

MIT Open Access Articles

Remaining popular: power-law regularities in network dynamics

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

Citation: EPJ Data Science. 2022 Dec 15;11(1):61

As Published: <https://doi.org/10.1140/epjds/s13688-022-00373-3>

Publisher: Springer Berlin Heidelberg

Persistent URL: <https://hdl.handle.net/1721.1/146908>


Version: Final published version: final published article, as it appeared in a journal, conference proceedings, or other formally published context

Terms of use: Creative Commons Attribution





Remaining popular: power-law regularities in network dynamics

Shahar Somin^{1,2}, Yaniv Altshuler², Alex ‘Sandy’ Pentland² and Erez Shmueli^{1,2*} 

*Correspondence:

¹Department of Industrial Engineering, Tel Aviv University, Tel Aviv, Israel

²MIT Media Lab, Cambridge, MA, USA

Abstract

The structure of networks has been a focal research topic over the past few decades. These research efforts have enabled the discovery of numerous structural patterns and regularities, bringing forth advancements in many fields. In particular, the ubiquitous power-law patterns evident in degree distributions, graph eigenvalues and human mobility patterns have provided the opportunity to model many different complex systems. However, regularities in the dynamical patterns of networks remain a considerably less explored terrain. In this study we examine the dynamics of networks, focusing on stability characteristics of node popularity, and present our results using various empirical datasets. Specifically, we address several intriguing questions – for how long are popular nodes expected to remain so? How much time is expected to pass between two consecutive popularity periods? What characterizes nodes which manage to maintain their popularity for long periods of time? Surprisingly, we find that such temporal aspects are governed by a power-law regime, and that these power-law regularities are equally likely across all node ages.

1 Introduction

The study of complex systems and their structure has seen a growing interest in the past few decades. Discovering the existence of seemingly ubiquitous meta-structures such as the power-law patterns evident in degree distributions [1–3], graph eigenvalues [4] and human mobility patterns [5] has heralded the use of a network oriented approach for modeling, analyzing and predicting the macroscopic and mesoscopic behavior of “real-world” systems in a myriad of everyday fields and applications.

Equally important is the quest for patterns and regularities in network *dynamics*, since these could be used for analyzing and predicting the dynamics within a broad range of domains. To date research of network dynamics has focused on three main categories. The first and the most studied one is the spreading dynamics over a static network structure [6–11]. The second prevalent research direction involved analyzing the dynamics of individual-level user activity, for instance by establishing inter-event power-law distributions [12–17]. The third line of studies entailed exploring network dynamics-related characteristics, on a system-level perspective. These include shrinking diameters and network densification patterns [18], spectral evolution [19], community formation dynamics [20–22] and system-level bursty dynamics [23, 24].

© The Author(s) 2022. **Open Access** This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

This study pertains to the third category, aiming at examining regularities in network dynamics, while focusing on network stability patterns, as are manifested through the popularity¹ of nodes. In principle, throughout the network's lifetime, non popular nodes may become popular and popular nodes might lose their status. In [25] we suggest a theoretical modeling which might explain such popularity changes. In the heart of this study, we analyze temporal aspects of node popularity periods, such as the time-span for which a "popular node" is expected to maintain its popularity status, the number of consecutive popularity periods per node, and the time it takes for a node to regain its popularity after losing it. We show clear statistical regularities with regards to all aforementioned processes, in the form of an adherence to a power-law model, across various and distinct empirical datasets. We further show that such power-law regularities are equally likely across all node ages.

In order to provide a complete view of this phenomena, we also examine two generative models and assess their ability to account for our empirical findings. Specifically, we inspect the prevalent Barabasi-Albert (BA) model (employing two parameter constellations) and the Temporal Preferential Attachment (TPA) model, which accounts for the system's aging processes. We find that while many of the dynamic patterns are also captured by the BA model, it fails to accurately account for the characteristics of highly popular nodes, for all different examined constellations. In particular, while empirical evidence advocates that long-term popularity applies to all node ages, the BA model is highly biased towards early-joiners (i.e. nodes which have joined the network at its early stages). The TPA model brings forth some advancement, as it is able to better qualitatively reproduce age-related phenomena, however the low statistical significance of its results implies for the necessity of further research. Our findings may shed a new light on node ranking dynamics, enhancing the understanding of node popularity shifts on the one hand and their fortification on the other, regardless of node ages.

2 Materials and methods

2.1 Data

In this study we analyze empirical datasets from three distinct domains, as elaborated below:

2.1.1 Amazon ranking dataset

The *Amazon Product Rankings* dataset [26] contains product reviews and metadata from Amazon, including 142.8 million reviews spanning August 1997–July 2006. This dataset includes reviews (ratings, text, helpfulness votes), product metadata (descriptions, category information, price, brand, and image features), and links (also viewed/also bought graphs). We construct weekly bipartite temporal networks, containing product ratings from the Amazon online shopping website. In each such temporal network, nodes represent users and products. An edge between a user and a product is formed if the user rated the specific product within the given timespan. Previous studies that used this dataset for the modeling of network properties can be found for example in [27–30]

¹Throughout this paper, a node is referred to as popular if it is amongst the top-degree nodes in the network.

2.1.2 ERC20 Ethereum blockchain ledger

Launched in July 2015 [31], the Ethereum Blockchain is a public ledger that keeps records of all Ethereum related transactions. The ability of the Ethereum Blockchain to store not only ownership, similarly to Bitcoin, but also execution code, in the form of “*Smart Contracts*”, has recently led to the creation of a large number of new types of “tokens”, based on the Ethereum ERC20 protocol. These tokens are “minted” by a variety of players, for a variety of reasons, having all of their transactions carried out by their corresponding Smart Contracts, publicly accessible on the Ethereum Blockchain. As a result, the ERC20 ecosystem constitutes a fascinating example of a highly varied financial ecosystems whose entire activity is publicly available from its inception. This dataset was used in several network theory related studies [32–34] including financial assets adoption [35] and Malware and BOTs detection [36].

In order to preserve anonymity in the Ethereum Blockchain, personal information is omitted from all transactions. A user, represented by their wallet, can participate in the economy system through an address, which is obtained by applying *Keccak-256* hash function on his public key. The Ethereum Blockchain enables users to send transactions in order to either send Ether to other wallets, create new Smart Contracts or invoke any of their functions. Since Smart Contracts are scripts residing on the Blockchain as well, they are also assigned a unique address. A Smart Contract is called by sending a transaction to its address, which triggers its independent and automatic execution, in a prescribed manner on every node in the network, according to the *data* that was included in the triggering transaction.

We have retrieved all transactions spanning from February 2016 to January 2019, resulting in 179,488,619 transactions, performed by 27,888,847 unique wallets, trading 79,451 distinct tokens. We construct weekly bipartite temporal networks, containing crypto-tokens transactions on top of the Ethereum Blockchain. Nodes represent trading *wallets* and crypto-tokens. An edge between a wallet and a token is formed if the wallet bought or sold the given token in the examined timespan.

2.1.3 eToro financial trading dataset

The financial transaction data used in this work was received from an online social financial trading platform for foreign exchanges, equity indices and commodities, called *eToro* [37, 38]. This trading platform allows traders to take both long and short positions, with a minimal bid of as low as a few dollars, thus providing access for retail traders to investment activities that until recently were only available for professional investors. A key feature of eToro’s “social trading” platform is that each trader can easily see the complete trading history of other investors. Investors can then set their accounts to copy one or more trades made by any other investors, in which case the social trading platform will automatically execute the trade(s). Accordingly, there are three types of trades: (i) Single (or non-social) trade: Investor A places a normal trade by himself or herself; (ii) Copy trade: Investor A places exactly the same trade as investor B’s single trade; (iii) Mirror trade: Investor A automatically executes Investor B’s every single trade, i.e., Investor A follows exactly investor B’s trading activities (and implicitly their investment decisions). Both (ii) and (iii) are hereafter referred to as social trading, and can be regarded as decision making that is based on information received through the common social medium.

The data that was analyzed for this work encompasses approximately 3 million registered accounts, containing over 40 million trades during a period of 3 years. We construct

weekly temporal networks, based on the *mirroring* activity of users on top of this platform.² The nodes represent traders (followers and followees). An edge (v_1, v_2) between two traders is formed if trader v_1 mirrored the trading activity of trader v_2 . Previous studies of this dataset can be found in [39–46].

2.2 Methods

2.2.1 Temporal networks

We define *temporal networks* as follows.

Definition 2.1 The temporal graph for a given timestamp t , $G_t(V_t, E_t)$ is the directed graph constructed from all transactions performed during the time period $[t - \Delta, t)$. The set of vertices V_t consists of all entities participating in the network activity during that period:

$$V_t := \{v \mid v \text{ participates in network activity during } [t - \Delta, t)\} \quad (1)$$

and the set of edges $E_t \subseteq V_t \times V_t$ is defined as:

$$E_t := \{(u, v) \mid \text{An activity between } u \text{ and } v \text{ occurred during } [t - \Delta, t)\}. \quad (2)$$

The temporal degree of a vertex $v \in V_t$ is defined as:

$$\text{deg}_t(v) = |\{u \in V_t \mid (u, v) \in E_t\}|. \quad (3)$$

Definition 2.2 Given a time-stamp $t \in [T]$, the rank assigned to node $v \in V_t$, according to its degree $\text{deg}_t(v)$, is denoted by $\text{rank}_t(v)$. The ranking is performed in a descending order, such that the rank of 1 corresponds to the highest degree node in V_t . Specifically, $\text{rank}_t(v) = r$ if there are $r - 1$ nodes with a higher degree than v :

$$\text{rank}_t(v) = |\{u \in V_t \text{ such that } \text{deg}_t(u) > \text{deg}_t(v)\}| + 1. \quad (4)$$

Ties are broken randomly, by a random internal ranking of groups containing identical degree nodes. Given a threshold \hat{T} , a node v will be referred to as popular if its rank is amongst the top- \hat{T} nodes:

$$1 \leq \text{rank}_t(v) \leq \hat{T}. \quad (5)$$

2.2.2 Fitting a heavy-tailed model

In this work, we find that the examined distributions present heavy-tailed patterns. In order to substantiate this hypothesis and determine the exact model best representing these distributions, we compared four plausible heavy-tailed models:

1. The power-Law model: $P(k) = k^{-\alpha}$
2. The Truncated Power-Law model: $P(k) = k^{-\alpha} \cdot e^{-\lambda k}$

²We chose mirroring actions as the underlying edges in the network building process, since they imply strong trust in user B for user A. In particular, users who are mirrored and followed the most are likely to be the best traders, highly reflecting their popularity [39].

3. The Exponential model: $P(k) = \lambda e^{-\lambda k}$
4. The Lognormal model: $\frac{1}{k\sigma} \exp\left(-\frac{(\ln(x)-\mu)^2}{2\sigma^2}\right)$

To this end we applied a prevalent statistical framework [1] encompassing two main stages. Namely, given a heavy tailed model M :

Stage 1: calculate goodness of fit

1. Fit the empirical data to the given model M , estimate its parameters P_M using maximum likelihood estimators, and calculate the Kolmogorov-Smirnov (KS) statistic³ for this fit – KS_M .
2. Calculate the goodness-of-fit between the data and the examined model. This stage incorporates:
 - (a) Generate $n = 1000$ synthetic data sets with the P_M parameters.
 - (b) Fit each synthetic data set individually to its own model M and calculate the KS statistic KS_i for each one relative to its own model.
 - (c) Calculate the p-value, being the fraction of times that KS_i ($\forall i \in n$) is larger than KS_M .
3. If the resulting p-value was greater than 0.05 the model M was considered to be a plausible hypothesis for the data, otherwise it was rejected.

Stage 2: compare plausible models Compare all plausible models which were not rejected in the previous step using a likelihood ratio test. The log likelihood ratio test calculates the likelihood of the given data between two competing distributions. The logarithm of this ratio is positive or negative depending on which model presents a better fit, or is zero if a tie is obtained. The sign of the log likelihood ratio is subject to statistical instability and when close to zero, the fluctuations can change its sign. In order to establish the statistical significance of the log likelihood ratio sign, we calculate its standard deviation and corresponding p-value, where small p-values indicate that the established sign is a reliable estimate of model compatibility.

2.2.3 Barabasi-Albert (BA) model simulations

Introduced in 1999 [47], the BA Model was based on the discovery that a common property of many large networks is that vertex connectedness follows a scale-free power-law distribution. This feature appears generically in expanding networks where new vertices attach preferentially to already well connected sites. The proposed model managed to reproduce various stationary scale-free distributions, indicating that the development of large networks is governed by robust generic self-organizing phenomena that are agnostic to the particularities of the examined system. The BA Model has served as the basis of numerous studies in various scientific fields, including social networks analysis [9, 48, 49], computer communication networks [50], biological systems [51], transportation [52, 53], IOT [54], emergency detection [55], financial trading systems [40, 41, 44] and many others.

In order to substantiate our empirical findings, we examine the dynamics established by the Barabasi-Albert (BA) model. First we generate a Barabasi-Albert scale-free network $G(V, E)$ over $n = 100,000$ nodes. A single node is added at every iteration, each outputting

³The KS statistic is the maximum distance between the CDFs of the data and the fitted model.

Algorithm 1 Tagged Barabasi-Albert Scale-Free Network (n, m)**Input** n : Number of requested nodes. m : Number of edges to attach from each new node to existing nodes.**Output** A scale-free network, with edges tagged with the iteration number at which they were added to the network.

```

1:  $G \leftarrow \mathbf{Graph}(\text{nodes} = m)$            {Create a graph with  $m$  nodes and zero edges.}
2: for  $i = m + 1 : n$  do
3:   addNode( $G, v_i$ )
4:   for  $j = 1 : m$  do
5:      $u \leftarrow \mathbf{selectNode}(G)$            {A target node is chosen with a probability
                                                    proportional to its degree.}
6:     addEdge( $G, (v_i, u), \text{tag} = i$ )       {An edge is added and tagged
                                                    with the iteration number.}
7:   end for
8: end for
9: return  $G$ 

```

$m = 20$ edges, which are preferentially attached to existing nodes, proportionally to their degree. Algorithm 1 depicts this standard Barabasi-Albert process for scale-free network generation [47], with the minor alteration of tagging the added edges with the iteration at which they were added to the network. Next, in order to construct temporal networks from $G(V, E)$ we follow def. 2.1 with varying Δ choices:

1. $\Delta = 100$
2. $\Delta \sim \text{Norm}(\mu, \sigma)$ for $\mu = 100, \sigma = 20$

Specifically, each temporal network $G_t(V_t, E_t)$ is composed of all edges whose tags are within the following range:

$$E_t = \{(e, \text{tag}) \in E \mid \text{tag} \in [t - \Delta, t]\} \quad (6)$$

and all the nodes that participated in these iterations:

$$V_t = \{u \in V \mid (u, v) \in E_t \text{ or } (v, u) \in E_t\}. \quad (7)$$

2.2.4 Trendy preferential attachment (TPA) model simulations

The Trendy Preferential Attachment model [56] is a forget-based extension to the BA model. It presents a network evolution model where edges become less influential as they age. The diminishing influence is modeled by a monotonically decreasing function $f(\tau)$ of their age τ . We have chosen to apply $f(\tau) \propto 1/(\tau^2)$. As such, the probability of a new node to connect to another node v in time t is proportional to its *time-weighted degree* as follows:

$$\widehat{\text{deg}}_t(v) = f(1) \cdot \text{deg}_{t-1}(v) + f(2) \cdot \text{deg}_{t-2}(v) + \dots + f(t-1) \cdot \text{deg}_1(v), \quad (8)$$

where $\text{deg}_t(v)$ is the actual (not time-weighted) degree of node v at time t .

We start by generating a TPA network $G(V, E)$ over $n = 100,000$ nodes. A single node is added at every iteration, each outputting $m = 20$ edges, which are preferentially attached to existing nodes, proportionally to their time-weighted degree. Similarly to the process we have performed with the BA model, we tag each edge with the iteration at which it was added to the network. Next, in order to construct temporal networks from $G(V, E)$ we follow def. 2.1 with $\Delta = 100$.

3 Results

3.1 Empirical analysis

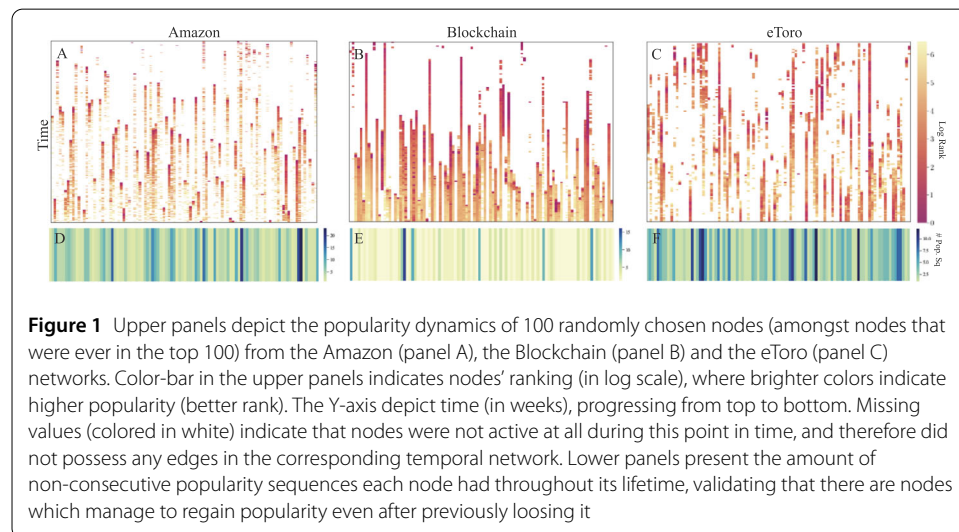
We examine the dynamics of nodes popularity levels over time, as measured by their degree-based rankings. In particular, for each temporal graph G_t we rank all nodes in V_t according to their associated in-degree, in a descending order (consider def. 2.2 for further details) and examine these ranks over time. Table 1 presents a description of the ranked items in each dataset.

We start by presenting the popularity dynamics of several randomly chosen nodes from three real-world datasets, as qualitatively depicted in Fig. 1. We observe that nodes' popularity periods vary considerably in length, and that certain nodes can regain high levels of popularity even after massive drops in popularity.

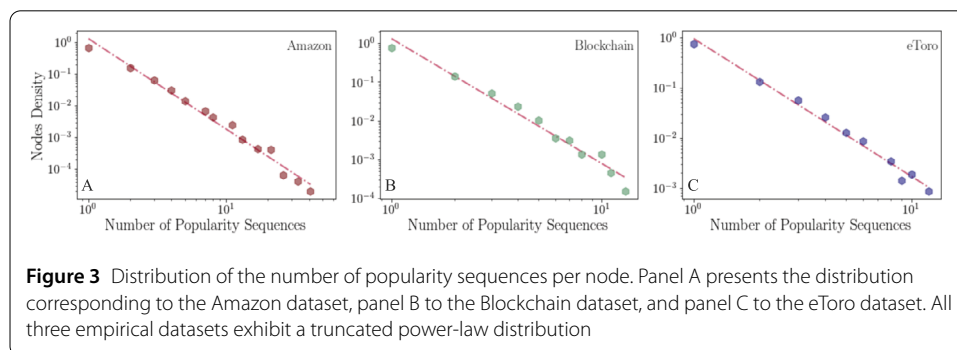
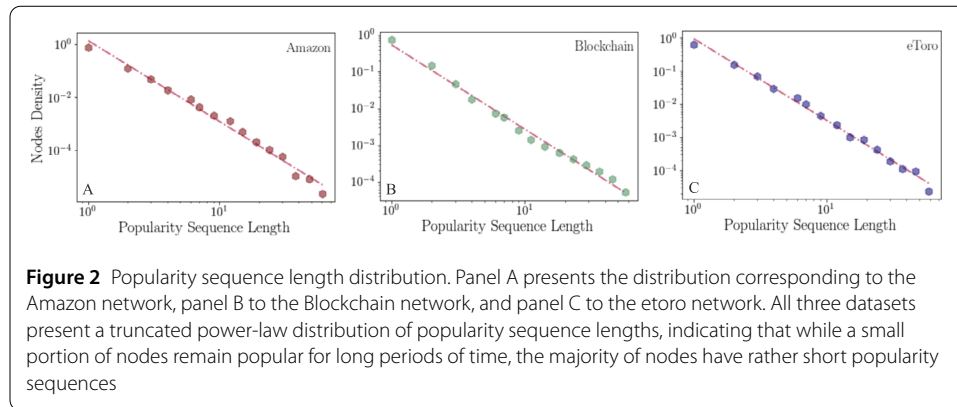
In the rest of this section, we examine the dynamics of popularity from a system perspective. For a given temporal graph, a node is considered to be popular if it was among the top 100 highest degree nodes.⁴ We start by examining the distribution of node popularity

Table 1 Dataset Network Representation

Dataset	Edges	Ranked Nodes	Popularity Significance
Amazon ratings	Users \times Products	Products	Amount of users who ranked the product.
Ethereum Blockchain	Traders \times Crypto-tokens	Crypto-tokens	Amount of traders who bought the token.
eToro: Social Trading Platform	Traders \times Traders	Mirrored traders	Amount of traders who mirrored user's trading activity.



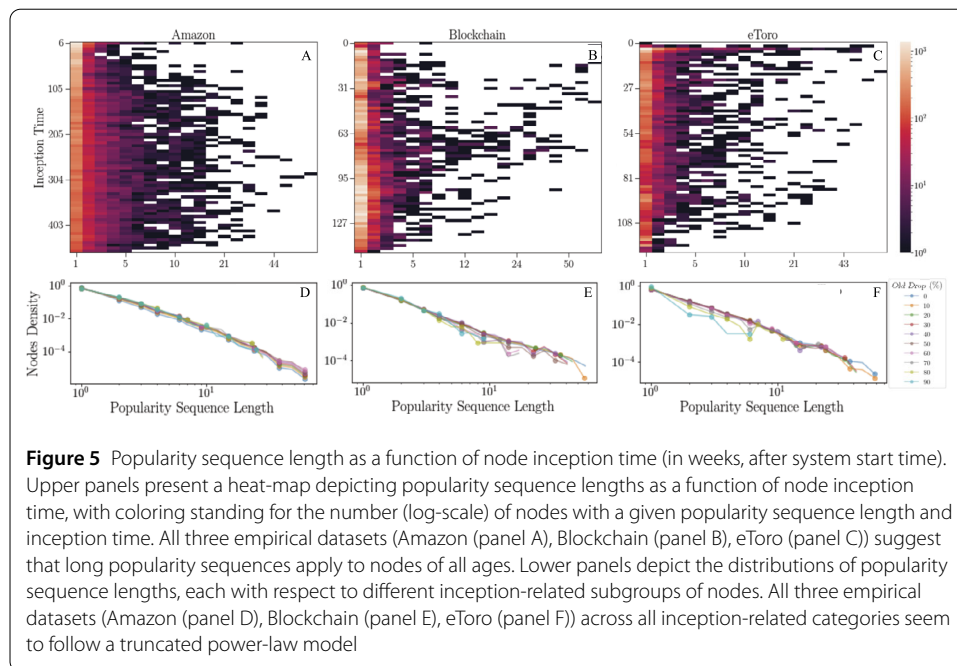
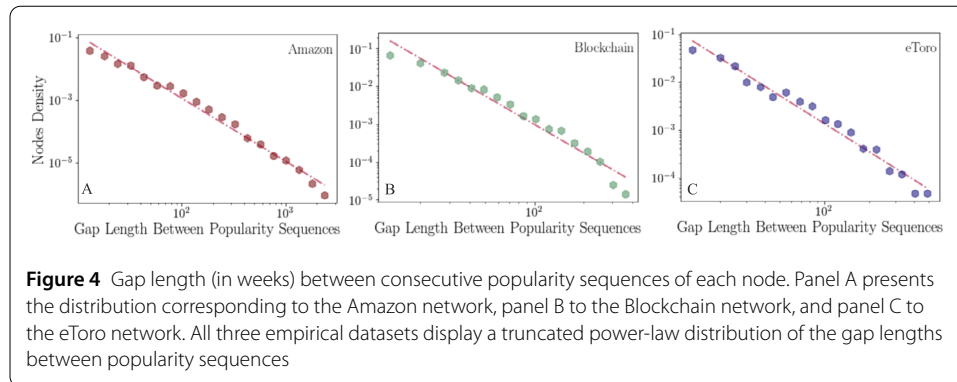
⁴See Additional file 1, Sect. 6.1 for other thresholds $\hat{T} \in (10, 25, 50)$ analyses.



sequence lengths, where the popularity sequence of a node is defined as the consecutive time-span for which it was popular. As depicted in Fig. 2, all three empirical datasets (Amazon (panel A), Blockchain (panel B) and eToro (panel C)) present a truncated power-law distribution of popularity sequence lengths (See Supportive statistical analysis in Tables 2, 3, Additional file 1). This result is rather surprising, as one might expect that once a node manages to join the “most popular list” it would maintain its popularity status for long periods of time. Real-world systems however do not abide by these rules. Instead, we note that the vast majority of popular nodes remain popular only for short periods of time and only a minority of nodes manage to preserve their popularity status for long periods of time.

Next, we analyze the distribution of the number of popularity sequences per node. As presented in Fig. 3, all three empirical datasets follow a truncated power-law distribution (See supportive statistical analysis in Tables 2, 3, Additional file 1). Interestingly, this implies that the vast majority of nodes, after losing their popularity status, will remain non-popular. However, there are the selected few who manage to regain popularity over and over.

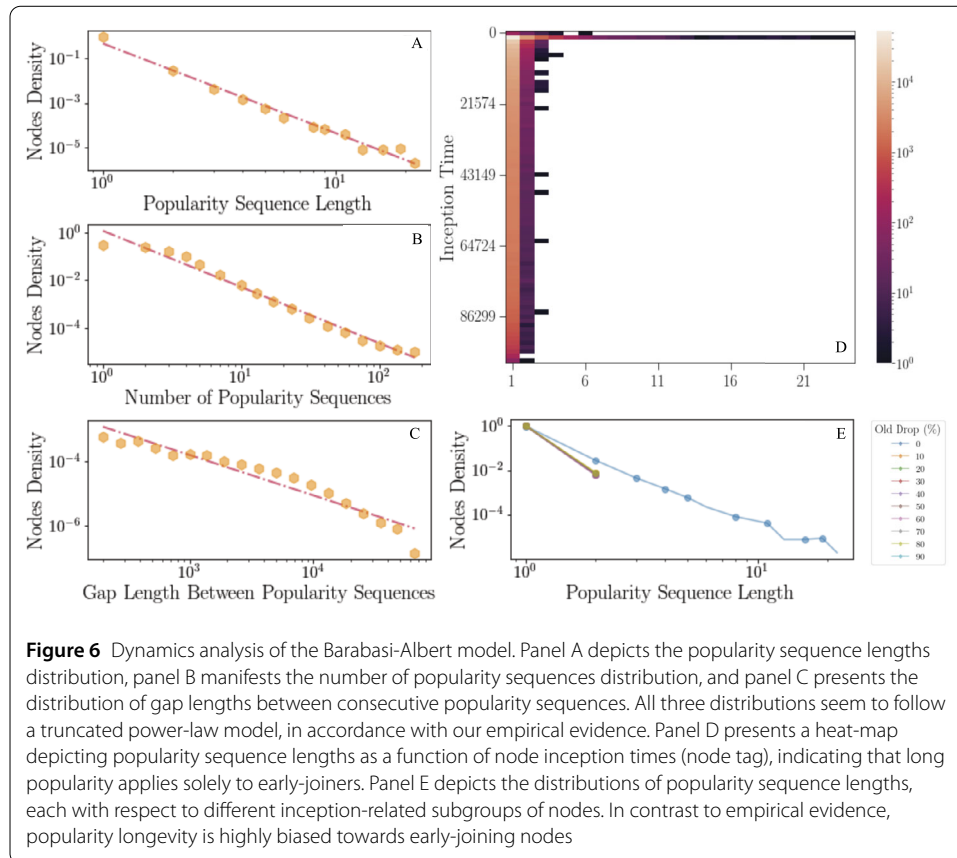
We also examine the distribution of gap lengths (in weeks) between consecutive popularity sequences of each node. As depicted in Fig. 4, all three empirical datasets (Amazon (panel A), Blockchain (panel B) and eToro (panel C)) follow a truncated power-law distribution (See supportive statistical analysis in Tables 2, 3, Additional file 1). This result suggests that most nodes which manage to regain their popularity, do so after short periods of time. Nonetheless, and somewhat counter-intuitively, few nodes manage to ‘resurrect’ and become popular again even after a rather long time.



Finally, we are interested in examining the characteristics of long-term popular nodes. Specifically, Fig. 5 depicts the connection between popularity sequence lengths and node “inception” times (the time on which a node was first introduced to the network). We first observe that long-term popularity has an approximately uniform spread across nodes of all ages, for the examined datasets (panels A–C). Furthermore, we find that the scale-free nature of the distribution of popularity sequence lengths (as exhibited in Fig. 2) is also “age-free” (panels D–F). Namely, even when removing $X \in \{0\%, 10\%, \dots, 90\%\}$ of the oldest nodes from the network, the popularity distribution associated with the remaining sub-network still follows a truncated power-law model (see 6.4, Additional file 1 for statistical support).

3.2 Generative models analysis

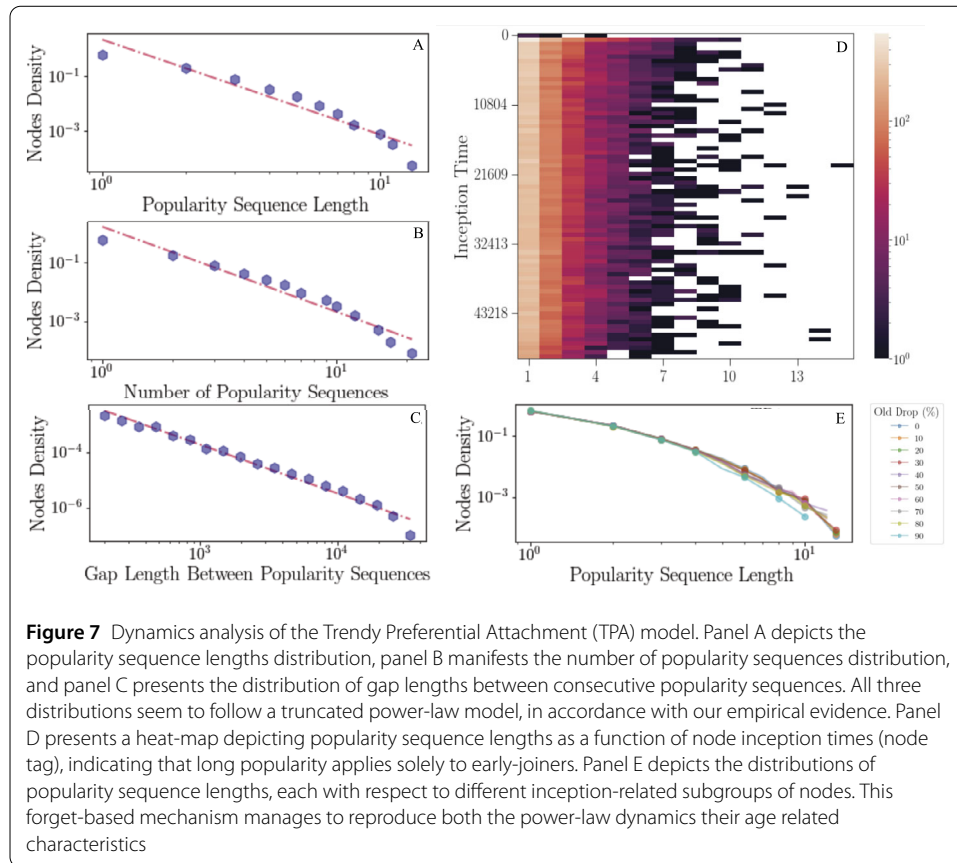
We next examine the popularity dynamics established by two network-evolution models. We start by exploring the well-known Barabasi-Albert (BA) model. This model, being one of the most prevalent and well-studied models for network evolution, was the first to ac-



count for the formation of power-law patterns in network *structure*⁵ (namely, heavy-tailed degree distributions). It is therefore interesting to verify the extent to which it manages to reproduce the power-law distributions observed in the network *dynamics* and their temporal characteristics. To this end, we construct temporal networks from a BA model simulation (see Methods section for details). Interestingly, we note that the BA model succeeds in capturing all three distributions (panels A-C, Fig. 6). However, it fails to reproduce the connection between nodes' popularity and their inception times, as seen in our empirical examples. In particular, the BA model is highly biased towards early-joiners becoming popular for long periods of time (panel D, Fig. 6) and its inception-time related distributions do not generally follow a power-law distribution (panel E, Fig. 6). Consider Sect. 6.4, Additional file 1, for supportive statistical analysis and Sect. 6.2 for a further analysis of a different BA temporal network configuration.

We continue by analyzing the dynamics established by a forget-based extension to the BA model. We believe that a mechanism which allows recent activity to have a heavier impact on the edge attachment process, might prevent the heavy popularity tilt towards early-joining nodes. In particular, we analyze the Trendy-Preferential Attachment model (TPA) [56] which presents a network evolution model where edges become less influential as they age. We construct temporal networks from a TPA model simulation (see Methods section for further details). Interestingly, we note that such forget-based mechanism

⁵Consider Fig. 9, Additional file 1 which shows that the temporal degree distributions of the empirical datasets are indeed heavy-tailed.



manages to qualitatively reproduce both the power-law distributions and their age characteristics (Fig. 7). However, when employing GOF analysis to the results, we find that both the dynamical distributions and the age-related distributions obtain rather low statistical significance for their power-law fit (see Sect. 6.4, Additional file 1.)

4 Discussion

In this study, we examined various characteristics related to the dynamics of node popularity in networks. In particular, we have analyzed the lengths of time periods for which nodes attain high popularity, the number of such periods per node, and the distribution of time gaps between two such consecutive periods. We have shown that truncated power-law patterns accurately describe these characteristics of network dynamics within three distinct empirical datasets, providing what may be the first evidence for these particular power-law regularities in network dynamics. We further examined the characteristics of long-term popular nodes. We show that across all three examined datasets, node tendency towards long popularity periods is not affected by their joining time to the network. Furthermore, we found that this scale-free property is also “age-free”, as the power-law distribution is evident across all age categories.

While the Barabasi-Albert (BA) model manages to capture some of these dynamics-related characteristics, it fails to accurately account for the connection between popularity dynamics and node ages. In particular, it shows a considerable bias towards early-joiners, having long popularity periods, in sharp contrast to the real-world networks we have examined. It is important to note however, that the BA temporal networks in our simulations

differ from real-world temporal networks since they are restricted to the exact number edges, resulting in upper bounded temporal degrees. Further research is required in order to fully comprehend the effect this restriction has on the established results. Nevertheless, a preliminary analysis we have performed (consider Sect. 6.2, Additional file 1 for further details) examined the effect of Gaussian noise added to the amount edges each temporal network consists of, and presented results consistent with the original BA specifications. We speculate that the origins of this mismatch between the BA model and the empirical evidence are rooted in the likelihood of popular nodes to be of a given age. Indeed, while the BA framework is heavily skewed towards early-joining popular nodes, the empirical datasets exhibit a roughly uniform distribution of inception times among popular nodes (see supportive analysis in Sect. 6.5, Additional file 1). This suggests different forces are behind the empirically established power-law distributions.

Employing a forget-based extension of the BA model (TPA), we found that it is able to qualitatively reproduce the examined dynamical patterns, and has a better agreement with the age characteristics of popularity. Nevertheless, the low statistical significance of its results suggests the need of further research in order to understand the forces and mechanisms behind the observed dynamics and their age-related characteristics. Such efforts might include examining other recent network evolution models [56–60] and developing new generative models to account for these findings.

Furthermore, since the presented analysis was focused on economy-related datasets, it is intriguing to verify whether the established regularities are a specific characterization of economical networks, or whether they actually describe any social network, regardless of its domain. Accompanied by the increasing availability of temporal empirical data, these research directions could enable much deeper understanding of dynamical regularities, and impact domains ranging from biology to social science.

Supplementary information

Supplementary information accompanies this paper at <https://doi.org/10.1140/epjds/s13688-022-00373-3>.

Additional file 1. Supplementary information (PDF 1.5 MB)

Acknowledgements

Not applicable.

Funding

Not applicable.

Availability of data and materials

The datasets used and analysed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests

The authors declare that they have no competing interests.

Author contribution

SS performed the analysis. SS, ES and YA interpreted the data. SS, ES, YA and SP contributed to writing the manuscript. All authors read and approved the final manuscript.

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 16 May 2022 Accepted: 28 November 2022 Published online: 15 December 2022

References

1. Clauset A, Shalizi CR, Newman ME (2009) Power-law distributions in empirical data. *SIAM Rev* 51(4):661–703
2. Adamic LA, Huberman BA, Barabási A, Albert R, Jeong H, Bianconi G (2000) Power-law distribution of the world wide web. *Science* 287(5461):2115
3. Redner S (1998) How popular is your paper? An empirical study of the citation distribution. *Eur Phys J B, Condens Matter Complex Syst* 4(2):131–134
4. Faloutsos M, Faloutsos P, Faloutsos C (1999) On power-law relationships of the Internet topology. *Comput Commun Rev* 29(4):251–262
5. Gonzalez MC, Hidalgo CA, Barabasi A-L (2008) Understanding individual human mobility patterns. *Nature* 453(7196):779–782
6. Artime O, Ramasco JJ, San Miguel M (2017) Dynamics on networks: competition of temporal and topological correlations. *Sci Rep* 7(1):1–10
7. Lloyd AL, May RM (2001) How viruses spread among computers and people. *Science* 292(5520):1316–1317
8. Leskovec J, McGlohon M, Faloutsos C, Glance N, Hurst M (2007) Patterns of cascading behavior in large blog graphs. In: Proceedings of the 2007 SIAM international conference on data mining. SIAM, Philadelphia, pp 551–556
9. Barrat A, Barthelemy M, Vespignani A (2008) Dynamical processes on complex networks. Cambridge University Press, Cambridge
10. Barthélemy M, Barrat A, Pastor-Satorras R, Vespignani A (2005) Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *J Theor Biol* 235(2):275–288
11. Karsai M, Kivela M, Pan RK, Kaski K, Kertész J, Barabási A-L, Saramäki J (2011) Small but slow world: how network topology and burstiness slow down spreading. *Phys Rev E* 83(2):025102
12. Barabasi A-L (2005) The origin of bursts and heavy tails in human dynamics. *Nature* 435(7039):207–211
13. Vázquez A, Oliveira JG, Dezső Z, Goh K-I, Kondor I, Barabási A-L (2006) Modeling bursts and heavy tails in human dynamics. *Phys Rev E* 73(3):036127
14. Dewes C, Wichmann A, Feldmann A (2003) An analysis of Internet chat systems. In: Proceedings of the 3rd ACM SIGCOMM conference on Internet measurement, pp 51–64
15. Kleban SD, Clearwater SH (2003) Hierarchical dynamics, interarrival times, and performance. In: SC'03: proceedings of the 2003 ACM/IEEE conference on supercomputing. IEEE, Los Alamitos, pp 28–28
16. Candia J, González MC, Wang P, Schoenharl T, Madey G, Barabási A-L (2008) Uncovering individual and collective human dynamics from mobile phone records. *J Phys A, Math Theor* 41(22):224015
17. Dezső Z, Almaas E, Lukács A, Rácz B, Szakadát I, Barabási A-L (2006) Dynamics of information access on the web. *Phys Rev E* 73(6):066132
18. Leskovec J, Kleinberg J, Faloutsos C (2005) Graphs over time: densification laws, shrinking diameters and possible explanations. In: Proceedings of the eleventh ACM SIGKDD international conference on knowledge discovery in data mining, pp 177–187
19. McGlohon M, Akoglu L, Faloutsos C (2011) Statistical properties of social networks. In: Social network data analytics. Springer, Berlin, pp 17–42
20. Backstrom L, Huttenlocher D, Kleinberg J, Lan X (2006) Group formation in large social networks: membership, growth, and evolution. In: Proceedings of the 12th ACM SIGKDD international conference on knowledge discovery and data mining, pp 44–54
21. Kumar R, Novak J, Tomkins A (2010) Structure and evolution of online social networks. In: Link mining: models, algorithms, and applications. Springer, Berlin, pp 337–357
22. McGlohon M, Akoglu L, Faloutsos C (2008) Weighted graphs and disconnected components: patterns and a generator. In: Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining, pp 524–532
23. Rybski D, Buldyrev SV, Havlin S, Liljeros F, Makse HA (2012) Communication activity in a social network: relation between long-term correlations and inter-event clustering. *Sci Rep* 2(1):1–8
24. Karsai M, Jo H-H, Kaski K et al (2018) Bursty human dynamics. Springer, Berlin
25. Somin S, Altshuler Y, 'Sandy' Pentland A, Shmueli E (2022) Beyond preferential attachment: falling of stars and survival of superstars. *R Soc Open Sci* 9(8):220899
26. (2016) Amazon ratings network dataset – KONECT
27. Lim E-P, Nguyen V-A, Jindal N, Liu B, Lauw HW (2010) Detecting product review spammers using rating behaviors. In: Proc. Int. conf. on information and knowledge management, pp 939–948
28. Mishra M, Chopde J, Shah M, Parikh P, Babu RC, Woo J (2019) Big data predictive analysis of Amazon product review. In: KSIIT the 14th Asia Pacific international conference on information science and technology (APIC-IST), pp 141–147
29. Woo J, Mishra M (2021) Predicting the ratings of Amazon products using big data. *Wiley Interdiscip Rev Data Min Knowl Discov* 11(3):e1400
30. Haque TU, Saber NN, Shah FM (2018) Sentiment analysis on large scale Amazon product reviews. In: 2018 IEEE international conference on innovative research and development (ICIIRD). IEEE, Los Alamitos, pp 1–6
31. Buterin V et al (2014) A next-generation smart contract and decentralized application platform. In: White paper
32. Victor F, Lüders BK (2019) Measuring Ethereum-based ERC20 token networks. In: International conference on financial cryptography and data security. Springer, Berlin, pp 113–129
33. Somin S, Altshuler Y, Gordon G, Shmueli E et al (2020) Network dynamics of a financial ecosystem. *Sci Rep* 10(1):1–10
34. Somin S, Gordon G, Pentland A, Shmueli E, Altshuler Y (2020) Network dynamics of a tokenized financial ecosystem. In: Building the new economy, 0 edn. vol 4. <https://wip.mitpress.mit.edu/pub/dnb7e62x>
35. Morales AJ, Somin S, Altshuler Y, Pentland A (2020) User behavior and token adoption on ERC20. [arXiv:2005.12218](https://arxiv.org/abs/2005.12218)
36. Zwang M, Somin S, Pentland AS, Altshuler Y (2018) Detecting bot activity in the ethereum blockchain network <https://www.etoro.com/>
37. <https://www.etoro.com/>
38. Assia Y (2016) Etoro—building the world's largest social investment network. In: The FinTech book: the financial technology handbook for investors, entrepreneurs and visionaries, pp 196–197

39. Pan W et al (2015) Reality hedging: social system approach for understanding economic and financial dynamics. PhD thesis, Massachusetts Institute of Technology
40. Altshuler Y, Pan W, Pentland A (2012) Trends prediction using social diffusion models. In: International conference on social computing, behavioral-cultural modeling and prediction. Springer, Berlin, pp 97–104
41. Shmueli E, Altshuler Y et al (2014) Temporal dynamics of scale-free networks. In: International conference on social computing, behavioral-cultural modeling, and prediction. Springer, Berlin, pp 359–366
42. Altshuler Y, Pentland AS, Gordon G (2015) Social behavior bias and knowledge management optimization. In: Social computing, behavioral-cultural modeling, and prediction. Springer, Berlin, pp 258–263
43. Liu Y-Y, Nacher JC, Ochiai T, Martino M, Altshuler Y (2014) Prospect theory for online financial trading. *PLoS ONE* 9(10):e109458
44. Pan W, Altshuler Y, Pentland A (2012) Decoding social influence and the wisdom of the crowd in financial trading network. In: Privacy, security, risk and trust (PASSAT), 2012 international conference on and 2012 international conference on social computing (SocialCom). IEEE, Los Alamitos, pp 203–209
45. Dorfleitner G, Fischer L, Lung C, Willmeringer P, Stang N, Dietrich N (2018) To follow or not to follow—an empirical analysis of the returns of actors on social trading platforms. *Q Rev Econ Finance* 70:160–171
46. Krafft PM, Shmueli E, Griffiths TL, Tenenbaum JB et al (2021) Bayesian collective learning emerges from heuristic social learning. *Cognition* 212:104469
47. Barabási A-L, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512
48. Newman ME (2003) The structure and function of complex networks. *SIAM Rev* 45(2):167–256
49. Newman ME (2005) Power laws, Pareto distributions and Zipf's law. *Contemp Phys* 46(5):323–351
50. Pastor-Satorras R, Vespignani A (2007) Evolution and structure of the Internet: a statistical physics approach. Cambridge University Press, Cambridge
51. Barabasi A-L, Oltvai ZN (2004) Network biology: understanding the cell's functional organization. *Nat Rev Genet* 5(2):101–113
52. Shmueli E, Mazeh I, Radaelli L, Pentland AS, Altshuler Y (2015) Ride sharing: a network perspective. In: International conference on social computing, behavioral-cultural modeling, and prediction. Springer, Berlin, pp 434–439
53. Altshuler Y, Puzis R, Elovici Y, Bekhor S, Pentland AS (2015) On the rationality and optimality of transportation networks defense: a network centrality approach. In: Securing transportation systems, pp 35–63
54. Altshuler Y, Fire M, Aharony N, Elovici Y, Pentland A (2012) How many makes a crowd? On the correlation between groups' size and the accuracy of modeling. In: International conference on social computing, behavioral-cultural modeling and prediction. Springer, Berlin, pp 43–52
55. Altshuler Y, Fire M, Shmueli E, Elovici Y, Bruckstein A, Pentland AS, Lazer D (2013) The social amplifier—reaction of human communities to emergencies. *J Stat Phys* 152(3):399–418
56. Mokryn O, Wagner A, Blattner M, Ruppig E, Shavitt Y (2016) The role of temporal trends in growing networks. *PLoS ONE* 11(8):e0156505
57. Bianconi G, Barabási A-L (2011) Competition and multiscaling in evolving networks. In: The structure and dynamics of networks. Princeton University Press, Princeton, pp 361–367
58. Dorogovtsev SN, Mendes JFF, Samukhin AN (2000) Structure of growing networks with preferential linking. *Phys Rev Lett* 85(21):4633–4636
59. Leskovec J, Backstrom L, Kumar R, Tomkins A (2008) Microscopic evolution of social networks. In: Proceedings of the 14th ACM SIGKDD international conference on knowledge discovery and data mining, pp 462–470
60. Oikonomou P, Cluzel P (2006) Effects of topology on network evolution. *Nat Phys* 2(8):532–536

Submit your manuscript to a SpringerOpen[®] journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

Submit your next manuscript at ► [springeropen.com](https://www.springeropen.com)
