*International Journal of Mathematics and Statistics Invention (IJMSI)*
*E-ISSN: 2321 – 4767 P-ISSN: 2321 - 4759*
*www.ijmsi.org Volume 5 Issue 7 || August. 2017 || PP 11-16*

# On Small-World and Scale-Free Properties of Complex Network

A. Bharali[1], A. Doley[2]
[1]*Department of Mathematics, Dibrugarh University,Dibrugarh-786004, India*
[2]*Department of Mathematics, DHSK College, Dibrugarh-786001, India*
*a.bharali@dibru.ac.in[1], amitav1987doley@gmail.com[2]*
*Corresponding Author: A. Doley*

**ABSTRACT:** *In this paper we study two popular properties exhibited by many complex networks, viz., small-world and scale-free. The object of this paper is twofold. First, we present a brief historical account on small-world and scale-free properties of complex network. Second, we adopt complex network approach to study few network datasets in connection with small-world and scale-free properties.*

***KEYWORDS***: Complex network, Network Model, Network topology, Small-world, Scale-free.

## I. INTRODUCTION

In nature many system can be described by models of complex networks, which are structures consisting of nodes or vertices connected by links or edges. For example, the Internet, which is a complex network of routers or domains, the World Wide Web (WWW), which is a complex network of websites, the brain, which is a complex network of neurons, the power grids, airport network etc. Food webs and metabolic pathways can also be represented by networks, as can the relationships among words in a language.

In 1950, two mathematicians, *Erdös* and *Rényi* (ER) made described a network with complex topology by a random graph [12]. Their work had laid a foundation of the random network theory. Although intuition clearly indicates that many real-life complex networks are neither completely regular nor completely random, the ER random graph model was the only sensible and rigorous approach that dominated researchers' thinking about complex networks for nearly half of a century, due essentially to the absence of super-computational power and detailed topological information about very large-scale real-world networks. In the past few years, the computerization of data acquisition and the availability of high computing power have led to the emergence of huge databases on various real networks of complex topology.

In 1998, in order to describe the transition from a regular lattice to a random graph, *Watts* and *Strogatz* (WS) introduced the concept of small-world network. In their 1998 seminal paper, *Watts* and *Strogatz* described networks, which are "highly clustered, like regular lattices, yet have small characteristic path lengths, like random graphs" [3]. A small-world network is a network which has clustering similar to a regular lattice and path length similar to a random network. It is notable that the small-world phenomenon is indeed very common. An interesting popular manifestation of the "small-world effect" is the so-called "six degrees of separation" principle, suggested by a social psychologist, *Milgram*, in 1960s [17]. Although this point remains controversial, the small-world pattern has been shown to be ubiquitous in many real networks. A prominent common feature of the ER random graph and the WS small-world model is that the degree distribution of a network peaks at an average value and decays exponentially. Such networks are called "exponential networks" or "homogeneous networks," because each node has about the same number of link connections.

Another path-breaking contribution in the field of complex networks is the scale-free networks by *Barabási* and *Albert* [13,14]. The degree distributions of scale-free networks follow a power-law regime that is independent of the network scale. The field has received growing attention from scientists and researchers working in many walks of life. Scale-free networks seem to match real-world applications much better than ER network models. Differing from an exponential network, a scale-free network is inhomogeneous in nature; most nodes have very few link connections and yet a few nodes have many connections. The discovery of the small-world effect and scale-free feature of complex networks has led to dramatic advances in complex networks theory in the past few years. In 2001 *Strogatz* has imparted about complex networks through his paper "Exploring complex networks" [16], *Barabási* and *Albert* in 2002 through the paper "Statistical mechanics of complex networks" [15], and *Wang* in 2002 through his paper "Complex networks: topology, dynamics and synchronization" [21].

The rest of the paper is organized as follows: in the next section we present some preliminaries of network science and these notions will be expedited in the rest of the paper. In section III we present a historical account as well as discuss the research contributions related to small-world and scale-free properties. In section IV we discuss different models proposed in complex network analysis. In section V we consider few benchmark network datasets to study in connection with the two properties namely small-world and scale-free properties. In section VI conclusions are made.

## II. SOME PRELIMINARIES

A network is a system whose objects are somehow connected. The objects of the system are represented as nodes or vertices and the connections among interacting objects are known as ties, edges, arcs, or links. Based on the properties of the edges, networks can be further subdivided into Undirected, Directed, and Weighted networks. Before going further we need to know few basic network measures popularly used in network analysis [20].

**Degree Distribution**: For a graph $G = (V, E)$, the degree distribution of $G$ is $p_k =$ fraction of vertices with degree $k$, for $k = 1,2, \dots, n$. Thus if there are $n$ nodes in total in a network and $n_k$ of them have degree $k$ then $p_k = \frac{n_k}{n}$.

**Clustering Coefficient**: Consider a simple connected, undirected graph $G$ and vertex $v \epsilon V(G)$ with neighbor set $N(v)$. Let $n_v = |N(v)|$ and $m_v$ be the number of edges in the subgraph induced by $N(v)$, i.e., $m_v = |E(G[N(v)])|$. The Clustering coefficient, $cc(v)$ for node v with degree $\delta(v)$ is defined as

$$cc(v) \stackrel{\text{def}}{=} \begin{cases} \frac{m_v}{\binom{n_v}{2}}, \delta(v) > 1 \\ \text{undefined, otherwise} \end{cases}$$

The Clustering coefficient $CC(G)$ for the entire network is defined as the average over all (well defined) clustering coefficients of its nodes. In case of a random network of size $n$ the clustering coefficient will be $CC_{rand} = \frac{\langle k \rangle}{n}$ where $\langle k \rangle$ is the mean degree of the network.

**Average Path Length**: The average path length L of a network measures the average number of links along the shortest paths for all possible pair of nodes in the network. If $d(i, j)$ is the minimum distance between nodes $i$ and $j$ then L can be defined as:

$$L = \frac{1}{n(n-1)} \sum_{i \neq j} d(i, j), \text{ where n is the number of nodes in the network.}$$

In case of a random network of size n the average path length will be $L_{rand} = \frac{\log n}{\log \langle k \rangle}$, where $\langle k \rangle$ is the mean degree of the network.

## III. SMALL-WORLD EFFECT AND SCALE-FREE PROPERTY

3.1 Small world effect: During 1960s *Stanley Milgram* carried out the famous experiment [17], in which letters passed from person to person were able to reach a designated target individual in only a small number of steps. This result is one of the first direct demonstrations of the small-world effect, the fact that most pairs of nodes inmost networks seem to be connected by a short path through the network.

On one hand, the small-world effect has obvious implications for the dynamics of processes taking place on networks. For example, if one considers the spread of information, or indeed anything else, across a network, the small-world effect implies that that spread will be fast on most real-world networks. And on the other hand, the small-world effect is also mathematically obvious. If the number of vertices within a distance r of a typical central node grows exponentially with r then the value of average path length will increase as log n, and this is true of many networks, including the random graph. In recent years, the term “small-world effect” has thus found a more precise meaning, which can be presented as follows: A network is said to exhibit the small-world effect if the value of average path length scales logarithmically and the clustering coefficient is very high as compared to a random network of same size. A brief account of theoretical work on “Small-world” phenomenon may be found in [3, 4, 10].

**3.2 Scale free property:** One way of presenting degree data is to make a plot of the cumulative distribution function. It is the best way to present the degree data as it reduces the noise in the tail of the curve. In case of many networks this distribution follows power-laws in the tails. The term "scale-free" refers to any functional form f(x) that remains unchanged to within a multiplicative factor under a rescaling of the independent variable x. In effect this means power-law forms, since these are the only solutions to f(ax) = bf(x), and hence "power-law" and "scale-free" can be used synonymously. The most notable characteristic in a scale-free network is the relative commonness of nodes with a degree that greatly exceeds the average. The highest-degree nodes are often called "hubs", and are thought to serve specific purposes in their networks.

Networks with power-law degree distributions draw a great deal of attention in the literature of complex network research [1, 4, 9, 15, 18, 19]. They are sometimes referred to as scale-free networks, although it is only their degree distributions that are scale-free; one can and usually does have scales present in other network properties.

## IV. FEW NETWORK MODELS

In this section we discuss few classical network models [11], using which one can obtain networks with small-world effect and scale-free property.

**4.1 Small World Model:** Many real-world networks have a geographical component to them; the nodes of the network have positions in space, and in many cases it is reasonable to assume that geographical proximity will play an important role in deciding which nodes are connected to which others. The small-world model starts from this idea by positing a network built on a low-dimensional regular lattice and then "rewiring" edges to create a low density of "shortcuts" that join remote parts of the lattice to one another. The rewiring procedure involves going through each edge in turn and, with probability p, moving one end of that edge to a new location chosen uniformly at random from the lattice, except that no double edges or self-edges are ever created.

4.1.1 **Watts-Strogatz Model**: A *Watts-Strogatz* network is a random graph that can be obtained by rewiring links in a circle in which only neighbors are connected initially. *Watts-Strogatz* networks possess small-world properties as the rewiring probability p is big enough. The *Watts-Strogatz* model was first introduced by *Duncan J. Watts* and *Steven Strogatz* in their joint work published in Nature in 1998 [2]. Fig. I shows that an increasing the value of p results in increasing randomness [6]. A *Watts-Strogatz* network can be generated using the following steps:

- Construct a regular ring lattice with N nodes each connected to 2m neighbors (m on each side).
- For every node $n_i$, $i = 1,2,\dots,N$ take every edge $(n_i, n_j)$ with $i < j$, and rewire it with probability $p$. Rewiring is done by replacing $(n_i, n_j)$ with $(n_i, n_k)$ where k is chosen with uniform probability from all possible values that avoid self-loop and link duplication.

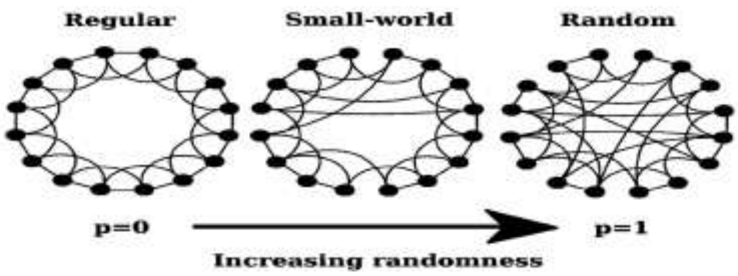


Fig. I: The value of p can tune the randomness

***4.2 Models of Network Growth***

4.2.1 **Price's Model**: This model is a mathematical model for the growth of citation networks. The model picked up the ideas of the *Simon* model [5] reflecting the concept of rich get richer. *Derek Price* took the example of a network of citations between scientific papers and expressed its properties. His idea was that the way how an old node (existing paper) gets new edges (new citations) should be proportional to the number of existing edges (existing citations) the node already has. This was referred to as cumulative advantage, now also known as preferential attachment. *Price*'s work is also significant in providing the first known example of a scale-free network.

The *Price*'s Model may be stated as follows [11]: Consider a directed graph with $n$ nodes. Let $p_k$ denote the fraction of nodes with degree $k$ so that, $\sum_k p_k = 1$. Each new node has a given out-degree (namely those papers it cites) and it is fixed in the long run. This does not mean that the out-degrees cannot vary across nodes, simply we assume that the mean out-degree $m$ is fixed over time. It is clear that, $\sum_k kp_k = m$, consequently $m$ is not restricted to integers. The most trivial form of preferential attachment means that a new node connects to an existing node proportionally to its in-degrees. The main problem of such idea is that no new node is connected when it is joined to the network so it is going to have zero probability of being connected in the future. To overcome this, *Price* proposed that an attachment should be proportional to some $k + k_0$ with $k_0$ constant. In general $k_0$ can be arbitrary, yet *Price* proposed $k_0 = 1$, in that way an initial connection is associated with the new node (so the proportionality factor is now k + 1 instead of k). The probability of a new edge connecting to any node with a degree k is $\frac{(k+1)p_k}{\sum_k (k+1)p_k} = \frac{(k+1)p_k}{m+1}$.

### 4.2.2 Barabási-Albert Model

In the *Barabási-Albert* (BA) model a network can be created by using the following steps.

- Start from a small number $m_0$ of nodes.
- At each step add a new node $u$ to the network and connect it to $m (\leq m_0)$ of the existing nodes $v$ with probability $p_u$, where

$$p_u = \frac{k_u}{\sum_w k_w}$$

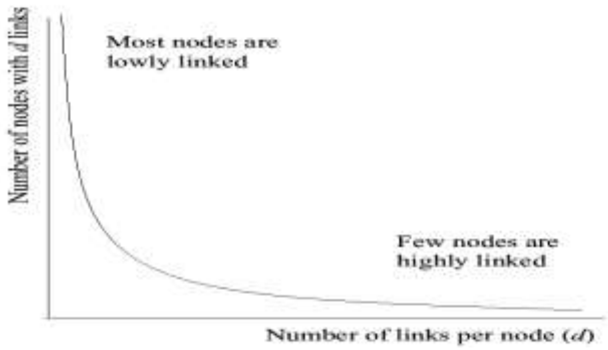


Fig. II: Power law degree distribution of a network generated using BA model.

We can assume that we start from a connected random network of the ER type with $m_0$ nodes. In this case the BA process can be understood as a process in which small inhomogeneities in the degree distribution of the ER network growths in time.

## V. CASE STUDY

Complex networks can be classified into four broad categories [11]. For case study we consider four network data sets one from each of the four classes of network. The network data sets are obtained from network databases KONECT [7] and SNAP [8].

### 5.1 Networks Description

5.1.1 **Protein Network**: This undirected network is a representation of protein interactions contained in yeast. A node in the network stands for a protein and an edge represents a metabolic interaction between two proteins.

5.1.2 **Email-Eu-Core Network**: The Email-Eu network was generated using email data from a large European research institution. A node represents a network whereas an edge means an exchange of at least one email. Email-Eu-Core is the "Core" of Email-Eu network, which contains links between members of the institution.

5.1.3 **Facebook (NIPS) Network**: This directed network contains Facebook user-user friendships. A node represents a user. An edge indicates that the user represented by the left node is a friend of the user represented by the right node.

5.1.4 **US Air Traffic Control Network (USATC)**: This network was constructed from USA's FAA national flight data centre, preferred routes database. Nodes in this network represent airports or service centres and links are created from strings of preferred routes recommended by National Flight Data Centre.

**5.2 Small-world effect of the networks**

The average path lengths and clustering coefficients of the networks are calculated using MatlabBGL toolbox. The obtained results are presented in the Table I. It is seen that the networks that fall in the four broad categories of complex networks have small average path length and high clustering coefficient in comparison with the corresponding random networks of same size. In case of Facebook (NIPS) network, the small-world effect is very prominent whereas it is marginal in case of Email-Eu-Core network as per the average path length is concerned. The clustering coefficients of all the networks are found to very high as compared to corresponding random networks. These observations suggest that the networks under study exhibit small-world property. Since small-world networks usually have almost equal degree for all the nodes, it is robust to targeted attack.

| **Networks** | **Node** | **Edge** | **L** | **CC** | **Corresponding Random Network** | |
|---|---|---|---|---|---|---|
| | | | | | **L** | **CC** |
| **Facebook** | 2888 | 2981 | 0.0066 | 0.0136 | 251.4265 | 0.0004 |
| **USATC** | 1226 | 2615 | 5.216 | 0.0404 | 9.3878 | 0.0017 |
| **Protein** | 1870 | 2277 | 0.144 | 0.0336 | 38.2013 | 0.0007 |
| **Email-Eu-Core** | 1005 | 25571 | 2.0834 | 0.3657 | 2.1359 | 0.0253 |

Table I: Comparison of average path length (L) and clustering coefficient (CC) of the networks with random networks of same size.

**5.3 Scale-free property of the networks**

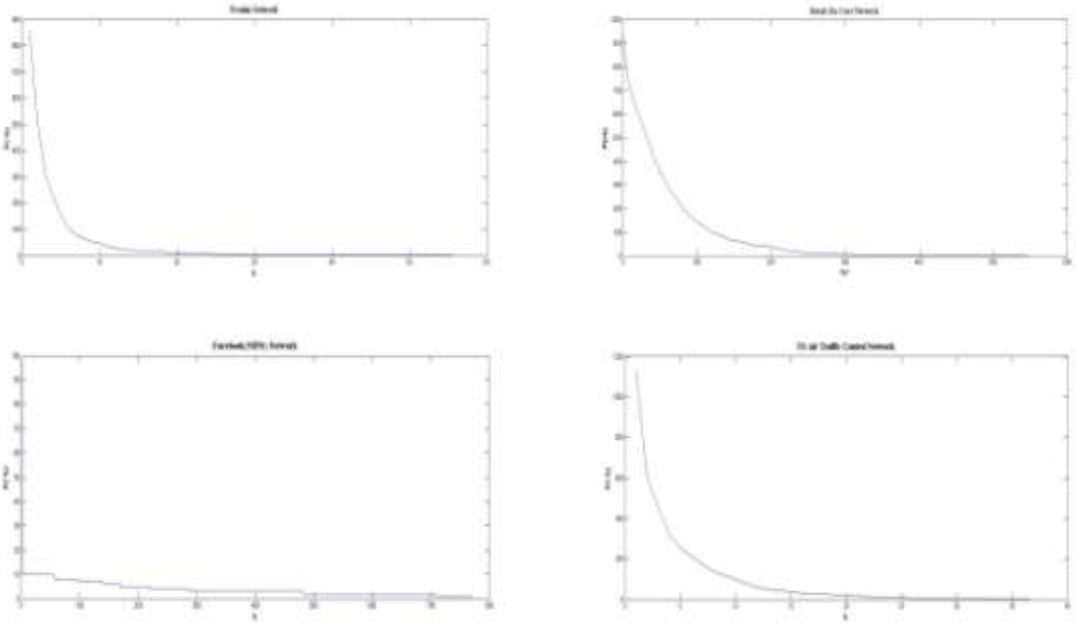

Fig. III: Illustration of the characteristic power-law degree distribution of the networks.

In Fig. III, we can observe that the degree distribution of the networks follow power-laws in the tails. This observation suggests that the networks under study exhibit scale-free property. In case of Facebook (NIPS) network, the distribution curve is noisy unlike the other three networks under consideration. Thus we can say it obeys a segmented power-law, which indicates the presence of many communities within the network. In this kind of networks where there is large variation among communities, global values of statistical measures can be misleading. The presence of modular structure may also alter the way in which dynamical processes (e.g., spreading processes and synchronization) unfold on the network. The scale-free property strongly correlates with the network's robustness to random failure. But these networks are vulnerable to targeted attacks, because if the hubs are attacked the network will literally collapse.

## VI. CONCLUSION

In this paper we have discussed the chronological development of Small-world effect and Scale-free property which are very common in complex network analysis. We also study four popular network datasets in connection with these properties. These properties can be very helpful in understanding the network topology and its evolution. Scale-free networks show properties of real-world networks and are often a better fit for modeling real-world processes. We have found that like many other real-world networks, both small-world and scale-free properties are prominent in the four networks that we have considered in this study.

## REFERENCES

[1] A. L. Barabási, Scale-Free Networks: A Decade and Beyond, Science, 325, 2009, 412-413.
[2] D. J. Watts, S. H. Strogatz, Collective dynamics of 'small world' networks, Nature, 393, 1998, 440-442.
[3] E. Estrada, Introduction to complex networks: structure and dynamics, in J. Banasiak (eds), Evolutionary Equations with Applications to Natural Sciences (NY Springer, 2015) 93-131.
[4] F. W. Xia, G. R. Chen, Complex networks: Small-world, Scale-free and Beyond", IEEE circuits and systems magazine, 3(1), 2003, 6-20.
[5] H. A. Simon, On a class of skew distribution functions, Biometrika, 42, 1955, 425-440.
[6] H. Lavicka, Simulations of agents on social networks, doctoral Thesis, Czech Technology University, Prague, 2010
[7] J. Kunegis, Konect network dataset -- KONECT, April 2017, http://konect.uni-koblenz.de
[8] J. Leskovec, J., A. Krevl, SNAP datasets: Stanford Large dataset collection, 2014, http://snap.stanford.edu/data
[9] K. Konstantin, M. E. Victor, Growing scale-free networks with small-world behaviour, Physical Review E. 65 (5), 2002, 057102.
[10] M. E. J. Newman, Models of the small world: A Review, J. Statist. Phys., 101, 2000, 819–841.
[11] M. E. J. Newman, The structure and function of complex networks, SIAM Rev., 45(2), 2003, 167-256.
[12] P. Erdös, A. Rényi, On the evolution of random graphs, Publ. Math. Inst. Hung. Acad. Sci., 5, 1959, 17-60.
[13] R. Albert, A. L. Barabási, Emergence of scaling in random networks, Science, 286, 1999, 509-512.
[14] R. Albert, A. L. Barabási, H. Jeong, Mean-field theory for scale-free random networks", Physica A, 272, 1999, 173-187.
[15] R. Albert, A. L. Barabási, Statistical mechanics of complex networks, Review of Modern Physics, 74, 2002, 47-91.
[16] S. H. Strogatz, Exploring complex networks, Nature, 410, 2001, 268-276.
[17] S. Milgram, The small-world problem, Psychology Today, 2, 1967, 60-67.
[18] S. N. Dorogovtsev, J. F. Mendes, Evolution of networks, Adv. in Phys., 51, 2002, 1079–1187.
[19] S. N. Dorogovtsev, J. F. Mendes, Scaling properties of scale-free evolving networks: Continuous approach, Phys. Rev. E., 63, 2001, 056125.
[20] S. Wasserman, K. Faust, Social Network Analysis: Methods and Applications (NY: Cambridge University Press, 1994).
[21] X. F. Wang, Complex networks: topology, dynamics and synchronization, Int. J. Bifurcation & Chaos, 12(5), 2002, 885-916.