

Scalability and Egalitarianism in peer-to-peer networks

Fabio Caccioli, Giacomo Livan and Tomaso Aste

Department of Computer Science, University College London

Gower Street, London, WC1E 6BT, UK

and Systemic Risk Centre, London School of Economics and Political Sciences,

London, WC2A2AE, UK

April 22, 2016

1 Introduction

We are currently witnessing the rise of a new form of distributed economy, which emerges from the combination of digital communication infrastructures and the big data revolution. Peer-to-peer decentralized economies and finance have the potential to provide citizens with direct control over their activities, by removing intermediation layers and fostering inclusion. Blockchain is providing the technology to make this happen in a secure and reliable way. Distributed systems are being constructed around an egalitarian ethos according to which “peers” freely exchange goods and information without the need of a central authority to establish trust, verify identity, or enforce commitments. Yet, we are witnessing that many of such idealistic egalitarian forms of economic organization are changing their nature as they evolve. In fact, these systems show a strong tendency to naturally evolve towards structures where a small portion of nodes has a large influence over the whole system. This has been for instance observed in the evolution of the mining pools in bitcoin, where the system has evolved from a fairly egalitarian network of miners, in which individuals were able to mine their coins at home, to highly specialized and concentrated industrial-scale mining activities. A similar evolution has been observed in the world wide web, which started from a distributed community of people and companies, and evolved into a highly centralized system where. For instance, Facebook owns 1.49 billion

active users profiles [1] and 99.9% of web searches in US are run through 5 search engines only, with Google accounting for over 64% [1] of them. This concentration is due to simple economic rules that demand greater efficiency and lower costs. This return-to-scale economic law introduces however new forms of information asymmetry [2] and new kinds of risk related to the presence of very large quantities of personal information held in a few places only. Within the context of distributed systems that generate consensus with majority vote, such tendency towards concentration can be very dangerous.

In this chapter we discuss the relation between the structure of communication network and the functional properties of peer-to-peer systems. In particular, we discuss the relation between level of equality between nodes in the network, and efficiency and scalability.

The chapter is structured as follows: we present in Section 2 a short introduction to complex networks, and we discuss in Section 3 how the properties of information spreading processes depend on the network topology. We then present in section 4 an application to bitcoin blockchain, and we study in particular how the occurrence of blockchain forks is related with the properties of the underlying network. In Chapter 4 we present conclusions.

2 A brief introduction to complex networks

Networks are the most general way to represent systems made of several entities characterized by pairwise interactions. A network is defined in terms of a set of nodes $\{1, 2, \dots, N\}$ and a set of links $\{e_1, e_2, \dots, e_M\}$, where each link connects a pair of nodes. Nodes connected by a link are said to be *neighbors*, and the number of neighbors of a node is the node's *degree*. A convenient way of representing a network is in terms of its adjacency matrix \mathbf{A} , whose element A_{ij} is 1 if node i is connected to node j , and zero otherwise. Each node in a network can interact with its neighbors, that are the nodes connected to it through links.

Several dynamical processes take place on networks. Certain processes, such as searches on the Internet or propagation in power grids, happen on physical networks. Other processes are mediated by non-physical network structures, such as, e.g., the spreading of epidemics or the word of mouth diffusion of information in a social system. In both cases the topological properties and the heterogeneity of the underlying networks play a crucial role in driving the evolution and, possibly, determining the outcome of the dynamics at hand. Countless network topology models, and their impact on dynamical processes, have been studied in the dedicated literature (see for instance [3–6]). In the following, we outline the properties of the two most common

random network classes. In this context, random does not mean chaotic or disorganized. On the contrary, we shall focus on ensembles whose individual realizations all share a number of well defined *statistical* properties. In particular, it has been shown that the degree distribution $P(k)$ of a network, that defines the probability of a node to have degree k , strongly affects the dynamical properties of processes taking place on networks [6–8]. Here we are interested in exploring the effect of the network structure on its capability to efficiently propagate information, which is at the basis of the validation mechanism in blockchain systems and is directly related with efficiency and egalitarianism. We will therefore focus on two classes of networks with very different degree distributions: the Poisson distribution, that is rather equalitarian, and the power-law distribution, that is dis-equalitarian, being characterized by a few hub nodes with very large degrees, and a multitude of other nodes with small degree.

2.1 Erdős-Rényi networks

The Erdős-Rényi (ER) random network model was devised in the late fifties, and it represents the most popular benchmark model for networks featuring *mild heterogeneity*. It consists of N nodes, and each of the $N(N - 1)/2$ pairs of nodes in the network have a fixed probability p to be connected. Clearly, this formation scheme creates, on average, a total number of $pN(N - 1)/2$ links, with an average degree of $\langle k \rangle = p(N - 1)$.

What about individual nodes? The probability that a given node i will be connected to exactly k neighbors (in network jargon this is referred to as node i having degree $k_i = k$) is proportional to the probability of independently “hitting” k nodes (each with probability p) and “missing” the other $N - k - 1$, which is a bimodal distribution:

$$\text{Prob}(k_i = k) = \binom{N - 1}{k} p^k (1 - p)^{N - k - 1} . \quad (1)$$

From the above equation it is straightforward to compute that the probability of hitting a fraction αN of nodes becomes exponentially small in N when $\alpha \rightarrow 1$, i.e. $\text{Prob}(k_i = \alpha N) \propto \exp(\alpha N \log p)$. Taking the large N approximation of equation (1), one can also show that the degree distribution of a large network is very well approximated by a Poisson distribution

$$P(k) = \frac{(pN)^k}{k!} e^{-pN} . \quad (2)$$

The above result means that both the average and the variance of the ER model’s degree distribution are given by pN . All in all, these results justify the previous hint

at mild heterogeneity or, in other words, egalitarianism: in the ER model there are no nodes that dominate the network by linking to a disproportionately large fraction of peers, and the Poisson degree distribution (2) concentrates most of the probability within a limited set of degrees around the average.

Processes on ER networks can reach the entirety of the network by starting from a single node in a number of steps that is proportional to the logarithm of the total number of nodes N . This is called the *small world* effect. Intuitively, given that each node has on average pN neighbors, one expects the number of nodes at a distance d from a given node i to scale as $(pN)^d$: on average, node i has pN neighbors, and so do they and the nodes higher upstream from i . For sufficiently large values of d , the number of nodes found at distance d from i is a finite fraction of all nodes in the network. From this consideration, one finds that the average distance between pairs of nodes in an ER network scales as

$$d \simeq \frac{\log N}{\log(pN)} \simeq \frac{\log N}{\log\langle k \rangle} . \quad (3)$$

Therefore a process evolving on the network from any given node will reach any other node in average after a number $d \propto \log N$ of steps.

2.2 Scale-free networks

Many real-life networks are not egalitarian, and are dominated by a small fraction of hubs connected to a substantial fraction of nodes. In mathematical terms, this is best described in terms of a power law, hence scale-free, distribution of the degrees, i.e.

$$P(k) \sim k^{-\gamma} , \quad (4)$$

the tail index $\gamma > 0$ is a very important parameter. Networks with smaller values of γ have larger hubs and are less egalitarian as we shall see shortly.

Quite interestingly, modeling the topology of scale-free networks requires describing their growth and evolution, in a way which is quite revealing about the mechanisms leading to the emergence of dominant nodes. Such description is provided by the Barabasi-Albert model of *preferential attachment* [9]. Starting from a small group of nodes, the network is built by adding nodes one at a time, and each newcomer connects to one of the already existing nodes. In particular, the new node connects to a given node i with a degree-dependent probability π that reads

$$\pi(k_i) = \frac{k_i}{\sum_{j=1}^M k_j} , \quad (5)$$

where k_i is node i 's degree, and M is the number of nodes present in the network. The preferential attachment growth rule (5) gives a competitive advantage to nodes that already have a high degree, in a rich-get-richer fashion. This mechanism indeed gives rise (for large N) to a power law degree distribution (4), with a tail exponent $\gamma = 3$. Exponents γ with values other than three can be achieved with generalizations of the preferential attachment rule (5) (see, e.g., [10]).

2.3 Network egalitarianism

The egalitarianism, or lack thereof, of a network can be measured directly from its degree sequence in terms of Gini coefficient, i.e. the most popular measure of inequality in a population. Despite having been originally intended and still being mostly used as a measure of wealth inequality, the Gini coefficient can be used to assess how unevenly a generic quantity is distributed across a given population. Let us then consider the sequence k_i , $i = 1, \dots, N$, of degrees in a network. The Gini index G of the sequence is defined as

$$G = \frac{\sum_{i < j} |k_i - k_j|}{N \sum_{i=1}^N k_i}. \quad (6)$$

The above expression can be easily manipulated to show that it corresponds to half the average absolute difference between degrees (normalized to the average degree in the network), i.e. $G = \langle |k_i - k_j| \rangle / (2 \langle k \rangle)$. The Gini coefficient is bounded between zero, which is achieved under complete equality ($k_i = k$, $\forall i$), and one, which is achieved asymptotically under complete concentration on one node (formally $k_i = k$ for a given node i and $k_j = 0$, for all other nodes $j \neq i$, even though this does not correspond to a feasible degree sequence).

An alternative inequality measure is represented by the Theil index, which is akin to an entropy measure and is also bounded between zero (complete equality) and one (complete concentration on one node). It reads

$$T = \frac{1}{N \log N} \sum_{i=1}^N \frac{k_i}{\langle k \rangle} \log \frac{k_i}{\langle k \rangle}. \quad (7)$$

In Figure 1 we show the average behavior of the Gini and Theil indices computed from networks with degree sequences distributed according to Eq. (4).

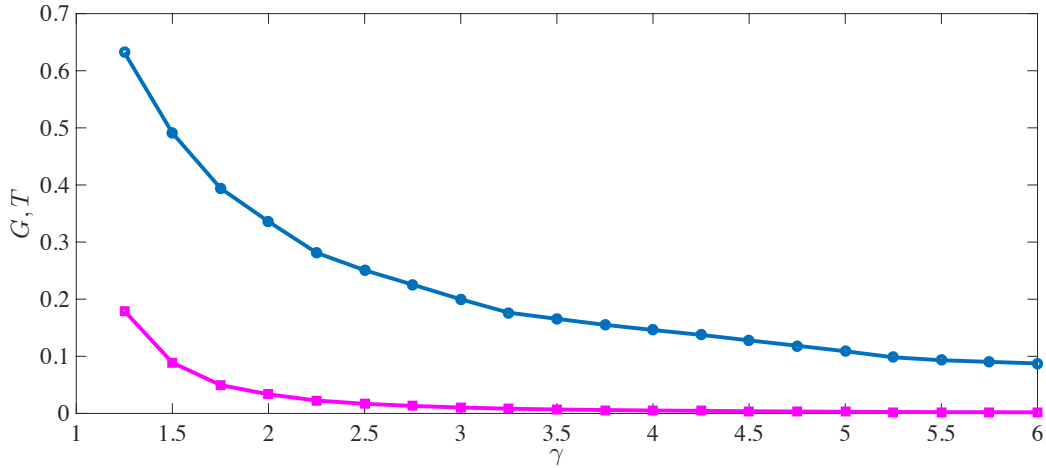


Figure 1: Gini (blue circles) and Theil (purple squares) indices as a function of the tail exponent γ of a power law distributed degree sequence (4). For each value of γ results are obtained by averaging over 100 networks made of 10^4 nodes, with a fixed average degree $\langle k \rangle = 8$ (the value does not affect the behavior of the Gini coefficient).

As Figure 1 reveals, networks with extremely heavy tailed degree sequences ($\gamma \simeq 1.25$) show very large Gini coefficient values ($G \sim 0.65$), comparable to those measured, e.g., for individual wealth distribution in those countries with the highest observed wealth inequality levels. This reflects the strong centralization induced by the degree distribution (5) for low values of γ , which generates a few hubs with $\mathcal{O}(N)$ links, and leaves the vast majority of the nodes with a few connections only. Increasing the tail exponent reduces centralization, which is very well captured by the corresponding monotonic decrease of the Gini coefficient shown in Figure 1. When the tail exponent is large enough to ensure the convergence of the first few moments (i.e. $\gamma \gtrsim 3.5$), the network degree distribution is *de facto* egalitarian, despite still being power law, as the Gini coefficient is quite close to zero and well within ranges that are naturally observed in systems characterized by mild heterogeneity. For instance an Erdős-Rényi network with $\langle k \rangle = 8$, as the one used in the example in Figure 1, has a Gini coefficient around 0.08. Similar considerations can be made from the behavior of the Theil index, which also decays to values very close to zero for $\gamma \gtrsim 3.5$.

The presence of heavily connected hubs in scale-free networks improves the communication between nodes by providing additional paths with respect to more homogeneous topologies, e.g., ER networks. In fact, for $\gamma \geq 3$, the average distance between nodes in scale-free networks scales logarithmically in N , as in (3), but it is

systematically smaller than that of ER networks when comparing systems with the same average degree (see [3]). When $2 < \gamma < 3$ the distance between nodes scales as $\log \log N$ and when $\gamma \leq 2$ the distance becomes a few steps only independently from the size of the network [5].

Facilitated communication between nodes comes at a price, as scale-free networks are much less resilient than their more egalitarian counterparts. Attacks aimed at destroying hubs can have severely disruptive effects on the transport properties. On the other hand, ER networks are more resilient to targeted attacks, as no node (or group of nodes) is particularly central in the topology.

2.4 Collective properties of networks

The topology of networks strongly affects their collective properties. For instance, the threshold for the emergence of a giant component is very different in ER vs scale-free networks. A network of N nodes has a giant component if the size S of its largest connected component (a connected component is a set of nodes such that any pair amongst them is connected by at least one path) is formed by an extensive number of nodes, i.e. $\lim_{N \rightarrow \infty} S/N > 0$. For uncorrelated networks (i.e. networks with no correlations between degrees of neighboring nodes), it is possible to show [4] that a giant component is present if

$$\frac{\langle k^2 \rangle}{\langle k \rangle} > 2, \quad (8)$$

where $\langle k \rangle$ and $\langle k^2 \rangle$ are the first and second moments of the degree distribution. From the above criterion, it follows that ER networks always have a giant component if $\langle k \rangle > 1$, whereas scale-free networks display a giant component if $\gamma \leq 3$.

Similarly, it is possible to show that disease spreading processes on scale-free networks always lead to an epidemic outbreak whenever $\gamma \leq 3$, while on ER networks this does not happen if the rate of contagion is low enough [11].

In the following section, we consider a simple model of information spreading, and we discuss how this process is affected by the structure of the network.

3 A network model of blockchain forks: efficiency and egalitarianism

In the context of bitcoin [12], each node is a server that keeps a ledger containing a record of all past transactions. Different ledgers are synchronized at regular intervals

by broadcasting a block created by one node. A block contains all transaction that have been recorded by a node since the last block was broadcasted and verified by all nodes in the system. When a new block, A , is found by a node, the node will broadcast it to its neighbors, these in turn will broadcast it to their neighbors, and so on. A conflict may arise if an alternative block B is independently discovered by a node before the block A has reached all nodes in the system. Such a conflict is called *fork*. Blockchain forks should be avoided because they effectively amount to inconsistencies in the system. In the following, we consider the model introduced in [13] to show how the structural properties of networks affect the propagation of information through the network.

We consider a model in which we have a network of N nodes. Each node is endowed with an integer variable $\sigma_i \in \{-1, 0, 1\}$. When $\sigma_i = 0$ no new block has been discovered or verified by node i since the previous synchronization, while states ± 1 represent two conflicting blocks. We consider a very simple dynamics in which at time $t = 0$ all nodes are in state 0 except a randomly chosen fraction, which are in state 1. The dynamics then runs as follows: at rate β , a node i in state +1 will broadcast its block to its neighbors in state 0, while potentially conflicting blocks can be discovered by any node in state 0 with rate η . Whenever a conflicting block is found, the node responsible for its discovery switches to state -1 and starts propagating to its neighbors in state 0 at rate β .

In the following, we will present an analytical description of the model for Erdős-Rényi random networks, for which we will compute the probability of observing a fork as a function of the average degree of the network, its size (i.e. the number of nodes), as well as the fraction of nodes to which the block has been broadcasted. We will then present a comparison between the performance of Erdős-Rényi and scale-free networks.

3.1 Erdős-Rényi networks

We are interested in computing the probability that the initial block is broadcasted to a fraction n^* of nodes in the system before an alternative block is discovered. Until no new blocks are discovered, the dynamics of the system is equivalent to that of a Susceptible-Infected (SI) process [6]. The SI model on an Erdős-Rényi random network can be described in terms of the following rate equation:

$$\frac{dn_A(t)}{dt} = \beta n_A(t)(1 - n_A(t)) , \quad (9)$$

where $n_A(t)$ is the fraction of nodes to which block A has been broadcasted to (informed nodes), and β is the rate at which the signal coming from an informed node reaches one of its uninformed neighbors. For Erdős-Rényi networks we can write $\beta = \langle k \rangle c p_0 / \langle l \rangle$, where c is the speed at which the signal travels, $\langle l \rangle$ is the average distance between two nodes, and p_0 the rate at which the informed node sends the signal. Equation (9) can be intuitively understood as follows: In a time interval of length dt each node in state +1 will broadcast the block B to its neighbors with probability βdt , hence the term $\beta n_A(t)$ in the equation. A node that is broadcasting will propagate the block to its neighbors that have not been yet informed, whose number can be approximated as $\langle k \rangle (1 - n_A(t))$, i.e. the average number of neighbors of a node times the probability that a node is in state 0.

The solution of equation (9) is given by (see for instance [6])

$$n_A(t) = \frac{n_A(0)e^{\beta t}}{1 + n_A(0)(e^{\beta t} - 1)}. \quad (10)$$

If $n_A(0)$ is the fraction of nodes that know the block A at time zero, we can compute the time $T(n^*)$ needed for block A to be broadcasted to a fraction $n^* \in [n_A(0), 1 - 1/N]$ of the system. This time can be computed by solving the equation $n_A(T(n^*)) = n^*$, which gives

$$T(n^*) = \beta^{-1} \log \left(\frac{n^*(1 - n_A(0))}{n_A(0)(1 - n^*)} \right). \quad (11)$$

From this equation we can compute the time needed for a block to be broadcasted to 50% of the nodes without alternative blocks being discovered. If we assume the process to start from a single informed node (i.e. $n_A(0) = 1/N$) we have

$$T(50\%) \simeq \frac{\log N}{\beta}. \quad (12)$$

From this we see that the time needed for information to be propagated in the network increases logarithmically with the size of the network and it is inversely proportional to $\beta = \langle k \rangle c p_0 / \langle l \rangle$.¹ Therefore, propagation speed can be increased by increasing the average connectivity.

¹More in general, it is possible to show that the initial phases of the propagation process are characterized by an exponential behavior for the fraction of informed nodes over time. This increases with a characteristic time $\tau = 1/\beta$, with $\beta = \frac{c p_0}{\langle l \rangle} \frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle}$. Note that, for a scale-free network with a tail exponent $2 < \gamma \leq 3$, if we consider the natural cut-off $N^{1/(\gamma-1)}$ for the degree distribution [4], we find that β scales as $\beta \sim \langle k^2 \rangle / \langle k \rangle \sim N^{1/(\gamma-1)}$ for large values of N . This would alter the dependence of $T(0.5)$ on N in equation (12), speeding up the propagation.

Equation (11) can be calibrated to obtain an estimate for the minimum time $T_{\min}(50\%)$ needed for a block to reach 50% in a realistic situation. A rough estimate of $T_{\min}(50\%)$ can be given as follows: Let us consider N nodes that are randomly placed on a spherical surface of radius R_0 , the radius of the Earth. The average distance between these nodes is $\pi R_0/2$. Since information cannot propagate faster than light, we have that the minimum average time for a block to be propagated between two nodes is $\frac{\pi R_0}{2c}$ and therefore the minimum time to reach 50% of the nodes in the network is

$$T_{\min}(50\%) = \frac{\pi R_0 \log N}{2c \langle k \rangle} \approx 0.033 \frac{\log N}{\langle k \rangle} \text{ seconds}; \quad (13)$$

considering that in the blockchain there are $N \approx 6000$ nodes, and $\langle k \rangle \approx 10$ we have $T_{\min}(50\%) \simeq 30$ milliseconds. From this result we note that in a network with $\langle k \rangle > \log N$ the average time to reach 50% of the nodes can be faster than the time needed in average to reach all the neighbors of a given node. This is not a contradiction, and it is due to the fact that the first neighbors to be informed start broadcasting the information to their neighbors, and, among these, those which receive the information start broadcasting it to their neighbors, and so on.

From equation (11) we can also compute the probability that, given a discovering rate η , no new block is discovered before block A has been broadcasted to a fraction n^* of the system. In the continuous time limit we have that

$$q(n^*) = 1 - (1 - \eta)^N \int_0^{T(n^*)} (1 - n_A(t)) dt . \quad (14)$$

Plugging equation (11) into equation (14), we can write $q(n^*)$ as

$$q(n^*) = 1 - (1 - \eta)^{N f(n^*, \beta)} , \quad (15)$$

where

$$f(n^*, \beta, n_A(0)) = \beta^{-1} \log \left(\frac{n^*}{n_A(0)} \right) , \quad (16)$$

from which we easily find that

$$\begin{aligned} q(n^*) &= 1 - (1 - \eta)^{\beta N \log(n^*/n_A(0))} \\ &= 1 - \left(\frac{n^*}{n_A(0)} \right)^{\beta N \log(1 - \eta)} . \end{aligned} \quad (17)$$

Equations (15) and (16) can be used to characterize the performance of the network as a function of all its structural parameters, such as its average degree and size. In

the following, we will measure the efficiency of a network in terms of the probability $q(50\%)$ that block A reaches 50% of the network before block B is discovered. Other metrics could be used, for instance the fork probability can be computed from the above equations as the probability that the block A is broadcasted to $N - 1$ nodes before an alternative block is found².

In figures 2 and 3 we show the behavior of $q(50\%)$ as a function of the average degree of the network and its size. In both cases there is very good agreement between numerical simulations and analytical results from equation (17). Figure 2 highlights the role played by the network's *density*, as measured by the average degree. In fact, an increase in density (at fixed N) corresponds to an increase in the average number of connections each node can exploit to broadcast a newly discovered block, therefore speeding up the overall propagation process and ensuring a fast convergence to consensus. On the other hand, very sparse networks are rather inefficient, as low connectivity practically corresponds to low average numbers of paths connecting distant nodes, i.e. to the *de facto* separation between different regions of the network.

Figure 3 shows instead that network expansion at fixed density has adverse effects on the efficiency of block propagation, as the presence of larger numbers of nodes increases the likelihood of discovering conflicting blocks.

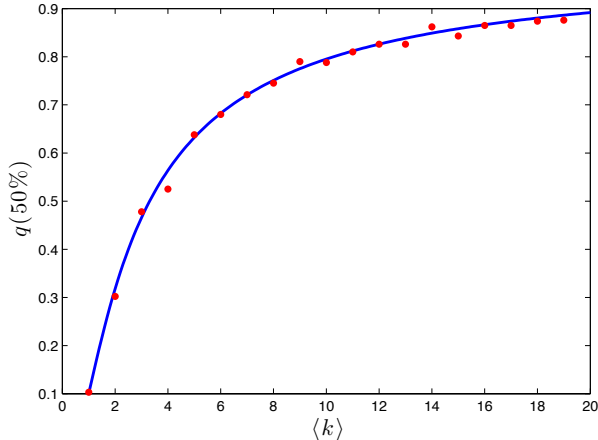


Figure 2: **Effect of connectivity on block propagation:** Probability of reaching 50% of the network before a conflicting block is discovered as a function of the network's average degree $\langle k \rangle$ for a system with $N = 1000$, $\eta = 10^{-3}$, and $n_A(0) = 0.1$. Dots: results from numerical simulations. Solid line: analytical result from Eq.17. There is an overall good match between analytical and numerical results, which show that an increase in connectivity is beneficial to the network's efficiency.

²Note here that the analytical equations are not defined for $n^* = 1$.

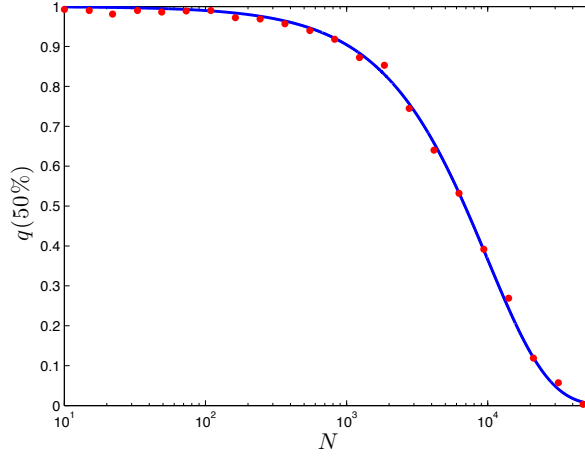


Figure 3: **Scalability with respect to network size:** Probability of reaching 50% of the network before a conflict as a function of network size for a system with $\langle k \rangle = 7$, $\eta = 10^{-3}$, and $n_A(0) = 0.1$. Dots: results from numerical simulations. Solid line: analytical results from Eq.17. As can be seen, there is a very good agreement between analytical and numerical results, which show that network expansion at fixed densities, i.e. at fixed average degree $\langle k \rangle$, increasingly prevents the likelihood of reaching consensus in the system.

Lastly, Figure 4 shows the probability of informing 50% of the network (at fixed density and number of nodes) conditional on having already informed a fraction $n_A(0)$ of nodes at an initial time $t = 0$. As one would expect intuitively, efficiency increases monotonically with $n_A(0)$. However, let us note that there are remarkable marginal gains in efficiency for relatively small increases of the initially informed fraction of nodes. This is shown in the inset in Figure 4, where we plot the discrete derivative of $q(50\%)$ that corresponds to incrementing the number of initially informed nodes by one node at a time.

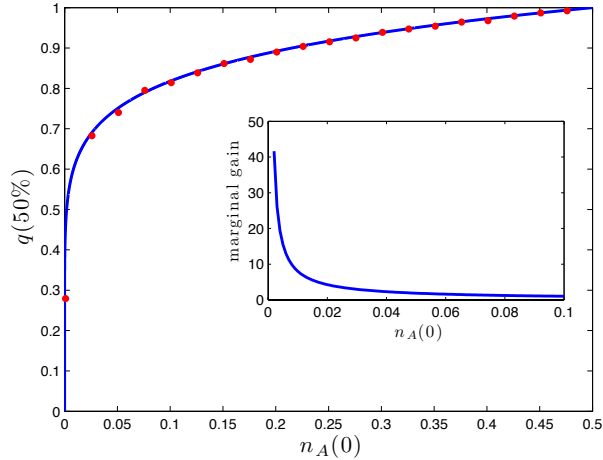


Figure 4: **Dependence on initial condition:** Probability of reaching 50% of the network conditional on a given fraction $n_A(0)$ of nodes being already informed for a system with $\langle k \rangle = 7$, $\eta = 10^{-3}$, and $N = 1000$. Dots: results from numerical simulations. Solid line: analytical results from Eq.17. Inset: zoom in the region of small $n_A(0)$. Controlling 5% of the system allows to reach the majority of the system without observing conflicts in about 70% of the cases.

3.2 Comparing the performance of different networks: Erdős-Rényi vs scale-free networks

In the previous section we have seen how properties such as the size of an Erdős-Rényi network and its average degree affect the probability that a block reaches the majority of the system before a conflict occurs. We now turn to the comparison of the performance of Erdős-Rényi with respect to that of scale-free networks. To this end we resort to numerical simulations, because there is no analytical solution for the case of scale-free networks. For scale-free networks, we consider networks generated through a linear preferential attachment rule [9]. This mechanism generates networks whose degree distribution has a tail exponent equal to $\gamma = 3$. This is a rather common exponent value for real networks [5]. In figure 5 we plot the ratio between the probability of reaching the majority of the network before a conflict in scale-free vs. Erdős-Rényi networks with the same average degree $\langle k \rangle = 8$, and for different values of N . From the plot, we see that networks generated through preferential attachment become more efficient as N increases. The increased efficiency of scale-free networks is due to the presence of hubs, that connect most of the nodes in the network through short paths. However, the presence of hubs also increases the fragility of the network with respect to targeted attacks. In fact, it has been shown [5]

that an attack protocol that seeks to remove nodes starting from those with the highest degree quickly breaks a scale-free networks into disconnected components. For instance, the fraction of nodes that need to be removed for the giant component to disappear is, in the case of scale-free networks, typically of a few percent [14, 15]. Although the scale-free networks we have used in our simulations have a tail exponent $\gamma = 3$, we expect the higher efficiency of scale-free networks to hold also for smaller values of γ , as lowering the tail exponent increases the probability of having hubs in the system.

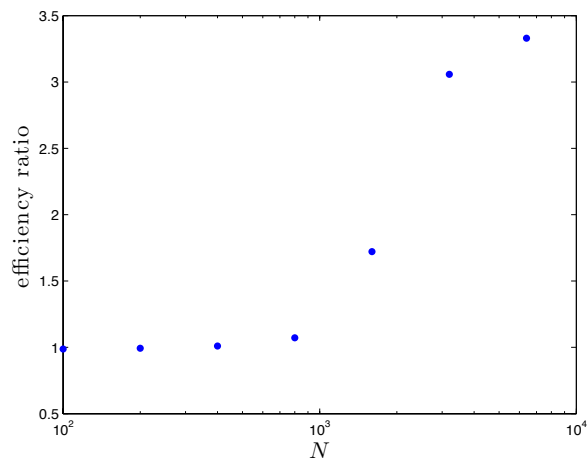


Figure 5: **Scalability of Erdős-Rényi vs scale-free networks:** *Ratio between the probability of reaching 50% of the networks before a new block is discovered on scale-free over Erdős-Rényi networks with the same average degree $\langle k \rangle = 8$ as a function of the network size. As the size increases scale-free networks become more efficient than Erdős-Rényi networks.*

4 Conclusion

Topology strongly affects the behavior of dynamical processes taking place on networks. In this chapter we have shown how egalitarianism is in conflict with efficiency. This was shown by reviewing the main features of two common ensembles of random networks, Erdős-Rényi and scale-free. Erdős-Rényi networks have a Poisson degree distribution, while scale-free networks have a power-law degree distribution. The latter implies the existence of hubs in the network, i.e. nodes with significantly more connections than the average degree, that are not present in Erdős-Rényi networks. Therefore, as suggested by the Gini coefficient associated with the degree distribution of these network ensembles, the two can be considered as prototypical examples

of egalitarian and non-egalitarian networks.

We have discussed that the difference in the degree distributions of these two classes of networks has profound consequences on their behavior, for instance regarding the emergence of a giant component of connected nodes, or the propagation speed of signals in the network. The focus of this chapter was to understand how the propagation of information on networks is affected by their topology. In particular, we have considered a stylized model of block-propagation in the blockchain. In the model, we assume that a new block has been discovered and has to be propagated to the whole network before an alternative block is found, leading to a conflict. Two competing processes take therefore place in the network. On one hand, nodes that have been informed about the new block broadcast it to their neighbors, on the other hand nodes that have not yet been informed can find an alternative block.

We have provided an analytical formula for the probability that a conflict arises in an Erdős-Rényi network, and we have characterized its dependence on the size of the network and its average degree. By means of numerical simulations, we have compared the performance of Erdős-Rényi and scale-free networks, and we have shown that the latter perform better as the size of the network increases. This finding suggests the existence of a trade-off between efficiency and nodes' equality.

References

- [1] http://www.comscore.com/Insights/Market_Rankings/comScore-Releases-June-2015-US-Desktop-Search-Engine-Rankings, 09 2015.
- [2] David Garcia, Claudio J Tessone, Pavlin Mavrodiev, and Nicolas Perony. The digital traces of bubbles: feedback cycles between socio-economic signals in the bitcoin economy. *Journal of the Royal Society Interface*, 11(99):20140623, 2014.
- [3] Réka Albert and Albert-László Barabási. Statistical mechanics of complex networks. *Reviews of modern physics*, 74(1):47, 2002.
- [4] Sergey N Dorogovtsev, Alexander V Goltsev, and José FF Mendes. Critical phenomena in complex networks. *Reviews of Modern Physics*, 80(4):1275, 2008.
- [5] Mark Newman. *Networks: an introduction*. Oxford University Press, 2010.
- [6] Alain Barrat, Marc Barthélemy, and Alessandro Vespignani. *Dynamical processes on complex networks*. Cambridge University Press, 2008.

-
- [7] Ginestra Bianconi. Mean field solution of the Ising model on a Barabási–Albert network. *Physics Letters A*, 303(2):166–168, 2002.
- [8] Daniele De Martino, Luca Dall’Asta, Ginestra Bianconi, and Matteo Marsili. Congestion phenomena on complex networks. *Phys. Rev. E*, 79:015101, Jan 2009.
- [9] Albert-László Barabási and Réka Albert. Emergence of scaling in random networks. *science*, 286(5439):509–512, 1999.
- [10] Paul L Krapivsky, Sidney Redner, and Francois Leyvraz. Connectivity of growing random networks. *Physical review letters*, 85(21):4629, 2000.
- [11] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic spreading in scale-free networks. *Physical review letters*, 86(14):3200, 2001.
- [12] Satoshi Nakamoto. Bitcoin: A peer-to-peer electronic cash system. *Consulted*, 1(2012):28, 2008.
- [13] Christian Decker and Roger Wattenhofer. Information propagation in the bitcoin network. In *Peer-to-Peer Computing (P2P), 2013 IEEE Thirteenth International Conference on*, pages 1–10. IEEE, 2013.
- [14] Reuven Cohen, Keren Erez, Daniel Ben-Avraham, and Shlomo Havlin. Breakdown of the internet under intentional attack. *Physical review letters*, 86(16):3682, 2001.
- [15] Sergey N Dorogovtsev and José FF Mendes. *Physical review letters*, 87(21):219801, 2001.