DOI:10.21276/sjeat

Saudi Journal of Engineering and Technology
Scholars Middle East Publishers
Dubai, United Arab Emirates
Website: http://scholarsmepub.com/

ISSN 2415-6272 (Print)
ISSN 2415-6264 (Online)

Review Article

Complex Networks and their Contribution to Real Network Design-A Survey

Hilary Ackah-Arthur
School of Computer Science and Telecom Engineering, Jiangsu University, Zhenjiang, 212013, China

***Corresponding Author:**
Hilary Ackah-Arthur
Email: hilaryaa@stmail.ujs.edu.cn

Abstract: Researches on complex networks have been performed across several fields of science, including technical networks, information networks, biological networks and social networks. Researchers have discovered different types of characteristics for complex networks such as regular-coupled networks, random networks, small-world networks and scale-free networks. Application of such network properties to solve real network problems, however, is still at the infancy stage. In this study, some key characteristics of complex networks are critically evaluated, the current state of complex networks and the application of theories of complex network to real network design. In particular, inspired by the average-path-length property and clustering coefficient, they are employed to evaluate the theories. It will also identify some of the challenges in applying these theories to real networks.
Keywords: Complex networks, Scale free, Small world, Random graphs, Regular coupled networks, Average path length, Clustering coefficient.

INTRODUCTION

Networks are all around us. The domain of the study of networks has mainly been a part of discrete mathematics referred to as *graph theory*.

The last decade has seen the origin of a new group of interest and studies in *complex networks*, referring to networks that have irregular structure, are complex and evolve dynamically in time, with the core purpose of migrating from small networks analysis to that of large systems, and with more attention to dynamic network properties. Complex networks studies has experienced significant achievements in various contexts, such as in the social sciences. This flurry of activity, initiated by two seminal papers by Watts and Strogatz, on small-world networks, and that of scale-free networks by Barabási and Albert, has contributed to the inclusion of the physics' community in the research area, and has surely been stimulated by the rise in computing powers and by the possibility to understand the properties of a plenty of large databases of real networks.

This paper reviews the present theories of complex networks and their applications to real network designs. It also identifies the some challenges in their application.

The paper is organized as follows. The first section assesses the various types of complex networks, followed by brief description of some basic concepts employed in complex networks. Next, it identifies the various complex network models, and analyzes the application of the different types of complex networks to real network application design. It then evaluates challenges to network design and finally provides conclusions.

TYPES OF COMPLEX NETWORKS

A complex network is composed of a set of items, also known as vertices or nodes that have connections between them, known as edges. Systems that have the form of networks are in all places of the world. Examples include social networks of acquaintance showing links between individuals, the World Wide Web, the Internet, organizational networks and networks of inter-business relations, neural networks, food webs, metabolic networks, networks of citations between papers, distribution networks such as blood vessels or postal delivery routes, and many others.

Networks are categorized into four less strict groups based on their common properties [1]: social networks, technological networks, information networks, and biological networks.

Social Network

A social network is a set of people or groups of people with some pattern of contacts or interactions between them (friendship, kinship, status, sexual, business or political) among them [2, 3]. The patterns of friendships among individuals [4, 5], business relationships between companies [6, 7], and intermarriages between families [35], all shows networks that have been studied in the past.

Information Networks

The second network category is what is called information networks (also sometimes called "knowledge networks"). The network of citations among academic papers is a typical example of an information network. Most learned articles cite previous works by others on related topics. The acyclic nature of citation networks are because papers can only cite other papers that have already been written, not those that have yet to be written. Thus the network cannot have closed loops since all edges point backwards in time.

The World Wide Web is another very important example of an information network, which refers to a network of Web pages containing information, linked together by hyperlinks from a page to another. The World Wide Web is cyclic, as compared to a citation network; there are no constraints and natural ordering of sites, which prevent the formation of closed loops. The Web obtains data from "crawls" of the network, which finds Web pages by links from other pages. For the web, a page is found only if another page in the web points to it; this is different from a citation network, where paper can show in the citation indices although it has never been cited.

Technological Networks

The third class of networks is technological networks, is usually man-made. It is designed typically for the distribution of some commodity or resource, such as information or electricity. The electric power grid (high-voltage network made up of three-phase transmission lines) is an excellent example of technological networks [8-10]. Other types of distribution networks comprise the network of airline routes [10] and networks of roads [11], railways [12], pedestrian traffic [13] and the telephone network. Another very widely studied technological network is the Internet, i.e., the network of physical connections between computers.

Biological Networks

A number of biological systems can be usefully represented as networks. The network of metabolic pathways is a typical example of a biological network which represents metabolic substrates and products, connected with directed edges if a known metabolic reaction exists to act on a given substrate and produce a given product. The network of mechanistic physical interactions between proteins is a separate network which is usually known as a protein interaction network.

Another much-studied example of a biological network is the food web, which has its vertices symbolizing species in an ecosystem and a directed edge from species A to species B indicating that A preys on B.

SOME BASIC CONCEPTS

Many proposed complex network measures exist, three spectacular concepts—the average path length, the clustering coefficient, and the degree distribution, play a major role in the recent study and development of theories in complex networks. This section provides a brief review of these important concepts.

Average Path Length

The number of edges linking two nodes along the shortest path in a network defines the distance between them. Whiles the largest of all distances between any pair of nodes in the network defines the diameter of a network, the network's average path length L, is then the average of the mean distance between two nodes, over all pairs of nodes. At this point, L shows the effective "size" of a network, which is the most characteristic separation of a pair of nodes therein. In a friendship network, for instance, L represents the average number of friends present in the shortest chain linking two individuals in the network.

Clustering Coefficient

There is likelihood that your friend's friend is as well your direct friend. This property describes the clustering of network. More clearly, a clustering coefficient C can be described as the average fraction of a node's neighbor pairs that also are neighbors of each other. Suppose that a network with k_i edges contains a node i, and this node is connected to k_i other nodes. All these nodes connected to node i are neighbors. Clearly, at most there can exist $k_i (k_i - 1)/2$ edges among them, moreover this occurs when node i has every neighbor connected to every other neighbor of it. The clustering coefficient of node i, denoted as C_i, is then defined as the ratio of the number edges E_i actually existing among the k_i nodes and the total possible number $k_i(k_i - 1)/2$, that is, $C_i = 2E_i/(k_i(k_i - 1))$. The entire network has a clustering coefficient C represented by the average of C_i over all i. Clearly, $C \leq 1$; and $C = 1$ given that it is a globally coupled network, which indicates that the network has every node connecting to every other node. In an entirely random network that consists of N nodes, $C \approx 1/N$, is very small in comparison to most real networks.

Degree Distribution

The simplest and possibly the most significant of the characteristics of a single node is its degree. The degree k_i of a node labeled i is usually defined as the total number of its connections to i.

Thus, the larger the degree, the "more important" the node is in a network. Also, the average of k_i over all i is referred to as the average degree of the network, and is represented by $< k >$. A distribution function P(k), which is the probability that a node selected at random has precisely has k edges, is used to characterize the spread of node degrees over a network. A regular lattice has all its nodes having the same number of edges, hence, has a simple degree sequence; and therefore a plot of its degree distribution contains a single sharp spike (delta distribution). For an entirely random network, the degree sequence follows the common Poisson distribution; as well as the shape of the Poisson distribution drops exponentially, further than the peak value $< k >$. Due to this exponential decline, the probability of locating a node having k edges will obviously be negligibly small for $k >> < k >$. In previous years, many empirical results showed that the degree distribution of many large-scale real networks deviates significantly from the Poisson distribution. In particular, for a number of networks, the degree distribution can very well be described by a power law of the form $P(k) \approx k^{-\gamma}$

Network Resilience

Resilience is an essential property of most complex systems, which refers to the ability of a system to adjust itself to retain its original functionality when errors, failures and environmental changes occur [14]. Networks vary in their level of resilience to vertex removal.

Albert, Jeong, and Barab´asi [15], studied the effect of vertex deletion in two example Networks which were power-law in form. They found for both networks that most networks are robust against random vertex removal but considerably less robust to targeted removal of the highest-degree vertices.

MODELS OF COMPLEX NETWORKS

A great number of attractive characteristics of real-world networks that have attracted the attention of several researchers in recent years, however, concern the manner in which the networks are not like random graphs. In a number of revealing ways, Real networks are non-random which shows the possibility of mechanisms that could direct network formation, in addition to probable ways in which network structure could be exploited to achieve certain aims. This section describes some features that appear to be common to many different networks models.

Regular Coupled Networks

Generally, a globally coupled network has the largest clustering coefficient the least average path length. Although the model of the globally coupled network includes the small-world in addition to large-clustering properties of several real networks, it is simple to notice its limitations: a globally coupled network with N nodes has $N(N-1)/2$ edges, whereas a good number large-scale real networks are sparse in nature, that is, a good number real networks are not completely connected and the number of edges they have is generally of order N rather than N^2.

A well researched, sparse, and regular network model is the nearest-neighbor coupled network (a lattice), comprise a regular graph in which every node is connected to only by a few of its neighbors. A minimal lattice is a simple one dimensional structure, similar to people holding hands in a row. There exists a regular network, which is sparse and clustered, but has a small average path length. A star-shaped coupled network is a common example, in which there is a center node, which connects directly to each of $(N-1)$ other nodes although the $(N-1)$ nodes do not among themselves.

Random Graphs

A network having a completely random graph, which was first studied by Erdös and Rényi [16], is found at the extremely opposite end of a completely regular network in the spectrum. Random graph theory has the aim to establish the probability of connection p that a particular graph property will most probably arise. It was discovered that significant random graph properties can emerge quite suddenly.

Let L_{rand} be the average path length of a random network. Generally in the random network, about $< k >^{L_{rand}}$ nodes are distanced at L_{rand} or very close to it. Hence, $N \sim < k >^{L_{rand}}$, means that $L_{rand} \sim \ln N / \ln < k >$. This increase in the average path length logarithmically is a typical small world effect with regards to the size of the network. Because $\ln N$ increases gradually with N, it contributes to a quite small average path length even in a large network. Alternatively, the clustering coefficient of the ER model is $C=p=<k>/N << 1$ in a totally random network. This means that a very large random network does not show clustering in general.

The Small-World Effect

The well-known experiments performed by Stanley Milgram in the 1960s, where letters were passed from one individual to another to reach a specified individual target using only a small number of steps—around six as published. This result is among the first direct illustrations of the small-world effect, the idea that most networks have pairs of vertices that seem to be linked by a short path through the network.

Watts and Strogatz [9] investigated small world networks by introducing increased amounts of disorder with probability p, in the rewiring of regular networks. Analysis is carried out to investigate the intermediate region between regular and random networks. From their analysis, Small-World Networks (the intermediate) can be highly clustered, akin to regular lattices, and can have smaller characteristic path lengths, like random graphs.

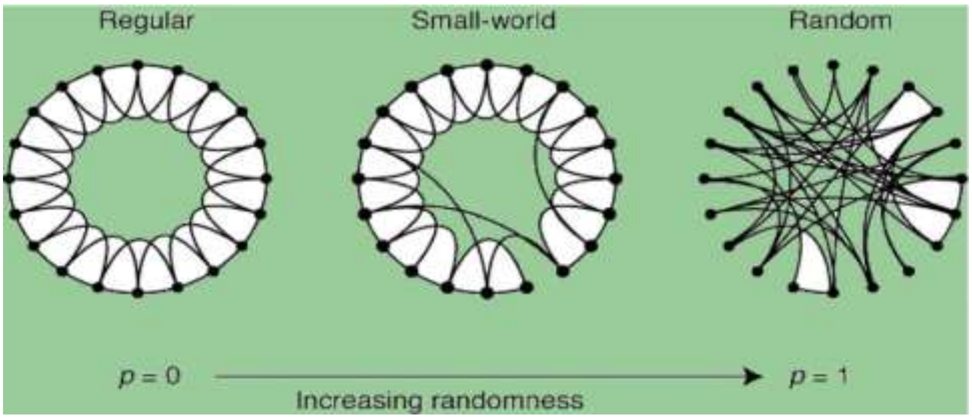

Fig-1: Random rewiring to interpolate between a regular ring lattice and a random network, without modifying the number of vertices or edges in the graph

The small-world effect has clear implications for the several of processes that take place on networks. If a rumor only takes six steps to spread from one person to any other, then the rumor will actually spread much faster than if it takes a hundred steps, or a million. This affects the number of "hops" on the Internet that a packet has to make to move from a computer to another, how many legs of a journey a traveler makes by air or by train, the time it takes for a disease to spread throughout a population, and so forth. The small-world effect as well underlies some popular parlor games, in particular the calculation of Erd˝os numbers [17] and Bacon numbers.

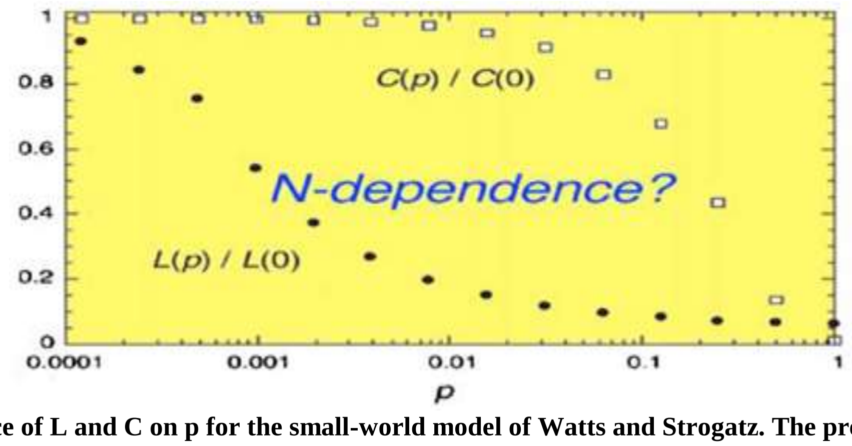

Fig-2: Dependence of L and C on p for the small-world model of Watts and Strogatz. The presence of the small-world is apparent for p 0:01, as L(p) rapidly approaches the random graph value, as C(p) remains in the ordered graph range. After Watts and Strogatz (1998)

Scale-Free Networks

Networks with power-law degree distributions have gained an immense interest in the literature [18, 19]. They are referred to as scale-free networks [20], although just their degree distributions are scale-free, one can and more often than not does include scales present in other network properties.

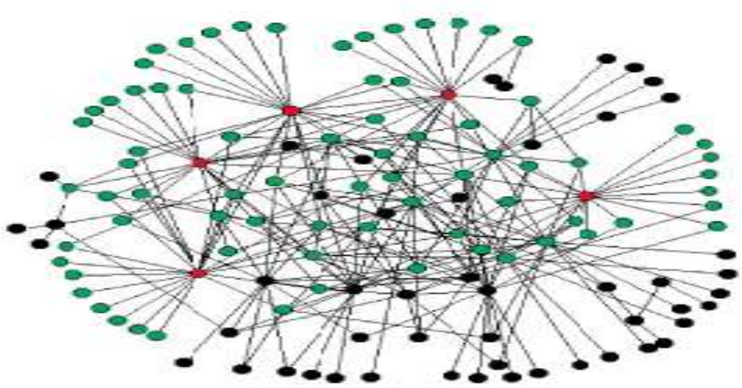

Fig-3: A scale-free network of 130 nodes, generated by the BA scale-free model. The five biggest nodes are given a red color, and they are in contact with 60% of other nodes (green)

The earliest scale-free network publication is probably Price's network of citations between scientific papers [21]. He proposed that the value of α = 2.5 to 3 representing his network's exponent. He quoted in a subsequent paper, a far more accurate figure of α = 3.04 [22]. For the out-degree of the network (which is the number of bibliographic entries existing in each paper), he also established a power-law distribution, even though later work has raised questions [23]. More recently, a host of other networks have observed power law degree distributions, including notably other citation networks [24], the World Wide Web [25], the Internet, metabolic networks [26], telephone call graphs, and the network of human sexual contacts [27].

A part of the (WWW) were mapped by Barabasi and Albert and result displayed as a graph. The result was surprising as the connectivity distribution of the WWW was not even. In contrast, majority of the nodes have one or two links whiles few nodes have a lot of links. If a node is connected to *k* other nodes, its probability was proportional to *k* to the power of a constant: $p(k)\infty k^{-r}$. Networks that demonstrate these characteristics are called scale-free networks (or power-law networks).These includes Social networks, Biological networks, Sexual partners in humans, Semantic networks, Airline networks, P2P networks, Computer networks as well as the internet, and the World Wide Web. For degree distribution, other widespread functional forms used are exponentials, for instance those observed in the power grid and railway networks, and power laws with exponential cutoffs, like those observed in the network of movie actors and some collaboration networks [28].

APPLICATIONS

Application to Social Networks

There have been some attempts to portray the social interactions in animals (association, aggression, submission, grooming) [29]; the networked memberships of football players, musicians, and movie actors [30]; or the interactions of fictional characters, such as the personages of Victor Hugo's Les Misérables, Tolstoy's Anna Karenina or Shakespeare's plays. Scientific co-authorships are examples of collaboration networks that have made available a better insight of the social mechanisms ruling collaborations among scientists. Related networks are the so called citation networks: scientific articles are associated to nodes and a direct connection from A to B signifies that the article B is cited by A.

Application to the World Wide Web

The WWW can be studied at the web-page level, where a node corresponds to a web-page and the hyperlinks are mapped into directed links between nodes. In this case, the degree of a node is made of incoming and outgoing connections. Another possible resolution is the site level, where a node corresponds to a site having a collection of Web pages, and two nodes are connected by undirected edges when there are hyperlinks between Web pages in the corresponding sites.

The WWW can be naturally decomposed in unified groups of pages, identified by the category of the page contents. Pennock et al. have analyzed the in-degree distributions for the following four categories of Web pages: computer science, universities, companies and newspapers [31].

Application to Metabolic, Protein

Jeong et al. have represented the three domains of life by studying 43 different organisms in terms of their metabolic reactions, and have used the metabolites as nodes and biochemical reactions as edges in constructing directed graphs [32].

A category of cellular networks that has acquired interest over the years is that of protein–

protein and protein–gene interaction networks. According to the demand of understanding the protein interaction map, several high-throughput experiments have been performed which provides evidence of a partial interaction map between proteins. The east two-hybrid screen method has been applied for revealing protein–protein interactions by Uetz *et al* [33]. and by Ito *et al* [34].

CHALLENGES

The units that comprise a complex system are highly adaptable and robust, as they do not have strictly defined roles. This feature of complex systems, however, increases the challenges in describing their structure and evolution. The prototypical challenges one faces when applying a complex system at different stages are:

- *The nature of the units*: Units of complex networks normally have complex internal structures; are not identical; and do not have stringently defined roles.
- *The nature of the interactions*: Units of complex networks may have nonlinear interactions; noise; and high complexity of the network interactions.
- *The nature of energy input:* Systems of complex networks may encounter poor quality distribution of external perturbations; poor quality correlations of external perturbations; and instability of external perturbations.

CONCLUSION

This study has analyzed the various complex networks and the models they employ. Critical evaluations of their implications and applications to real networks have been identified. Finally, various challenges to their applications have been defined. The application of complex network to real networks or to model real world scenarios can be of an enormous benefit to humanity. However, complex networks field is a broad area and still requires much research.

ACKNOWLEDGMENT

I want to express my sincere gratitude to my Professor Lu Liu, whose advice and attitude have influenced me very much in a short period of academic interactions. Without his support and guidance, this research could not have been completed.

REFERENCES

1. Newman, M. E. (2003). The structure and function of complex networks. *SIAM review, 45*(2), 167-256.
2. Scott, J., & Carrington, P. J. (2011). *The SAGE handbook of social network analysis*: SAGE publications.
3. Wasserman, S., & Faust, K. (1994). *Social network analysis: Methods and applications* (Vol. 8): Cambridge university press.
4. Moreno, J. (1934). Who shall survive? foundations of sociometry, group psychotherapy, and sociodrama. Beacon, NY: Beacon House: Inc.
5. Rapoport, A., & Horvath, W. J. (1961). A study of a large sociogram. *Behavioral Science, 6*(4), 279-291.
6. Mariolis, P. (1975). Interlocking directorates and control of corporations: The theory of bank control. *Social Science Quarterly*, 425-439.
7. Mizruchi, M. S. (1982). *The American corporate network, 1904-1974* (Vol. 138): Sage Publications, Inc.
8. Watts, D. J. (1999). *Small worlds: the dynamics of networks between order and randomness*: Princeton university press.
9. Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world'networks. *nature, 393*(6684), 440-442
10. Amaral, L. A. N., Scala, A., Barthelemy, M., & Stanley, H. E. (2000). Classes of small-world networks. *Proceedings of the National Academy of Sciences, 97*(21), 11149-11152.
11. Kalapala, V., Sanwalani, V., & Moore, C. (2003). The structure of the United States road network. *Preprint, University of New Mexico*.
12. Latora, V., & Marchiori, M. (2002). Is the Boston subway a small-world network? *Physica A: Statistical mechanics and its applications, 314*(1), 109-113.
13. Chowell, G., Hyman, J., & Eubank, S. (2002). Analysis of a real world network: The City of Portland.
14. Gao, J., Barzel, B., & Barabási, A.-L. (2016). Universal resilience patterns in complex networks. *nature, 530*(7590), 307-312.
15. Albert, R., Jeong, H., & Barabási, A.-L. (2000). Attack and error tolerance of complex networks. *nature, 406*(6794), 378-382.
16. Erd6s, P., & Rényi, A. (1960). On the evolution of random graphs. *Publ. Math. Inst. Hungar. Acad. Sci, 5*, 17-61.
17. De Castro, R., & Grossman, J. W. (1999). Famous trails to paul erdős. *The Mathematical Intelligencer, 21*(3), 51-53.
18. Albert, R., & Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of modern physics, 74*(1), 47
19. Strogatz, S. H. (2001). Exploring complex networks. *nature, 410*(6825), 268-276.
20. Barabási, A.-L., & Albert, R. (1999). Emergence of scaling in random networks. *science, 286*(5439), 509-512.
21. de Solla Price, D. J. (1965). Networks of scientific papers. *science, 149*(3683), 510-515.
22. Price, D. d. S. (1976). A general theory of bibliometric and other cumulative advantage

processes. *Journal of the American society for Information science, 27*(5), 292-306.
23. Radicchi, F., Fortunato, S., & Vespignani, A. (2012). Citation networks *Models of science dynamics* (pp. 233-257): Springer.
24. Redner, S. (1998). How popular is your paper? An empirical study of the citation distribution. *The European Physical Journal B-Condensed Matter and Complex Systems, 4*(2), 131-134.
25. Broder, A., Kumar, R., Maghoul, F., Raghavan, P., Rajagopalan, S., Stata, R., . . . Wiener, J. (2000). Graph structure in the web. *Computer networks, 33*(1), 309-320.
26. Jeong, H., Mason, S. P., Barabási, A.-L., & Oltvai, Z. N. (2001). Lethality and centrality in protein networks. *nature, 411*(6833), 41-42.
27. Jones, J. H., & Handcock, M. S. (2003). An assessment of preferential attachment as a mechanism for human sexual network formation. *Proceedings of the Royal Society of London B: Biological Sciences, 270*(1520), 1123-1128.
28. Newman, M. E. (2001). The structure of scientific collaboration networks. *Proceedings of the National Academy of Sciences, 98*(2), 404-409.
29. Newman, M. E., Forrest, S., & Balthrop, J. (2002). Email networks and the spread of computer viruses. *Physical review E, 66*(3), 035101.
30. Gleiser, P. M., & Danon, L. (2003). Community structure in jazz. *Advances in complex systems, 6*(04), 565-573.
31. Pennock, D. M., Flake, G. W., Lawrence, S., Glover, E. J., & Giles, C. L. (2002). Winners don't take all: Characterizing the competition for links on the web. *Proceedings of the National Academy of Sciences, 99*(8), 5207-5211.
32. Jeong, H., Tombor, B., Albert, R., Oltvai, Z. N., & Barabási, A.-L. (2000). The large-scale organization of metabolic networks. *nature, 407*(6804), 651-654.
33. Uetz, P., & Hughes, R. E. (2000). Systematic and large-scale two-hybrid screens. *Current opinion in microbiology, 3*(3), 303-308.
34. Ito, T., Chiba, T., Ozawa, R., Yoshida, M., Hattori, M., & Sakaki, Y. (2001). A comprehensive two-hybrid analysis to explore the yeast protein interactome. *Proceedings of the National Academy of Sciences, 98*(8), 4569-4574.
35. Padgett, J. F., & Ansell, C. K. (1993). Robust Action and the Rise of the Medici, 1400-1434. *American journal of sociology*, 1259-1319.