. Contrary to Bitcoin, in Ethereum, the balance of an address is recorded as an intrinsic property in the system; this way, spending is possible in any denomination, and does not require the “change” mechanism used in Bitcoin. Similarly to Bitcoin, a user can have an unlimited number of addresses; grouping these can be even more difficult, since there are less clear transaction patterns that reveal connections among addresses.

Beside addresses directly controlled by users, Ethereum allows the creation of smart contracts, that are essentially algorithms deployed in the networks, associated with addresses as well (Wood, 2014; Anderson et al., 2016; Kiffer et al., 2018; Victor and Lüders, 2019). After the creation of a smart contract, it exists independently from its creator; Ethereum users can interact with smart contracts via sending them money, and by “function calls,” requesting certain some of the functionality exposed via the smart contract interface to run. Smart contracts can react to such interactions by creating further transactions themselves. In our analysis, we separated addresses associated with smart contracts from addresses controlled by regular users.

3.2 Edge Lifetime

In the usual picture of growing complex networks, edges are typically considered static entities that represent existing connections that can be gained or lost over time. For transactions in cryptocurrencies, this picture is not accurate: since transactions are instantaneous events, the presence of an edge in our network indicates that at least one transaction took place between two addresses over the lifetime of the network. Given the timescales in our analysis, edges that correspond to transactions that happened a long time ago lose their relevance (e.g., if a user abandons using a certain address, as is often the case). To account for this, we can use an alternate network

¹Source code of our modified client is available at <https://github.com/dkondor/bitcoin/tree/0.19>

definition, where edges have a finite “lifetime”: they are created when a transaction happens between two addresses, and are removed if a certain time passes without repeated transactions between the same pair of addresses. Removal of an edge also decreases the network degree of the associated nodes. This means that activity is gradually “forgotten,” at least for the purpose of our analysis. This procedure is similar to some of the “aging” processes suggested in theoretical models of growing complex networks (Dorogovtsev and Mendes, 2000a; Dorogovtsev and Mendes, 2000b).

In this case, the indegree of a node naturally represents the number of distinct transaction partners it had in a recent time interval. We can choose this time interval to correspond to a presumption of “memory” in the dynamics between addresses. In practice, we created networks where the lifetime of edges was limited to 1 day and 30 days beside the fully time-aggregated network.

3.3 Preferential Attachment

In the current work, we consider a generalized non-linear form of preferential attachment (Krapivsky et al., 2000), where probabilities of connecting to a node with degree k are given by the following equation

$$p(k) \sim k^a \quad (1)$$

with appropriate normalization. In this case, the probability of connecting to *any* node with degree k is then

$$\Pi(k) \sim n(k)k^a \quad (2)$$

where $n(k)$ is the number of nodes with degree k in the network (i.e., the empirical degree distribution).

The case of Bitcoin and Ethereum is clearly more complex than the simple growing network models used in most theoretical works about preferential attachment. With the constant addition of new nodes and transaction partners, both Bitcoin and Ethereum can be regarded as a growing network. At the same time, it is less clear what to consider as the lifetime of nodes and edges. Essentially, a transaction represents an instantaneous interaction, thus representing it as the addition of an edge to a network might be misleading. At the same time, the number of past transaction partners as represented by the indegree of a node is a meaningful metric that can be indicative of a form of “fitness” that also related to the capacity to attract new transaction partners.

In this paper, we ask whether a form of preferential attachment is present in the evolution of the Bitcoin and Ethereum transaction network. We use nodes’ indegrees as the base metric that is assumed to be related to connection probability of new edges. We perform our analysis both on time aggregated networks over the whole lifetime of the systems, and also variants where we consider edges to have a limited lifetime, thus indegrees more directly correspond to a measure of “recency” or “hotness” (Dorogovtsev and Mendes, 2000a; Dorogovtsev and Mendes, 2000b; Nsour and Sayama, 2021).

In our analysis, we focus on a model of nonlinear preferential attachment described by Eqs 1, 2 (Krapivsky et al., 2000). Importantly, we do not restrict this process to links from new

nodes, as we expect a significant amount of links to be created between already existing nodes, a departure from the original Barabási-Albert model (Barabási and Albert, 1999), but a case considered in previous theoretical models as well (Albert and Barabási, 2000; Dorogovtsev et al., 2000a; Dorogovtsev et al., 2000b). We consider the case of $a = 0$ a null model, where connection probabilities are independent of node indegrees. We compare results for this case with $a > 0$ to assess the importance of node degrees in establishing new transaction partners.

In an evolving network, the degree distribution will change over time, making it difficult to compare probabilities of events that occur at different times with different network configurations. We overcome this problem by calculating the *transformed rank* of the target indegree for each linking event:

$$R \equiv \frac{\sum_{k=0}^{k_{\text{target}}} n(k)k^a}{\sum_{k=0}^{k_{\text{max}}} n(k)k^a} \quad (3)$$

where k_{target} is the indegree of the node receiving the new link. If our assumption about the preferential attachment process and the a exponent holds true, then empirical R values calculated for a set of linking events will be distributed in a uniform way over the $[0, 1]$ interval (Kondor et al., 2014a). Since the R transformed rank values are normalized this way, values from different time points (and thus different stages of the evolving network) can be analyzed together. Furthermore, by limiting the set of events considered to smaller time intervals, the role of the preferential attachment process in network evolution at different times can be easily compared.

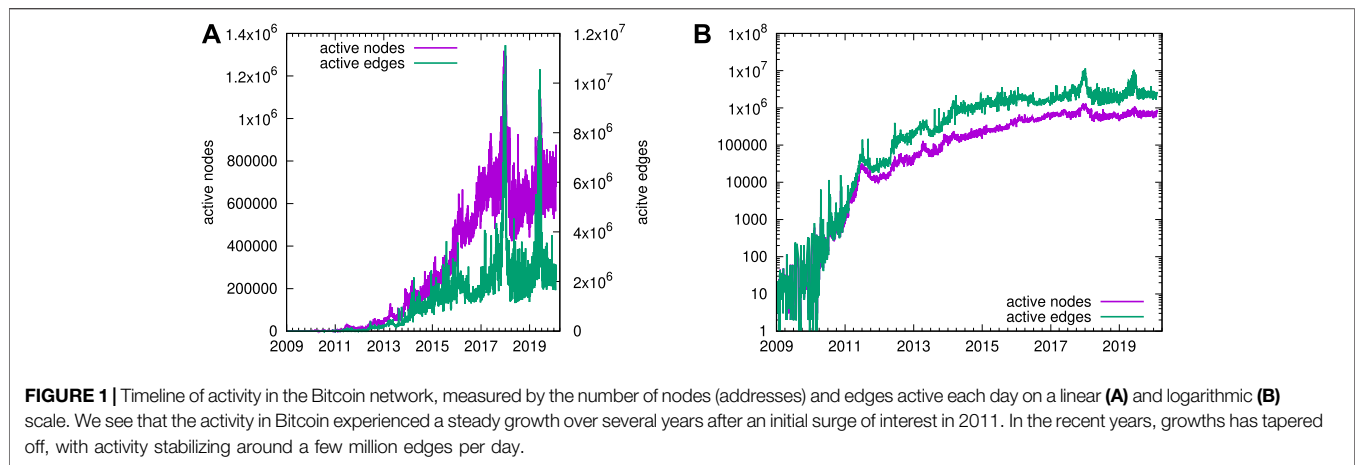
In practice, we can calculate transformed ranks for any value of the a exponent. In this article, we compare several a values and identify the one that best fits a uniform distribution. As noted above, we consider a null hypothesis of no preferential attachment (i.e., a case where network degree does not affect the probability of attracting new transaction partners) with $a = 0$.

Evaluating the statistics of preferential attachment requires calculating the R value in Eq. 3 for each “event,” based on the actual degree distribution in the network. Since the number of transactions is in the order of hundreds of millions for both networks, a direct summation over the degree distribution (that has a runtime complexity of $O(N)$ for a network of N nodes) is not feasible. However, using a properly augmented binary search tree as the data structure to store the degree distribution along with partial sums of k^a , we are able to perform the calculation of R values in $O(\log N)$ time complexity, making it possible to evaluate the distribution of R values over hundreds of millions of events. We describe the necessary tools used for this purpose in the Supplementary Material, while we publish the source code of an efficient augmented binary search tree implementation used for this purpose online (Kondor, 2020a; Kondor, 2020b).

4 RESULTS

4.1 Network Growth and Structure

Both Bitcoin and Ethereum has experienced a great amount of growth over their lifetime, including multiple “peaks,” where a



sudden surge of interest resulted in large upticks of both exchange price and network activity (Figures 1, 2) (Alabi, 2017). Since early 2018 when cryptocurrencies gained an unprecedented global attention, daily activity for both Bitcoin and Ethereum has had an approximately constant rate however, in contrast to previous periods of growth. This could be the consequence of getting close to the technical limits of transaction volume that the networks are able to handle, as both Bitcoin and Ethereum have hard limits on the amount of data, and thus the number of transactions that can be included in blocks: Bitcoin directly limits the block size, while Ethereum limits the maximum computational resources to be used in blocks (Zheng et al., 2016). Approaching this limit will result in transaction fees increasing since miners will prefer to include transactions with more fees. This functions as a natural feedback loop that discourages creating too many transactions and thus limits the network activity. Also, since the beginning of 2018, the total capitalization of cryptocurrencies (for simplicity, defined as the total value of coins in circulation based on the current exchange rate) have approached that of publicly traded stocks with the highest capitalization; this could limit further speculative investment in them.

We perform a simple characterization of structure by looking at the degree distribution of transaction networks. More specifically, we are interested in indegree distributions, since these can be interpreted as a measure of capacity to attract interaction with external entities. Both networks are characterized by fat-tailed distributions over their lifetime that are well approximated with power-laws (Figures 3, 4). The stability in shape of these distributions is especially remarkable considering that different stages of the networks depicted in Figures 3, 4 represent an over 100-fold increase in size (over 10,000-fold increase in the case of Bitcoin when comparing very early instances with the latest ones). We note that the presence of addresses with extremely high indegrees suggests that address reuse is common at least in some part of the user base of these cryptocurrency networks, despite the commonly cited recommendations against it (Bonneau et al., 2015; Wu et al., 2020; Fischer et al., 2021). We note that not all cryptocurrency users will have strong privacy requirements regarding the use of all of their addresses; there are many use cases where reuse of

well-known addresses is expected as part of normal operations. Nevertheless, it is still important to see that the subset of users who avoid address reuse (either manually or in an automated way implemented in a wallet software) is not large enough to statistically alter the properties of the indegree distribution.

4.2 Preferential Attachment

We test for the presence of preferential attachment by considering all transactions that add new links to the aggregated networks and calculating transformed ranks according to Eq. 3. In Figures 5, 7, we display the transformed ranks in order, i.e., as a function of their cumulative distribution function (CDF), for the case of the Bitcoin and Ethereum transaction networks, and for the evolution of Bitcoin balances. For each case, a perfect fit with the model of nonlinear preferential attachment (i.e., Eq. 2) would be a straight line, corresponding to the case where the transformed ranks are uniformly distributed in the $[0, 1]$ interval. Finding an exponent that best describes the process means finding a case where a straight line best approximates the distribution of transformed rank values. It has been suggested previously that deviations from a perfect fit can arise due to the large amount of automated and spam-like activity in cryptocurrencies (Di Francesco Maesa et al., 2018; Zwang et al., 2018; Liu et al., 2021).

In most cases, a significant feature is that the distributions do not start from zero. This means that there is a large number of transactions that target newly created addresses, in contrast to the original nonlinear preferential attachment model, where the probability of an edge targeting a non-existent node (i.e., a node with a degree of zero) is zero. This is understandable given that users can freely create any number of addresses, and are advised to often move their wealth to new addresses. Also, many service providers create unique addresses for their customers, which necessarily have zero degree then. Given this, we restrict the preferential attachment model to only apply to existing addresses, while we acknowledge that linking to new addresses is governed by more specific rules that are relevant to cryptocurrency system usage (Di Francesco Maesa et al., 2018; Fischer et al., 2021).

Given this observation, we only focus on nonzero transformed ranks when considering if they can be fitted with a uniform

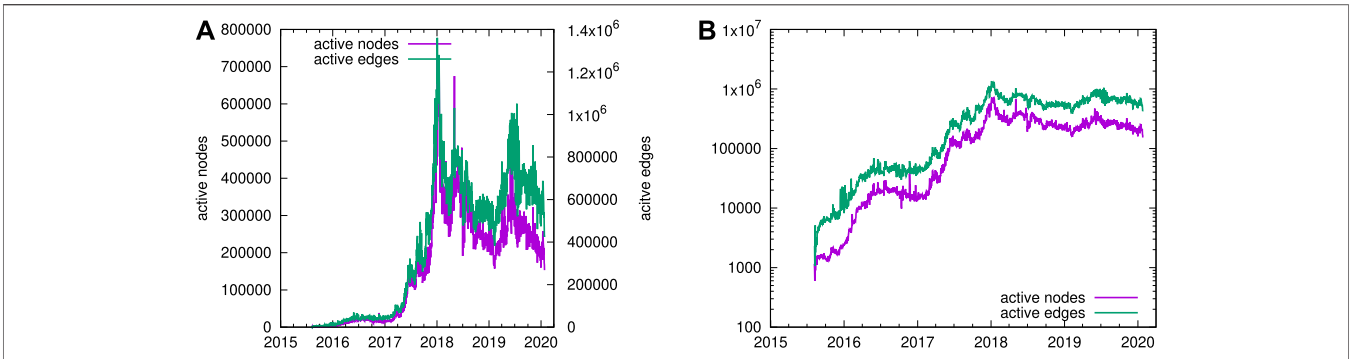


FIGURE 2 | Timeline of activity in the Ethereum network, measured by the number of nodes (addresses) and edges active each day on a linear **(A)** and logarithmic **(B)** scale. Growth of activity here is characterized by two distinct phases: an approximately exponential growth phase in the first 2.5 years, followed by an approximately constant level of activity in the past years.

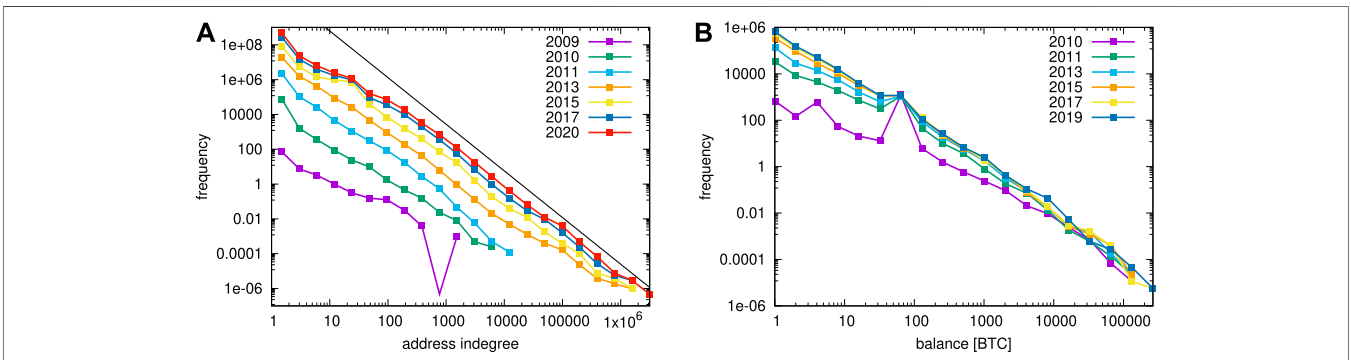


FIGURE 3 | Distribution of network indegrees **(A)** and address balances **(B)** for Bitcoin. Indegrees are determined by the total number of distinct transaction partners over the lifetime of the network. Both of these distributions are fat-tailed and are robust over the period of almost 10 years despite the size of the network increasing by multiple orders of magnitude. The black line in the left figure shows a power-law fit for the final distribution that has an exponent of -2.68 . The fit was carried out with the *plfit* package (Nepusz, 2020), based on the algorithm of Clauset et al. (2009).

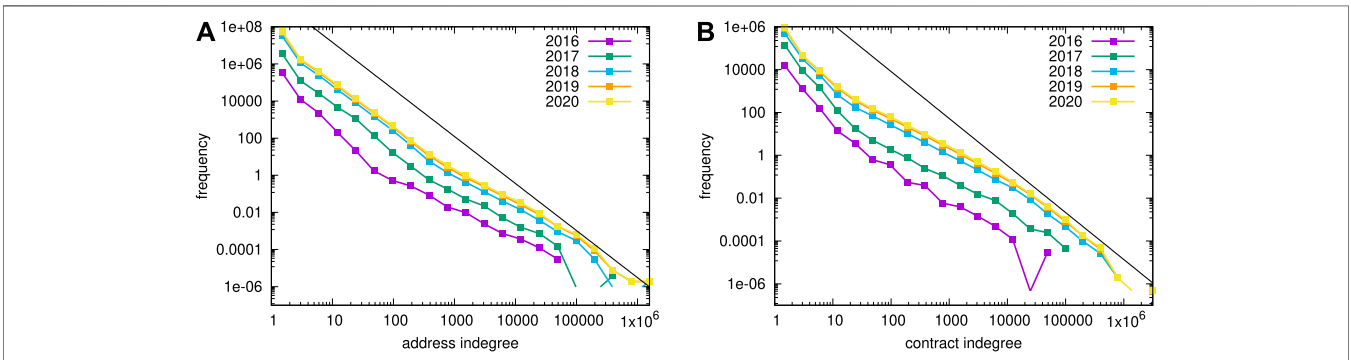


FIGURE 4 | Indegree distribution of regular addresses **(A)** and contract addresses **(B)** in Ethereum. These distributions are also characterized as fat-tailed ones, and are well approximated by power-laws, similarly to Bitcoin. Again, the time evolution is robust over a period of almost 5 years, during which the Ethereum network grew over 100-fold. Black lines show power-law fits for the final distribution, with exponents of -2.54 and -2.19 for addresses and contracts respectively. Fits were carried out with the *plfit* package (Nepusz, 2020), based on the algorithm of Clauset et al. (2009).

distribution. Graphically, this corresponds to starting the lines that represent such uniform distributions (the black lines in **Figures 5, 7**) from the CDF value that corresponds to the first nonzero transformed rank.

In each case, we verify the presence of preferential attachment by comparing the transformed rank distributions between the $a = 0$ and $a > 0$ cases. Using an exponent of $a = 0$ assumes that there is no relation among node degrees and connection probabilities,

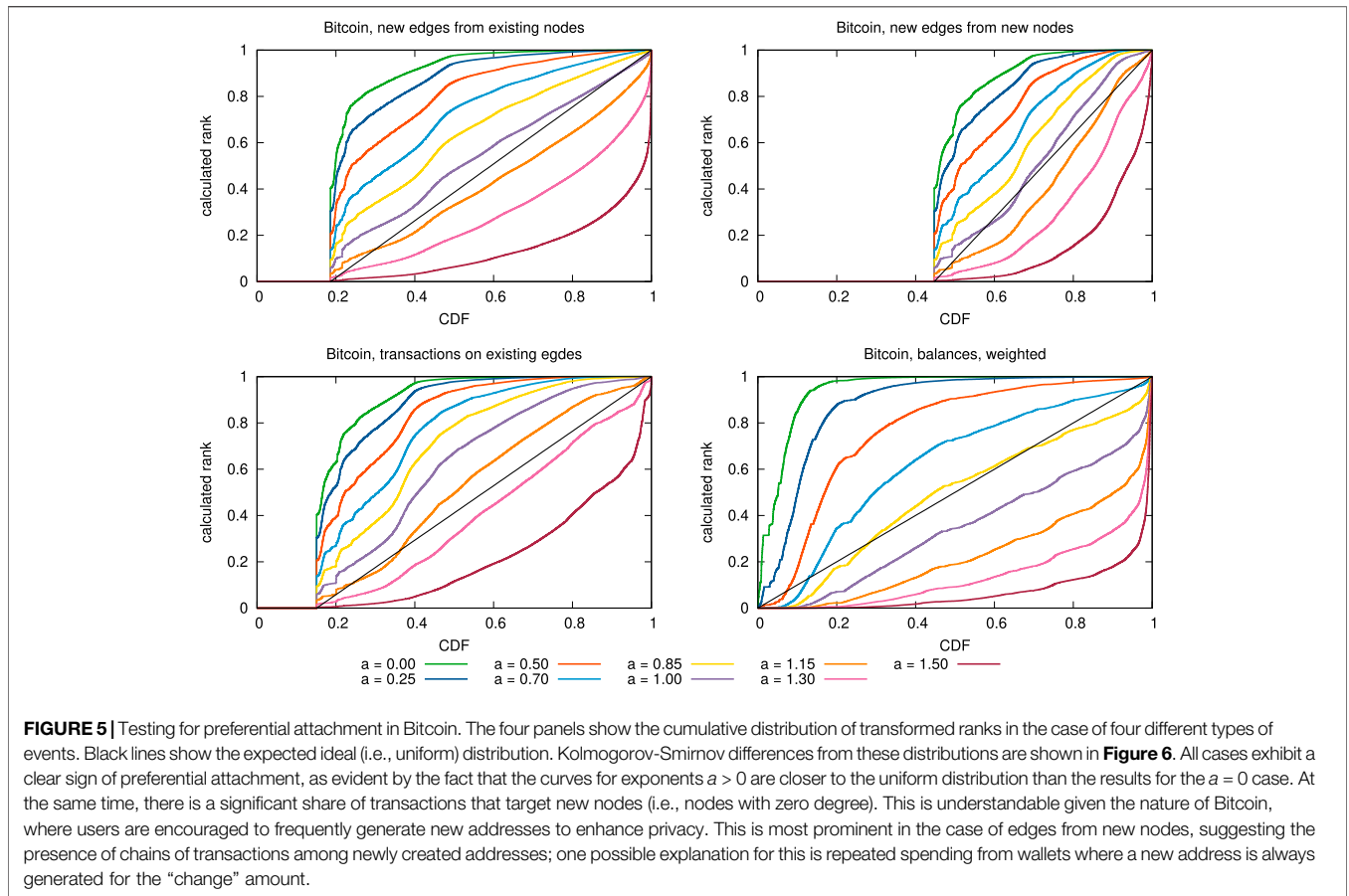


FIGURE 5 | Testing for preferential attachment in Bitcoin. The four panels show the cumulative distribution of transformed ranks in the case of four different types of events. Black lines show the expected ideal (i.e., uniform) distribution. Kolmogorov-Smirnov differences from these distributions are shown in **Figure 6**. All cases exhibit a clear sign of preferential attachment, as evident by the fact that the curves for exponents $a > 0$ are closer to the uniform distribution than the results for the $a = 0$ case. At the same time, there is a significant share of transactions that target new nodes (i.e., nodes with zero degree). This is understandable given the nature of Bitcoin, where users are encouraged to frequently generate new addresses to enhance privacy. This is most prominent in the case of edges from new nodes, suggesting the presence of chains of transactions among newly created addresses; one possible explanation for this is repeated spending from wallets where a new address is always generated for the “change” amount.

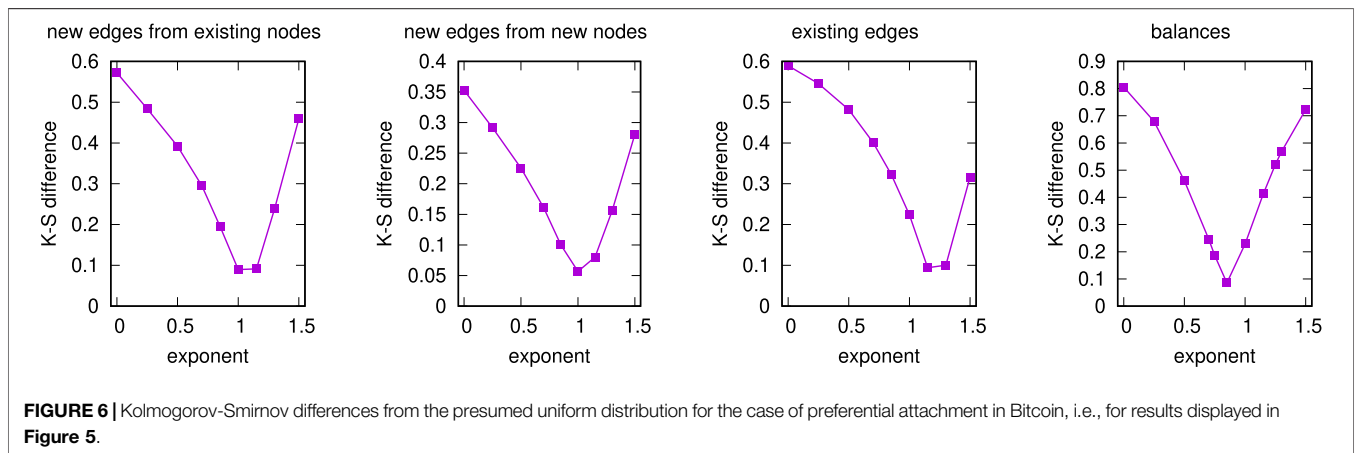
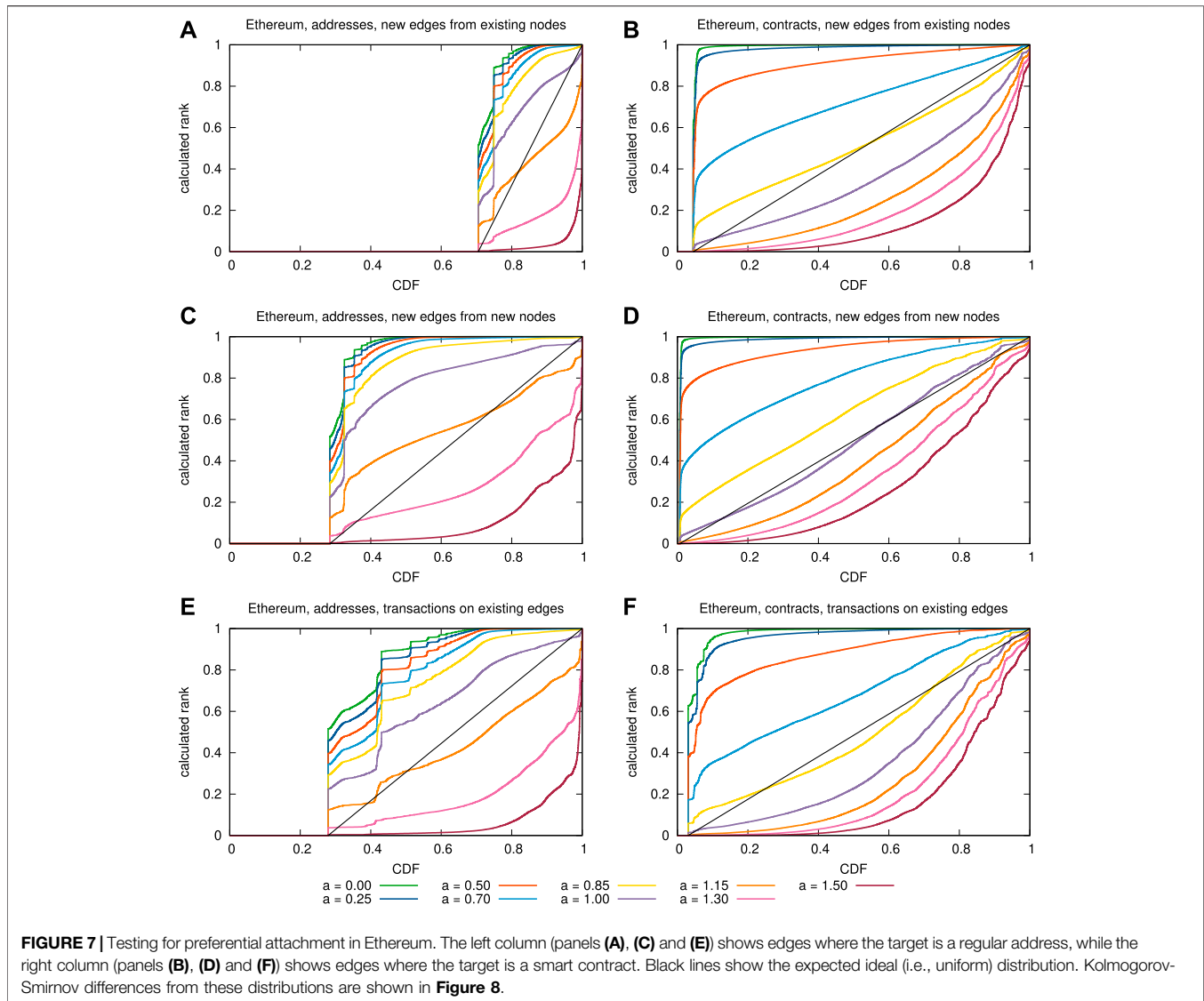


FIGURE 6 | Kolmogorov-Smirnov differences from the presumed uniform distribution for the case of preferential attachment in Bitcoin, i.e., for results displayed in **Figure 5**.

while an exponent of $a > 0$ assumes a positive correlation. For all cases considered in this study, we see strong evidence for the presence of a preferential attachment process as results for $a > 0$ always provide a much better fit to the assumed uniform distribution. Beside visual inspection of the fits, we calculate the Kolmogorov-Smirnov difference from a uniform distribution, and present this as a function of the a exponent in **Figures 6, 8**. Overall, exponents around $a = 1$ give the best fits;

however, there are some further interesting observations regarding typical values.

In the case of the Bitcoin transaction network, linear preferential attachment is the most plausible model for the case of newly created edges, either from new or from existing nodes. This is consistent with our earlier results (Kondor et al., 2014a) that were done for this network at a much earlier stage. For the case of repeated edges (i.e., repeated transactions on edges



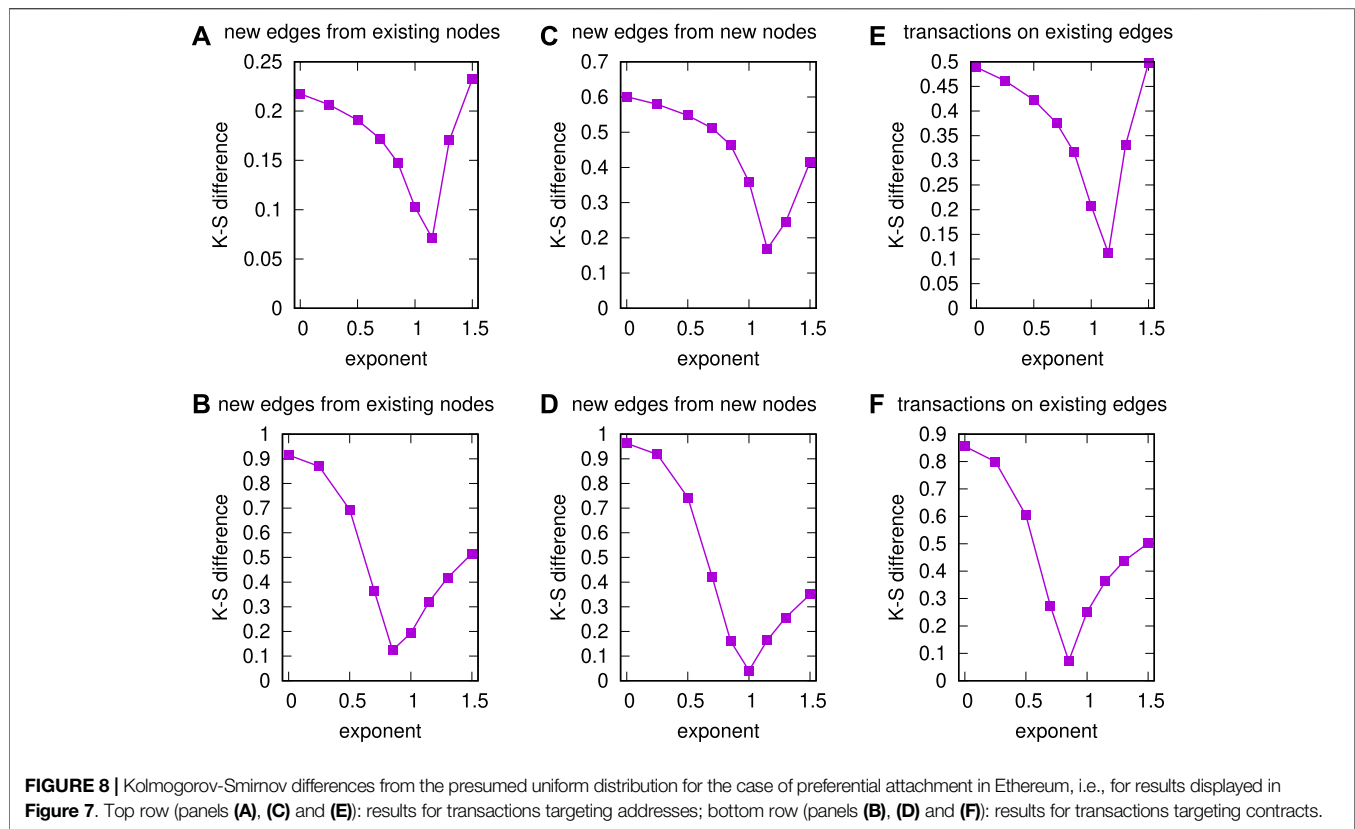
that appeared before), we see a slight superlinear case, with $a = 1.15$ and $a = 1.3$ both giving almost equally plausible fits. Furthermore, we also tested for preferential attachment in the case of money dynamics, i.e., related to the flow of Bitcoins. In this case, instead of node degrees, we considered the balance of the target address, and also weighted the CDF values with the transferred Bitcoin amount. We see evidence of slightly sublinear preferential attachment, with $a = 0.85$ being the most plausible exponent. This is again consistent with our earlier results (Kondor et al., 2014a) and work that found direct evidence of a “rich-get-richer” phenomenon in Bitcoin based on inferred wealth of users (Di Francesco Maesa et al., 2018).

In the case of Ethereum, we separately analyze the case where edges connect to regular addresses (left column in **Figure 7**; top row in **Figure 8**) and the case where the target of an edge is a smart contract (right column in **Figure 7**; bottom row in **Figure 8**). For regular addresses, we see some evidence of

superlinear preferential attachment ($a = 1.15$ being the most plausible exponent); nevertheless, a uniform distribution does not seem a very good fit in this case, as we see significant further features in the distribution of transformed ranks in **Figure 7**. Still, we can say that a form of preferential attachment is important in this process, since the case of $a = 0$ gives a much worse agreement with the empirical distribution of transformed ranks than any other case. For smart contracts, the distributions fit more nicely, and suggest a slightly sublinear process, with $a = 0.85$ being the most plausible exponent, with the exception of the case, where a newly created address initiates a transaction; in this case, $a = 1$ gives better fit.

4.2.1 Limited Lifetime Edges

We repeated the procedure of calculating the transformed ranks for variants of the transaction networks where edges are assumed to have limited lifetimes, i.e., 1 day or 30 days. This means that indegrees of nodes can decrease in the case when edges are



removed. Detailed results are shown in the **Supplementary Figures S1–S7**. These results are highly consistent with what we have obtained for the fully time aggregated network, showing an evidence of preferential attachment as well. Best fitting exponents are very similar in all cases for Bitcoin, while for Ethereum addresses, we see slightly higher exponents for short time intervals, hinting at a preference for addresses that already were the target of high activity recently, suggesting a phenomenon where recency has an importance in determining transaction dynamics (Dorogovtsev and Mendes, 2000a; Dorogovtsev and Mendes, 2000b; Nsour and Sayama, 2021).

4.2.2 Evaluating Changes in Exponents Over Time

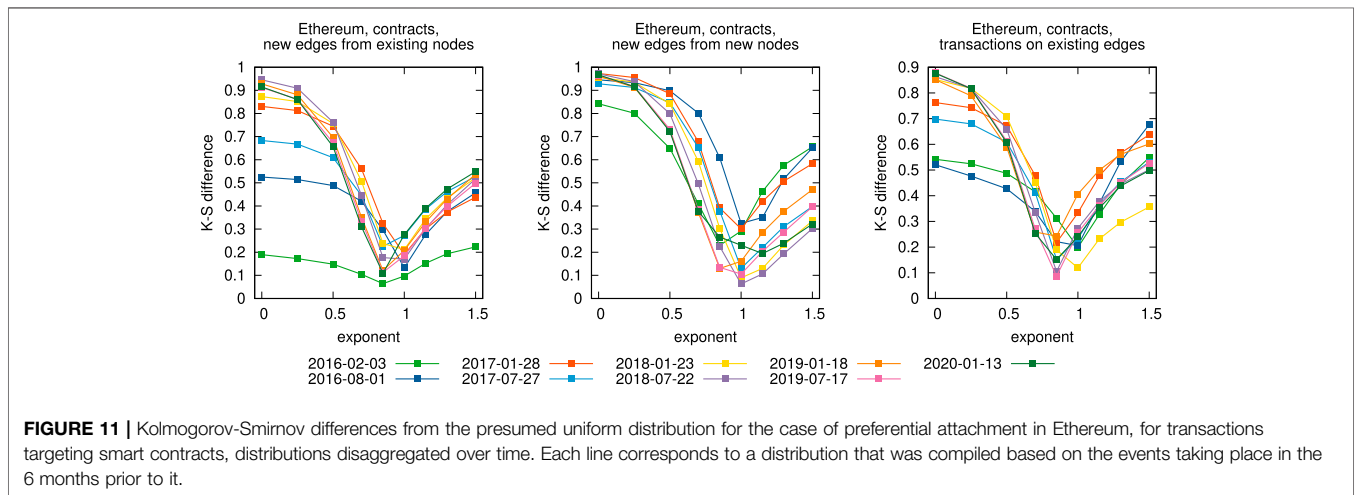
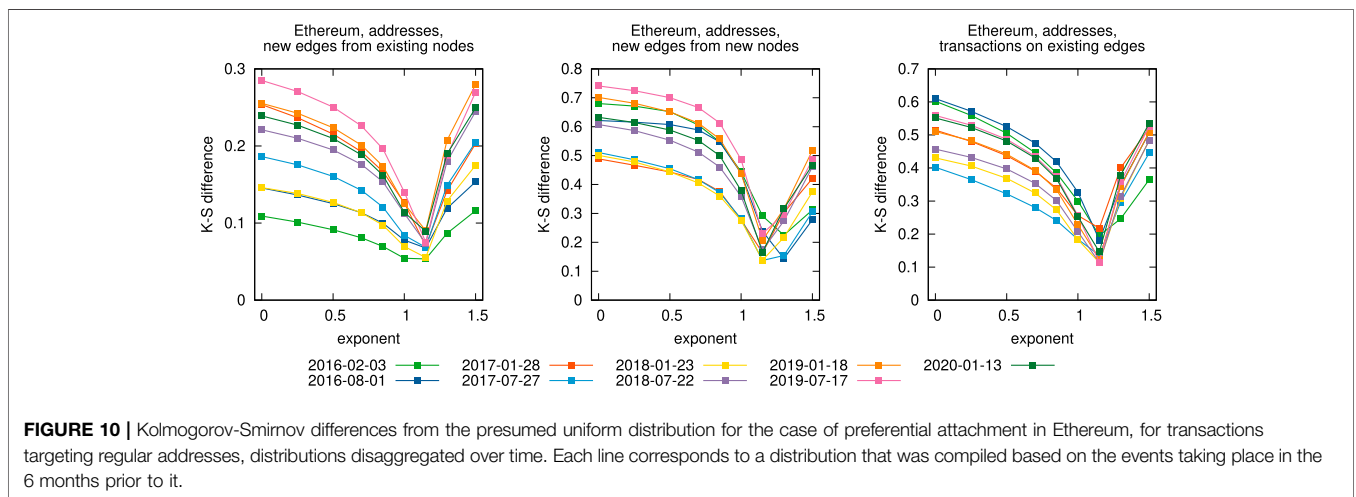
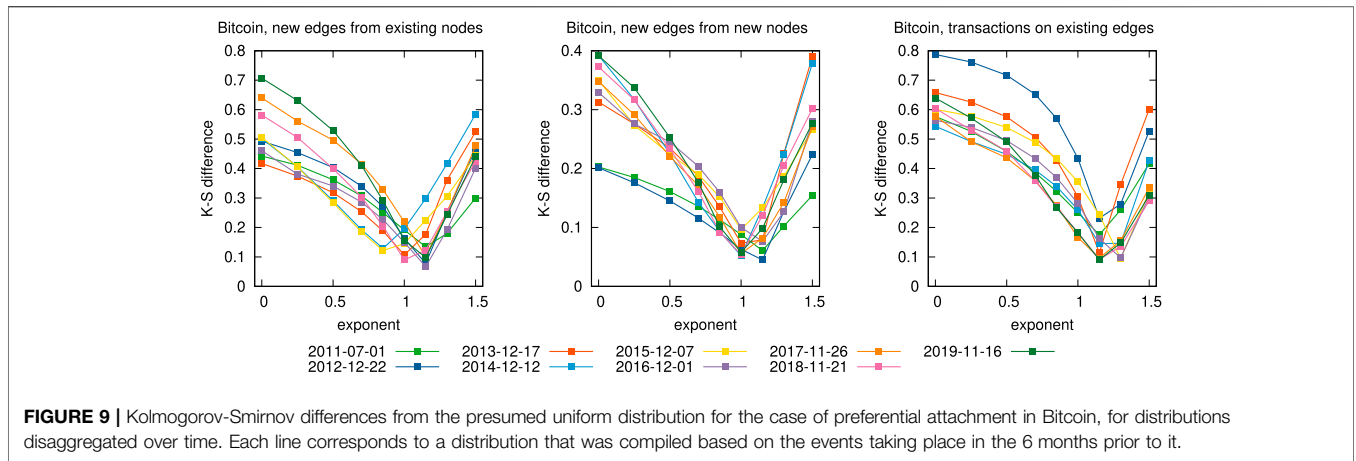
So far, we have evaluated statistics of preferential attachment in a time-aggregated fashion, i.e., we considered all transactions that happened over the lifetime of the cryptocurrency network when looking at the distribution of transformed ranks. To gain more insights into the process of network evolution, we evaluated the distribution of transformed ranks in shorter, half-year long time intervals, and show the Kolmogorov-Smirnov distances as a function of exponents in **Figures 9–11**. We see that while the best fit is achieved around the typical value of exponents as found previously (see **Figures 6, 8**), there is some noticeable variation, with some time periods showing slightly smaller or larger exponents as best fits. This hints that there might be important time-dependent processes shaping the evolution of the transaction networks beyond preferential attachment, as also evidenced by the deviations of the perfect fit of the transformed rank distributions.

5 DISCUSSION

Our results confirm that preferential attachment is a key component shaping the evolution of cryptocurrency transaction networks, contributing to the heavy-tailed degree distributions that arise. This is true regardless of the time scale considered, as focusing only on the subnetworks of recent transaction partners results in very similar statistics of edge creation and activity. While our previous results showed the presence of preferential attachment in the early Bitcoin network, it is remarkable that the same dynamic is present over a much longer time period that involved an almost 100-fold growth in terms of network size and several up- and downturns in the market.

Findings of preferential attachment and heavy-tailed degree distributions matches well with other findings about networks that describe interactions between complex and self-organizing social, technological or economical phenomena (Kullmann and Kertész, 2001; Albert and Barabási, 2002). It is also consistent with the picture of cryptocurrency networks being made up of a few very large players interacting with regular users who have limited activity, especially when considered on the level of individual addresses (Di Francesco Maesa et al., 2018; Fischer et al., 2021; Liu et al., 2021).

We note that there are several limitations in our current study. Firstly, we performed our analysis on the level of individual addresses and not attempted to infer the correspondence between addresses and users or other entities. While peculiarities of Bitcoin, such as the UTXO network model, present multiple



heuristics for this (Fischer et al., 2021), the situation with Ethereum is more complex; also, the presence of mixing services is a complicating factor for Bitcoin as well (Bonneau

et al., 2015; Heilman et al., 2017; Phetsouvanh et al., 2019; Fischer et al., 2021). Considering this limitation, we find the structural complexity emerging in our study especially remarkable.

Furthermore, while in the current work we established a relation among indegrees and the probability of attracting new connections, the situation in reality might be more complex. Even though the indegree of an address is a property that can be computed by anyone based on publicly available data, we do not assume that users actually base decisions directly on it. At the same time, in a statistical sense, indegree seems to be an efficient proxy for a number of latent properties of network nodes that determine their attractiveness for new transaction partners. While in early works about preferential attachment, node degrees were typically assumed to be a direct measure of node size and thus a determinant of “popularity” (Barabási and Albert, 1999; Barabási et al., 1999; Krapivsky et al., 2000; Dorogovtsev and Mendes, 2000a), recent works have looked at models where a more complex property determines the attractiveness of nodes (Pham et al., 2015; Pham et al., 2016; Topirceanu et al., 2018; Nsour and Sayama, 2021). Our work could be readily extended to consider other properties of addresses that can be extracted from blockchain data to uncover a more complete understanding of underlying transaction dynamics.

Our work suggests several future directions for research. Firstly, while we find that preferential attachment is consistently present in all of the studied networks over their lifetime, our results hint that the detailed dynamics of the process (as represented by the best fitting exponent, and also the shape of the distribution of transformed ranks) changes over time (see **Figures 9–11**). A more in-depth investigation of these changes could lead to new insights about different phases of cryptocurrency usage and how it is linked to structural properties of the transaction network.

Second, while the overall trend of preferential attachment is quite clear, there are systematic deviations from a perfect fit to the presumed form (**Eq. 2**). It is a question whether these could be explained by modifying the functional form or extending it to include readily available properties of nodes. Research in this direction could uncover more detailed driving forces of transaction network evolution and provide new, generalizable models of network growth (Naglić and Šubelj, 2019).

Finally, depending on availability of datasets, a comparison between cryptocurrencies and other types of economical or financial transaction networks could inform about the generalizability of our findings and also help in better understanding the role cryptocurrencies play in the

global economy (Alabi, 2017; Begušić et al., 2018; Seebacher and Maleshkova, 2018), a still widely debated subject. To facilitate further research, we publish the data and code used in the current work (Kondor et al., 2020; Kondor, 2020a; Kondor, 2020b; Kondor et al., 2021).

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: <https://doi.org/10.5281/zenodo.4543269> <https://doi.org/10.5061/dryad.qz612jmcf>.

AUTHOR CONTRIBUTIONS

DK, GV, and IC contributed to conception and design of the study. DK contributed software. NB, JS performed data collection and preprocessing. DK, NB, and JS analyzed data. DK performed further data analysis and drafted the paper. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

This research was supported by the Ministry of Innovation and Technology and the National Research, Development and Innovation Office within the Quantum Information National Laboratory of Hungary.

This research is supported by the Singapore Ministry of National Development and the National Research Foundation, Prime Minister’s Office, under the Singapore-MIT Alliance for Research and Technology (SMART) programme.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: <https://www.frontiersin.org/articles/10.3389/fbloc.2021.668510/full#supplementary-material>

REFERENCES

- Akcora, C. G., Dey, A. K., Gel, Y. R., and Kantarcioglu, M. (2018). “Forecasting Bitcoin Price with Graph,” in Pacific-Asia Conference on Knowledge Discovery and Data Mining, Melbourne, VIC, Australia, June 3–6, 2018 (Springer International Publishing), 765–776. doi:10.1007/978-3-319-93040-4
- Alabi, K. (2017). Digital Blockchain Networks Appear to Be Following Metcalfe’s Law. *Electron. Commerce Res. Appl.* 24, 23–29. doi:10.1016/j.elerap.2017.06.003
- Albert, R., and Barabási, A.-L. (2000). Topology of Evolving Networks: Local Events and Universality. *Phys. Rev. Lett.* 85, 5234–5237. doi:10.1103/physrevlett.85.5234
- Albert, R., and Barabási, A.-L. (2002). Statistical Mechanics of Complex Networks. *Rev. Mod. Phys.* 74, 47–97. doi:10.1103/revmodphys.74.47
- Anderson, L., Holz, R., Ponomarev, A., Rimba, P., and Weber, I. (2016). New Kids on the Block: An Analysis of Modern Blockchains. arXiv preprint.
- Aspembitova, A., Feng, L., Melnikov, V., and Chew, L. Y. (2019). Fitness Preferential Attachment as a Driving Mechanism in Bitcoin Transaction Network. *PLoS ONE* 14, e0219346. doi:10.1371/journal.pone.0219346
- Barabási, A.-L., and Albert, R. (1999). Emergence of Scaling in Random Networks. *Science* 286, 509–512. doi:10.1126/science.286.5439.509
- Barabási, A.-L., Albert, R., and Jeong, H. (1999). Mean-Field Theory for Scale-Free Random Networks. *Physica A: Stat. Mech. its Appl.* 272, 173–187. doi:10.1016/S0378-4371(99)00291-5
- Baur, D. G., Hong, K., and Lee, A. D. (2018). Bitcoin: Medium of Exchange or Speculative Assets? *J. Int. Financial Markets, Institutions Money* 54, 177–189. doi:10.1016/j.intfin.2017.12.004
- Begušić, S., Kostanjčar, Z., Eugene Stanley, H., and Podobnik, B. (2018). Scaling Properties of Extreme Price Fluctuations in Bitcoin Markets. *Physica A: Stat. Mech. its Appl.* 510, 400–406. doi:10.1016/j.physa.2018.06.131

- Bonneau, J., Miller, A., Clark, J., Narayanan, A., Kroll, J. A., and Felten, E. W. (2015). "SoK: Research Perspectives and Challenges for Bitcoin and Cryptocurrencies," in IEEE Symposium on Security and Privacy, San Jose, CA, United States, May 17–21, 2015, 104–121. doi:10.1109/SP.2015.14
- Caprolu, M., Pontecorvi, M., Signorini, M., Segarra, C., and Di Pietro, R. (2021). A Novel Framework for the Analysis of Unknown Transactions in Bitcoin: Theory, Model, and Experimental Results. arXiv preprint 2103.09459.
- Chen, L., Xu, L., Shah, N., Diallo, N., Gao, Z., Lu, Y., and Shi, W. (2017). "Unraveling Blockchain Based Crypto-Currency System Supporting Oblivious Transactions," in BCC 2017 - Proceedings of the ACM Workshop on Blockchain, Cryptocurrencies and Contracts, Abu Dhabi, United Arab Emirates, April 2, 2017 (co-located with ASIA CCS 2017), 23–28. doi:10.1145/3055518.3055528
- Clauset, A., Shalizi, C. R., and Newman, M. E. J. (2009). Power-Law Distributions in Empirical Data. *SIAM Rev.* 51, 661–703. doi:10.1137/070710111
- Conti, M., Sandeep Kumar, E., Lal, C., and Ruj, S. (2018). A Survey on Security and Privacy Issues of Bitcoin. *IEEE Commun. Surv. Tutorials* 20, 3416–3452. doi:10.1109/COMST.2018.2842460
- Di Francesco Maesa, D., Marino, A., and Ricci, L. (2018). Data-Driven Analysis of Bitcoin Properties: Exploiting the Users Graph. *Int. J. Data Sci. Anal.* 6, 63–80. doi:10.1007/s41060-017-0074-x
- Dixon, M. F., Akcora, C. G., Gel, Y. R., and Kantarcioglu, M. (2019). Blockchain Analytics for Intraday Financial Risk Modeling. *Digit. Finance* 1, 67–89. doi:10.1007/s42521-019-00009-8
- Dorogovtsev, S. N., and Mendes, J. F. F. (2000). Evolution of Networks with Aging of Sites. *Phys. Rev. E* 62, 1842–1845. doi:10.1103/PhysRevE.62.1842
- Dorogovtsev, S. N., and Mendes, J. F. F. (2000). Scaling Behaviour of Developing and Decaying Networks. *Europhys. Lett.* 52, 33–39. doi:10.1209/epl/i2000-00400-0
- Dorogovtsev, S. N., and Mendes, J. F. F. (2001). Language as an Evolving Word Web. *Proc. R. Soc. Lond. B* 268, 2603–2606. doi:10.1098/rspb.2001.1824
- Dorogovtsev, S. N., and Mendes, J. F. F. (2001). Effect of the Accelerating Growth of Communications Networks on Their Structure. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* 63, 025101. doi:10.1103/PhysRevE.63.025101
- Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. (2000). WWW and Internet Models from 1955 till Our Days and the "Popularity Is Attractive" Principle. arXiv preprint cond-mat/0009090
- Dorogovtsev, S. N., Mendes, J. F. F., and Samukhin, A. N. (2000). Structure of Growing Networks with Preferential Linking. *Phys. Rev. Lett.* 85, 4633–4636. doi:10.1103/physrevlett.85.4633
- Ferretti, S., and D'Angelo, G. (2020). On the Ethereum Blockchain Structure: A Complex Networks Theory Perspective. *Concurrency Comput.* 32, e5493. doi:10.1002/cpe.5493
- Fischer, J. A., Palechor, A., Dell'Aglio, D., Bernstein, A., and Tessone, C. J. (2021). The Complex Community Structure of the Bitcoin Address Correspondence Network. arXiv preprint 2105.09078. doi:10.3389/fphy.2021.681798
- Guo, D., Dong, J., and Wang, K. (2019). Graph Structure and Statistical Properties of Ethereum Transaction Relationships. *Inf. Sci.* 492, 58–71. doi:10.1016/j.ins.2019.04.013
- Gurcan, C., Dixon, M. F., Gel, Y. R., and Kantarcioglu, M. (2018). Bitcoin Risk Modeling with Blockchain Graphs. *Econ. Lett.* 173, 138–142. doi:10.1016/j.econlet.2018.07.039
- Heilman, E., Alshenibr, L., Baldimtsi, F., Scafuro, A., and Goldberg, S. (2017). "TumbleBit: An Untrusted Bitcoin-Compatible Anonymous Payment Hub," in Proceedings 2017 Network and Distributed System Security Symposium, San Diego, CA, United States, February 26–March 1, 2017. doi:10.14722/ndss.2017.23086
- Jeong, H., Néda, Z., and Barabási, A. L. (2003). Measuring Preferential Attachment in Evolving Networks. *Europhys. Lett.* 61, 567–572. doi:10.1209/epl/i2003-00166-9
- Kiffer, L., Levin, D., and Mislove, A. (2018). "Analyzing Ethereum's Contract Topology," in Internet Measurement Conference 2018, Boston, MA, United States, October 31–November 2, 2018, 494–499. doi:10.1145/3278532.3278575
- Kondor, D., Pósfai, M., Csabai, I., and Vattay, G. (2014). Do the Rich Get Richer? An Empirical Analysis of the Bitcoin Transaction Network. *PLoS ONE* 9, e86197. doi:10.1371/journal.pone.0086197
- Kondor, D., Csabai, I., Szüle, J., Pósfai, M., and Vattay, G. (2014). Inferring the Interplay between Network Structure and Market Effects in Bitcoin. *New J. Phys.* 16, 125003. doi:10.1088/1367-2630/16/12/125003
- Kondor, D., Pósfai, M., Csabai, I., and Vattay, G. (2020). *Bitcoin Transaction Network*. Dryad. [Dataset]. doi:10.5061/dryad.qz612jmcf
- Kondor, D., Bulatovic, N., Stéger, J., Csabai, I., and Vattay, G. (2021). *Ethereum Transaction Network*. [Dataset]. doi:10.5281/zenodo.4543269
- Kondor, D. (2020). Generalized Order-Statistic Tree Implementation. [Dataset] Available at: <https://github.com/dkondor/orbtree> (Accessed February 24, 2021).
- Kondor, D. (2020). Order-statistic Tree with Example Code for Preferential Attachment Testing. [Dataset] Available at: https://github.com/dkondor/patest_new (Accessed February 24, 2021).
- Krapivsky, P. L., Redner, S., and Leyvraz, F. (2000). Connectivity of Growing Random Networks. *Phys. Rev. Lett.* 85, 4629–4632. doi:10.1103/PhysRevLett.85.4629
- Kullmann, L., and Kertész, J. (2001). Preferential Growth: Exact Solution of the Time-dependent Distributions. *Phys. Rev. E Stat. Nonlin Soft Matter Phys.* 63, 051112. doi:10.1103/PhysRevE.63.051112
- Kunegis, J., Blattner, M., and Moser, C. (2013). "Preferential Attachment in Online Networks: Measurement and Explanations," in Proceedings of the 5th Annual ACM Web Science Conference WebSci '13, Paris, France, May 2–4, 2013, 205–214. doi:10.1145/2464464.2464514
- Kurbucz, M. T. (2019). Predicting the price of Bitcoin by the Most Frequent Edges of its Transaction Network. *Econ. Lett.* 184, 108655. doi:10.1016/j.econlet.2019.108655
- Liang, J., Li, L., and Zeng, D. (2018). Evolutionary Dynamics of Cryptocurrency Transaction Networks: An Empirical Study. *PLoS ONE* 13, e0202202. doi:10.1371/journal.pone.0202202
- Liu, X. F., Jiang, X.-J., Liu, S.-H., and Tse, C. K. (2021). Knowledge Discovery in Cryptocurrency Transactions: A Survey. *IEEE Access* 9, 37229–37254. doi:10.1109/access.2021.3062652
- Meiklejohn, S., Pomarole, M., Jordan, G., Levchenko, K., McCoy, D., Voelker, G. M., and Savage, S. (2013). "A Fistful of Bitcoins," in Proceedings of the Internet Measurement Conference - IMC '13, Barcelona, Spain, October 23–25, 2013, 127–140. doi:10.1145/2504730.2504747
- Miller, A., Moeser, M., Lee, K., and Narayanan, A. (2017). An Empirical Analysis of Linkability in the Monero Blockchain. arXiv preprint 1704.04299.
- Motamed, A. P., and Bahrak, B. (2019). Quantitative Analysis of Cryptocurrencies Transaction Graph. *Appl. Netw. Sci.* 4, 131. doi:10.1007/s41109-019-0249-6
- Naglić, L., and Šubelj, L. (2019). War Pact Model of Shrinking Networks. *PLoS ONE* 14, e0223480. doi:10.1371/journal.pone.0223480
- Nepusz, T. (2020). Plfit – Fitting Power-Law Distributions to Empirical Data. [Dataset] Available at: <https://github.com/ntamas/plfit> (Accessed January 10, 2021).
- Nick, D. J. (2015). Data-Driven De-Anonymization in Bitcoin. Ph.D. thesis. Zürich, Switzerland: ETH Zürich. doi:10.3929/ethz-a-010541254
- Nsour, F., and Sayama, H. (2021). Hot-Get-Richer Network Growth Model. *Stud. Comput. Intelligence* 944, 532–543. doi:10.1007/978-3-030-65351-4_43
- Oggier, F., Datta, A., and Phetsouvanh, S. (2020). An Ego Network Analysis of Sextortionists. *Soc. Netw. Anal. Min.* 10, 44. doi:10.1007/s13278-020-00650-x
- Paquet-Clouston, M., Haslhofer, B., and Dupont, B. (2019). Ransomware Payments in the Bitcoin Ecosystem. *J. Cybersecurity* 5, 1–11. doi:10.1093/cybsec/tyz003
- Perc, M. (2014). The Matthew Effect in Empirical Data. *J. R. Soc. Interf.* 11, 20140378. doi:10.1098/rsif.2014.0378
- Pham, T., Sheridan, P., and Shimodaira, H. (2015). PAFit: A Statistical Method for Measuring Preferential Attachment in Temporal Complex Networks. *PLoS ONE* 10, e0137796. doi:10.1371/journal.pone.0137796
- Pham, T., Sheridan, P., and Shimodaira, H. (2016). Joint Estimation of Preferential Attachment and Node Fitness in Growing Complex Networks. *Sci. Rep.* 6, 32558. doi:10.1038/srep32558
- Phetsouvanh, S., Datta, A., and Oggier, F. (2019). Analysis of Multi-Input Multi-Output Transactions in the Bitcoin Network. *Concurrency Comput.* 33, 1–20. doi:10.1002/cpe.5629
- Portnoff, R. S., Huang, D. Y., Doerfler, P., Afroz, S., and McCoy, D. (2017). "Backpage and Bitcoin," in Proceedings of the ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, Halifax, NS, Canada, August 13–17, 2017, 1595–1604. Part F129685. doi:10.1145/3097983.3098082

- Ron, D., and Shamir, A. (2013). "Quantitative Analysis of the Full Bitcoin Transaction Graph," in 17th Int. Conf. Financial Cryptogr. Data Secur., Okinawa, Japan, April 1–5, 2013, 6–24. doi:10.1007/978-3-642-39884-1_2
- Seebacher, S., and Maleshkova, M. (2018). "A Model-Driven Approach for the Description of Blockchain Business Networks," in Proceedings of the 51st Hawaii International Conference on System Sciences, Waikoloa Village, HI, United States, January 2–6, 2018. doi:10.24251/hicss.2018.442
- Seres, I. A., Gulyás, L., Nagy, D. A., and Burcsi, P. (2020). "Topological Analysis of Bitcoin's Lightning Network," in 1st International Conference MARBLE 2019 (Santorini, Greece: Springer), 1–12. doi:10.1007/978-3-030-37110-4_1
- Simon, H. A. (1955). On a Class of Skew Distribution Functions. *Biometrika* 42, 425–440. doi:10.2307/2333389
- Topircceanu, A., Udrescu, M., and Marculescu, R. (2018). Weighted Betweenness Preferential Attachment: A New Mechanism Explaining Social Network Formation and Evolution. *Sci. Rep.* 8, 10871. doi:10.1038/s41598-018-29224-w
- Victor, F., and Lüders, B. K. (2019). "Measuring Ethereum-Based ERC20 Token Networks," in 23rd International Conference on Financial Cryptography and Data Security, February 18–22, 2019 (Frigate Bay, St. Kitts and Nevis: Springer), 113–129. doi:10.1007/978-3-030-32101-7_8
- Wood, G. (2014). Ethereum: A Secure Decentralised Generalised Transaction Ledger. *Ethereum project yellow paper* 151, 1–32.
- Wu, J., Liu, J., Zhao, Y., and Zheng, Z. (2020). Analysis of Cryptocurrency Transactions from a Network Perspective: An Overview. arXiv preprint 2011.09318.
- Yli-Huumo, J., Ko, D., Choi, S., Park, S., and Smolander, K. (2016). Where Is Current Research on Blockchain Technology?—A Systematic Review. *PLoS ONE* 11, e0163477. doi:10.1371/journal.pone.0163477
- Yule, G. U. (1925). II.-A Mathematical Theory of Evolution, Based on the Conclusions of Dr. J. C. Willis, F. R. S. *Phil. Trans. R. Soc. Lond. B* 213, 21–87. doi:10.1098/rstb.1925.0002
- Zheng, Z., Xie, S., Dai, H. N., Chen, X., and Wang, H. (2016). Blockchain Challenges and Opportunities: A Survey. *Int. J. Web Grid Serv.* 14 (4), 1–24. doi:10.1504/IJWGS.2018.095647
- Zwang, M., Somin, S., Pentland, A. S., and Altshuler, Y. (2018). Detecting Bot Activity in the Ethereum Blockchain Network. arXiv preprint 1810.01591.

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Copyright © 2021 Kondor, Bulatovic, Stéger, Csabai and Vattay. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.