Department of Electrical Engineering

FINAL YEAR PROJECT REPORT

BENGEGU4-CDE-2019/20-Project Code

Strengthening a random graph against destructive attacks

Student Name: Chan Kin Lok Gerald
Student ID: 55698763
Supervisor: GRC – Prof CHEN, Guanrong
Assessor: TC – Prof CHOW, Tommy W S

Bachelor of Engineering (Honours) in
Computer and Data Engineering

Student Final Year Project Declaration

I have read the student handbook and I understand the meaning of academic dishonesty, in particular plagiarism and collusion. I declare that the work submitted for the final year project does not involve academic dishonesty. I give permission for my final year project work to be electronically scanned and if found to involve academic dishonesty, I am aware of the consequences as stated in the Student Handbook.

Project Title : Strengthening a random-

Graph network against destructive attacks

Student Name : Chan Kin Lok Gerald | Student ID: 55698763

Signature  | Date : 1/1/2022

Turnitin Originality Report:

Abstract:

In this research that is based on simulation, three different types of typical complex network models will be studied. These models are the random-graph network, the small-world network, and the scale-free network. Using this information, research is conducted on the following issue, especially random-graph networks: Assume that the designer of the random-graph network can increase the network's resistance against both random and targeted attempts to remove nodes and edges by introducing a specific percentage of new edges to the network. This project will simulate and verify the robustness of the strategy of randomly adding new edges to form as many triangles as possible in the original network. After that, it will compare the strategy of random rewiring to the strategy of randomly adding new edges to form as many triangles as possible in the original network. This will help verify that local triangle structure is advantageous in strengthening the robustness of complex networks. I have evaluated the robustness of the network by putting many different codes through MATLAB's simulations. In addition, one may extract a great deal of information from the internet by doing a variety of research projects. Based on the findings, we may reach the following conclusion: a local triangular structure has the potential to improve the resilience of a random graph network and produces superior results when compared to the method of random rewiring.

Acknowledgement

We show our great appreciation to all the authors who collected and shared the data, especially Prof. Guanrong Chen, and the department of Electrical Engineering, City University of Hong Kong, that supports this research.

Contents

List of Figures

List of Tables

Introduction

Research on complex networks is being conducted more often across a wide variety of scientific fields [1-3]. There is little room for question about the fact that many of the systems seen in nature are capable of being represented using intricate network models. Nodes, also known as vertices, are the building blocks of complex networks, which are assembled via connections, also known as edges. There are many instances to choose from. As an example, one may consider the Internet to be a network that is made up of routers or domains. The World Wide Web, sometimes known as WWW, is a network that consists of several individual websites that are linked to one another (Fig. 1). The global economy is made up of a network of national economies, each of which has its own network of markets. Together, these economies make up the global economy. In turn, markets are comprised of networks of producers and consumers that are linked with one another. Networks may be used to represent a wide variety of things, including food webs, metabolic pathways, the connections between words in a language, the subjects that are discussed in a conversation, and even the many approaches that might be used to solve mathematical problems.

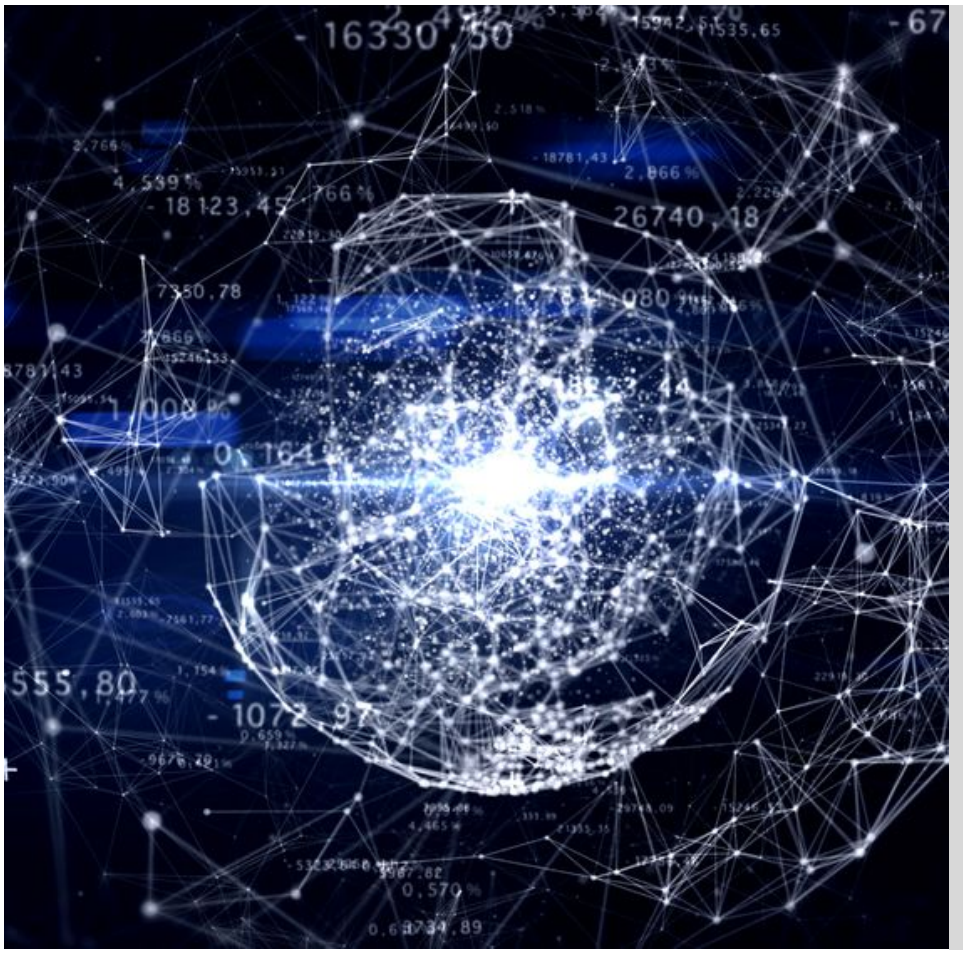

Fig.1 World Wide Web, network that consists of several individual websites that are linked to one another

These innovations include Google, Facebook, CISCO, Twitter (Fig. 2), and a great many more; the foundations of each of these and many other inventions are networks, and only networks. At the end of the day, networks are far more ubiquitous in sectors such as science and technology as well as business and nature than a first glance would imply. When it comes to the job, this is especially important to keep in mind. In addition, it is not feasible to comprehend complicated systems without first having a comprehensive comprehension of the networks that are necessary for their operation.

Fig. 2 Google, Facebook, Cisco, Twitter

In the first decade of the twenty-first century, there has been a revival of interest in network science. This is mostly due to the rising recognition that the underlying networks of all systems are subject to the same set of fundamental rules and principles. After ten years of dedicated investigation, this important discovery was discovered. Therefore, despite the enormous variability in the form, size (including the number of nodes), type (including age), and breadth (including the number of nodes), many actual networks are driven by a common set of organizing principles. This is the case even though there is an enormous amount of variation in each of these aspects. Networks that form have more in common with one another than they do differences between themselves as a result of a lack of attention to the distinctiveness of the components and interactions that make up the network. In the parts that are to follow, this emerging area of study, as well as the effects it has had on the fields of science and technology, as well as on society, will be examined in more depth.

In the first decade of the twenty-first century, there has been a revival of interest in network science. This is mostly due to the rising recognition that the underlying networks of all systems are subject to the same set of fundamental rules and principles. After ten years of

dedicated investigation, this important discovery was discovered. Therefore, despite the enormous variability in the form, size (including the number of nodes), type (including age), and breadth (including the number of nodes), many actual networks are driven by a common set of organizing principles. This is the case even though there is an enormous amount of variation in each of these aspects. Networks that form have more in common with one another than they do differences between themselves as a result of a lack of attention to the distinctiveness of the components and interactions that make up the network. In the parts that are to follow, this emerging area of study, as well as the effects it has had on the fields of science and technology, as well as on society, will be examined in more depth.

Because complex networks are so widely used in science and technology, it is only natural that a set of common and relevant research concerns have emerged about how network structure promotes and constrains network dynamical behaviors. Because traditional fields make such significant use of complicated networks, it has been common practice in academic research on traditional subjects to essentially disregard concerns like these. This is the nature of the difficulties that we come with daily, situations that need us to provide explanations and replies.

This assumption is implicit in many models, but it underpins the modeling of a wide variety of physical and non-physical systems over the course of more than a century. The basic idea behind these models is that the interactions between people can be mapped onto a regular and possibly universal structure, such as a Euclidean lattice. This assumption has been used as the foundation for modeling not just physical but also non-physical processes and systems. Erdos and Rényi, better known by their initials ER, were two mathematicians who, in the late 1950s, made significant contributions to the classic mathematical graph theory. They depicted a network with a complicated topology by using a graph that was completely random [4]. (Fig. 3) Their work laid the foundation for the random network theory, which has subsequently been the subject of intense research over the course of the last 40 years and is still being investigated to this day. Over the course of more than half a century, the ER random graph model was the preeminent way that academics thought about complex networks. This was the case even though many real-world complex networks are neither completely regular nor perfectly random. This occurred because there was a shortage of both supercomputers and thorough topological information on extremely large networks in the actual world.

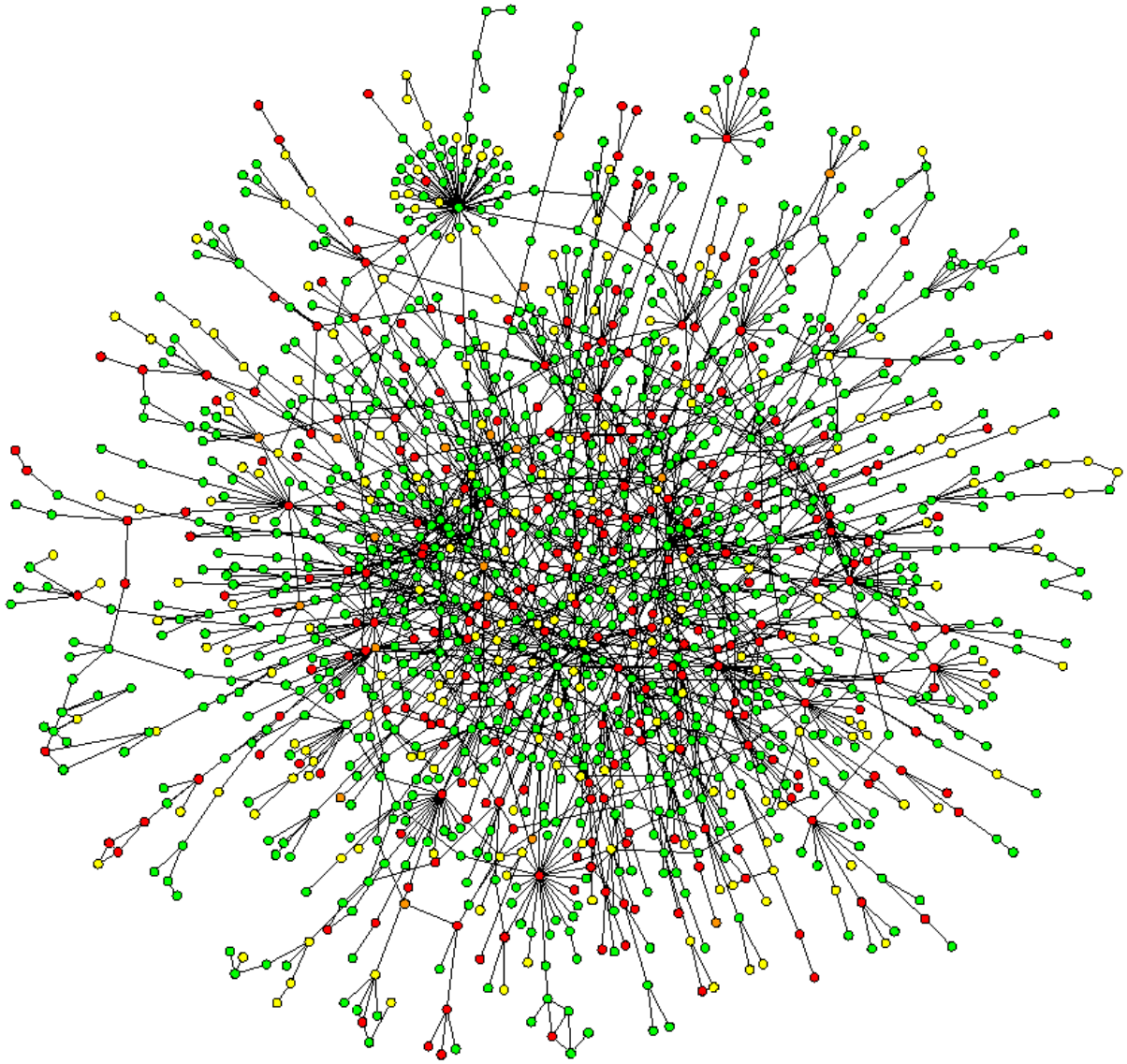

Fig. 3 Random Graph Dynamics

Large databases have been constructed on a wide range of real networks with intricate topologies as data collection processes have gotten increasingly automated and as high levels of processing power have been more easily accessible. As a result of the ease with which they can now get their hands on so much real data, an increasing number of people are interested by the parallels and contrasts that exist across the many kinds of sophisticated networks. Two recent discoveries in this area are the so-called "small-world effect" and the "scale-free feature," which can be found in most huge networks.

In 1998, Watts and Strogatz (WS) proposed the concept of a small-world network in order to describe the transformation from a regular lattice to a random graph. [5] This was done to quantify the change. [5] This was done in order to demonstrate how a network evolved over the course of time. It is quite important to keep in mind that the phenomenon of the little world is a widespread one. As a direct result of the realization that they are acquainted with the same person, two people who have never met before will remark, "What a little world!" upon their first encounter. Those who have this intuition are the ones who are taken aback when they learn that they are acquainted with another person in a similar capacity. The concept of "six degrees of separation" was made public by social scientist Stanley Milgram in the late 1960s [6]. Although the "small world effect" had been known for some time, the phrase "six degrees of separation" was. Even though this assumption is still up to debate, the small-world pattern has been seen in a significant number of real networks. This is because people tend to have a limited number of close connections. In both the ER random graph model and the WS small-world model, the connectedness distribution of a network starts with an average value, then gradually decreases from that point forward in an exponential fashion. This is a trait that appears rather often. These kinds of networks are often referred to as "exponential networks" or "homogeneous networks" (Fig. 4) since each node has about the same number of link connections.

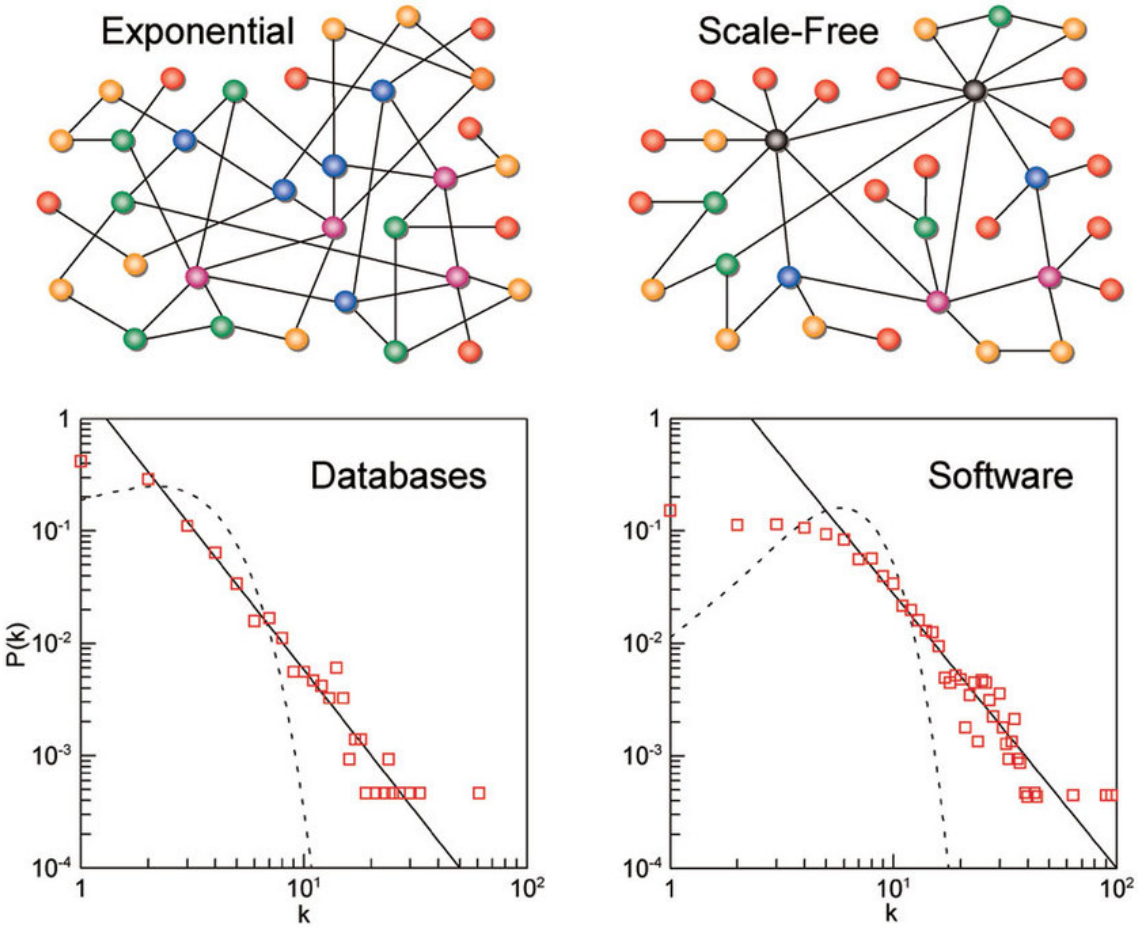

Fig.4 Exponential Network and Scale-Free Network

Another significant step forward in the development of the science of complex networks was the discovery that a great number of large-scale complex networks are scale-free. According to this, the pattern followed by the distribution of connections in these networks seems to be a power-law distribution [7, 8], independent of the size of the network. A scale-free network is not homogeneous, in contrast to an exponential network, which is more so. The majority of nodes in the network only have a few link connections, however there are a few nodes in the network that have a large number of link connections.

Over the course of the last several years, the findings concerning the small-world effect as well as the scale-free feature of complex networks (Fig. 5) have made a major contribution to the expansion of knowledge in the field of complex network theory. However, there are still many areas of it are waiting to be discovered.

Fig, 5 Complex Network

Following two decades of intensive and extensive research in a variety of scientific and technological domains, complex networks have recently gained popularity as a subject worthy of investigation on their own [1–3].

In today's hyper-connected world, the traditional systems control theory is developing in order to deal with large-scale networks of dynamical systems and networks and subnetworks of complex dynamical systems. This is being accomplished through the investigation of four aspects of these systems: their controllability, synchronizability, observability, and stability, as well as the design of efficient controllers for a variety of control objectives and applications.

Since the concept of state controllability has been introduced in directed networks, numerous new research difficulties as well as theoretical and technological barriers have emerged in the fields of network science as well as control systems. These barriers can be broken down into two categories: theoretical and technological.

In the last ten years, there has been a significant amount of progress made in the field of research pertaining to the robustness of complex networks. These advancements are documented and appraised in [4]–[6]. Investigations into network robustness have been carried out from a wide range of perspectives and on a wide variety of network topologies, such as random-graph networks, small-world networks, scale-free networks, and many more. Recently, academics have been concentrating their efforts on network controllability's resistance against a range of detrimental attacks on different network topologies (Fig. 6). This is a relatively new area of study.

As a result, I conducted this study to examine what if the designer of a random-graph network can increase the network's resistance against both random and targeted attempts to remove nodes and edges by introducing a specific percentage of new edges to the network and prove that local triangle structure is advantageous in strengthening the robustness of complex networks.

Fig. 6 network controllability's resistance against a range of detrimental attacks on different network topologies

In conclusion, networks have recently developed as a paradigmatic tool to show systems in which the pattern of interactions between the components is both complex and dynamic. Complex systems may be described in a variety of different ways, but networks have recently emerged as a paradigmatic technique. The movement of both people and merchandise is the primary purpose served by these dynamic networks. As a consequence of this, it is essential for us to have an understanding of the connection that exists between the structure of the network and the efficiency as well as the resilience of the transportation system. We will be able to better govern cellular communication once we have a more in-depth knowledge of cell signaling. This will make it possible for us to build more reliable systems for transporting information and energy, such as the Internet and the power grid. On the other hand, due to the inherent characteristics of network architecture, it is not feasible to get such an understanding. In this research, a variety of distinct complex networks are analyzed in terms of both their controllability and their resilience. A simple formula may be used to make a comparison between the controllability and resilience of different networks. The metric takes into consideration the network's controllability after the removal of each node and edge in the network.

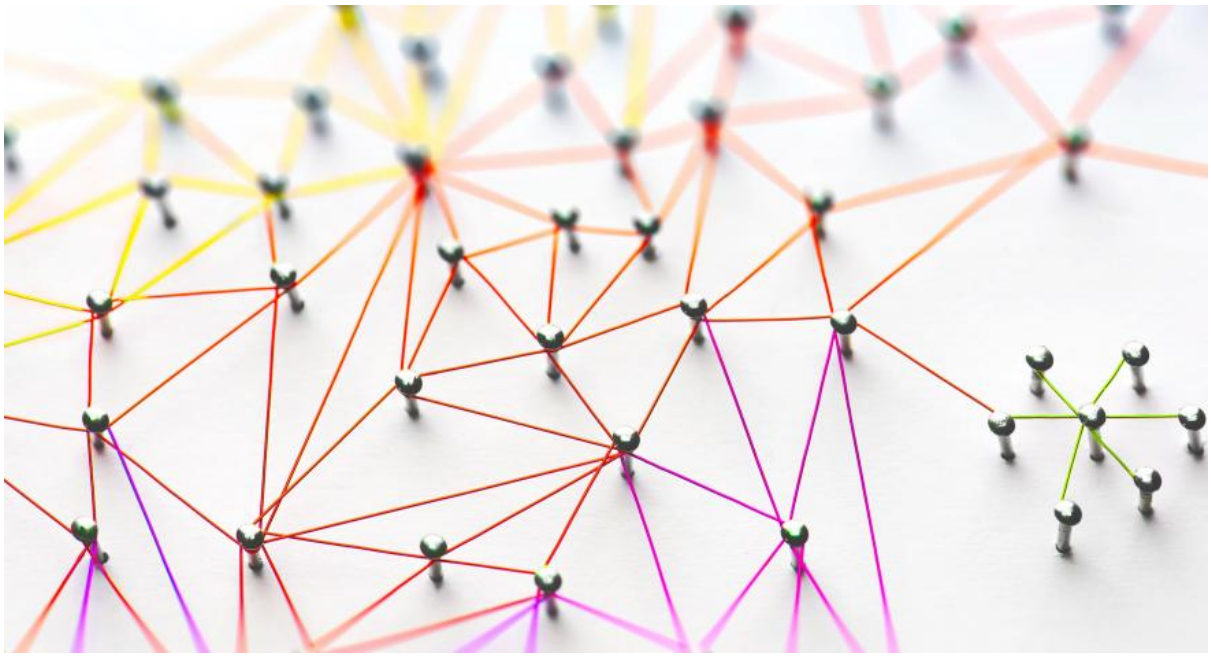

Fig. 7 A network of network

Background

Before I start my detail explanation, I would like to give a brief explanation about some special terms I would use about the network and graph theory.

In the last several decades, various quantities and measurements of complex networks have been suggested and studied, but the current research and development of complex networks theory is principally driven by three notable ideas: the average path length, the degree distribution, and lastly, the clustering coefficient. The development of the discipline is heavily dependent on these notions.

Watts and Strogatz [5] set out to create a network model with a short average route length as a random graph and a high clustering coefficient as a regular lattice in their study on small-world networks. The new network model that we use today evolved from this paradigm. It was only by looking at degree distributions that power laws in many actual networks could be discovered, which led to the discovery of scale free networks. No matter how many different systems and processes power law distributions have been studied in physics, the power law distributions provided the motivation for scaling-free networks to be discovered.

Average path length

One of the factors that is used in the process of determining the distance between two nodes, *i* and *j*, in a network is the number of edges that constitute the shortest path that can be taken between them. Because of this, the *diameter* of a network, denoted by the letter *D*, is measured as the farthest possible distance that may separate any two of its nodes. Because of this, the distance that separates any two nodes in the network is considered to be the network's *average path length*, which is abbreviated as *L*. The "effective size" of a network, also known as the distance between any two nodes, is determined by the metric known as *L*. A friendship network may be represented by the letter *L*, which stands for the typical number of friends who provide the most robust and immediate connection between any two individuals. Even though there are fewer nodes in the network, the majority of genuine complex networks have relatively short pathways. This contrasts with the situation in a typical globally linked network, which has the same number of nodes but longer paths. Even though there are still the same number of nodes in the network, this is the situation. Small-world networks got their name because the image they provide of a smaller world, which is mirrored in the size of the networks themselves, led to their being given this moniker.

Degree distribution

The *degree* of a single node is the quality that is both the most elementary and, potentially, the most significant. It is common practice to define the *degree* k_i of a node *I* as the total

number of connections that it has. Therefore, the "more essential" a node is in a network, the higher the *degree* at which it is located. The *average degree* of the network, indicated by the symbol $<k>$, is calculated by taking the value of k_i and averaging it across all i. The distribution function $P(k)$ is the chance that an arbitrarily chosen node has precisely k edges. This function is used to quantify the dispersion of node degrees over a network. Because all of the nodes in a regular lattice have the same number of edges, it is possible to calculate a simple degree sequence for the lattice. As a result, a plot of the degree distribution will only have one distinct spike which is called the delta distribution. The form of this peak will become wider if there is any unpredictability in the network. The typical Poisson distribution is followed by the degree sequence in the limiting situation of a totally random network. The form of the Poisson distribution decreases exponentially as it moves away from the peak value, which is denoted by the value $<k>$. As a result of this exponential fall, the chance of discovering a node that has k edges gets so low that it may be considered negligible when $k >> <k>$.

In recent years, the degree distribution of a great number of large-scale real networks has significantly departed from the Poisson distribution in a significant way. Many different types of networks have a degree distribution that is best described by a power law of the form $P(k) \sim k^{-\gamma}$. The power-law distribution is less abrupt than the exponential distribution; hence, it permits the occurrence of a few nodes with exceptionally high degrees. *Scale-free networks* are power-law degree distribution networks that do not have a defining scale and are thus referred to by that name.

Clustering Coefficient

It is possible that one of your direct friends in your friendship network is also a friend of one of your friends' friends; to put it another way, two of your friends may be friends. This characteristic makes reference to the clustering of the network. The average fraction of a node's neighbors who are also neighbors of each other may be used to construct a clustering coefficient, which is abbreviated as C. Let's say that one of the nodes in the network, I, has k_i edges that link it to k_i other nodes. These other nodes make up the perimeter of Node I. When every node I 's neighbor is connected to every other node I 's neighbor, there can be no more than $k_i (k_i - 1)/2$ edges between them. This is the only scenario in which this number of edges may exist. The number of edges that are really globally connected, which indicates that every node in the network connects to every other node, is referred to as the clustering coefficient C_i of the node I. This coefficient is defined as the number E_i of edges. In a completely random network with N nodes, the number $C \sim 1/N$ is a very low value in compared to the majority of real-world networks. It has been observed that the majority of large-scale real networks cluster, with clustering coefficients that are much larger than $O(1/N)$ but still significantly lower than one. As a direct consequence of this, the vast majority of real-world

complex networks are not completely determined by chance. Because of this, it is inappropriate to think of them as equally random and completely linked lattices.

Typical network models

The average route length (L), the clustering coefficient (C), and the degree distribution (P(k)) are some of the essential properties that must be measured in order to have a better understanding of the structure of a complex network. The next step is to create a mathematical model with a topology and statistical properties that are similar. This will enable mathematical analysis to be done on the resulting platform.

In order to create proper mathematical network models, which are required to understand the interactions between the topology and dynamics of a complex network, it is vital to research and appreciate the structural features of real-world complex networks. Only then would it be feasible to comprehend the linkages between a complicated network's structure and dynamics. Since Watts and Strogatz found small-world networks and Barabási and Albert discovered scale-free networks, a number of case studies on a range of real-world networks have been conducted and reported on from various perspectives. As a consequence of the aforementioned findings and experiences, a number of unique complex network models have been introduced. This section covers a variety of basic network models, such as regular coupled networks, random-graph networks, small-world networks, and scale-free networks.

Regular coupled networks

The mathematical concept of a "regular graph" does not necessarily apply when the word "regular network" is used in this context. A network that is entirely interconnected is one of the most prevalent forms of regular networks which is a complete graph. Edges exist between all nodes in a network of this kind (Fig, 8).

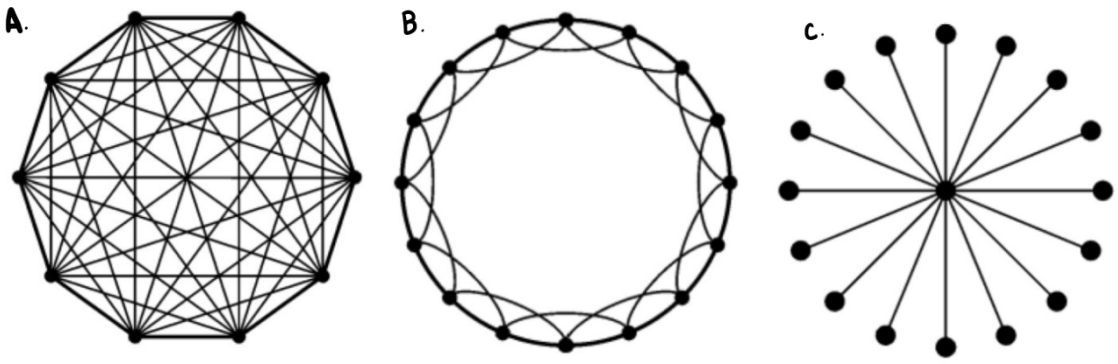

Fig. 8 Some simple networks: (a) fully-connected network (b) ring-shaped network (c) star-shaped network

On the surface, it would seem that a globally interconnected network would have the shortest average route length and the highest clustering coefficient possible. In spite of the fact that the globally connected network model can match the small-world and large-clustering properties of many real networks, its disadvantages are easy to spot: Large-scale networks seem to be sparse despite the fact that a globally connected network with N nodes has $N (N - 1)/2$ edges. Thus, most actual networks are not fully linked and contain a number of edges that is on the order of the magnitude N rather than the magnitude N^2.

The nearest-neighbor connected network is shown by the regular lattice graph. Each node is only linked to a tiny number of its neighbors in this kind of network. Network model that has been widely studied and is both sparse and regular has been developed. When you hear the phrase "lattice," you may picture a two-dimensional square grid, but there are a plethora of different possibilities. A line of individuals holding hands forms a minimum lattice, which only exists in one dimension. In a nearest-neighbor lattice with a periodic boundary condition, each of the N nodes in a ring is i close to its neighbors, resulting in a network of N nodes. If K is an even number, i may take on the values $1, 2, \ldots K/2$. In reality, the clustering coefficient of a network linked to its nearest neighbor is somewhere around $C = 3/4$ when the number of neighbors is big.

A small-world network, on the other hand, is not the same as a network that is linked to its immediate neighbors. Despite the fact that the average trip distance is so great, it is rapidly approaching infinity. As a result, any dynamic activity that requires global coordination, such as synchronization, would be difficult to do in a network made up only of locally connected nodes. What is the average length of individual routes in a typical network that is both sparse and clustered? That's the way it works. Consider a linked network in the shape of a star, where each of the other $N - 1$ nodes only connects to the center node and has no other connections. The average route length is shorter when there are more nodes in a network, coming in at 2, while the clustering coefficient is lower, coming in at 1. Many real-world networks, including their fascinating properties such as sparseness, clustering, and small-worldness, may be correctly represented using the star-shaped network model. As a result, it is a better representation of many well-known actual networks than a traditional lattice. It is hard to deny, however, that the great majority of real networks do not have the exact form of the star they seem to be.

Random-graph networks

On the opposite end of the spectrum from a perfectly regular network is a network with a completely random graph, which was first examined by Erdös and Rényi (Fig. 9) (ER) around 40 years ago [4].

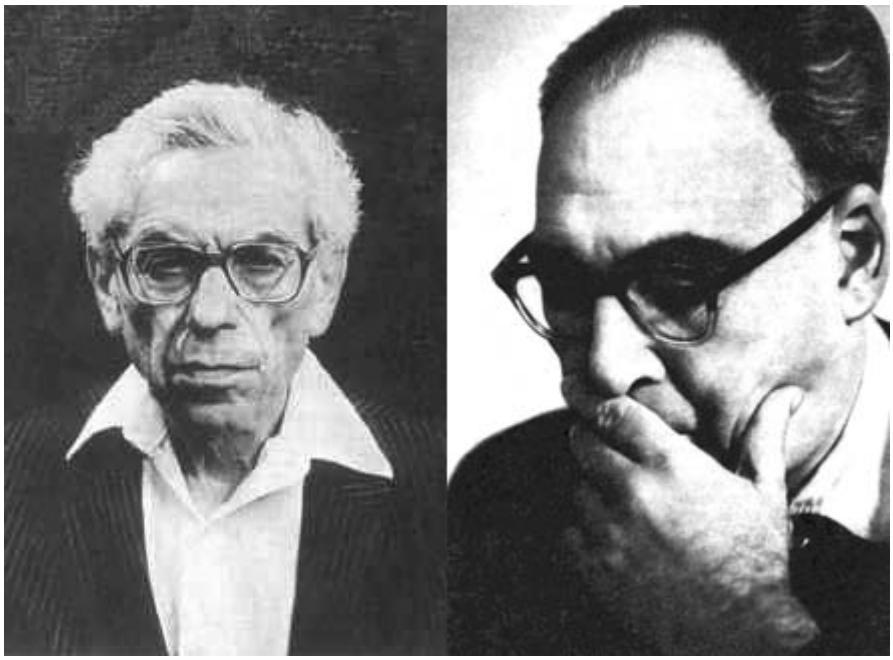

Fig. 9 Erdös and Rényi

Assume there are many buttons strewn across the floor. This integer, where N >> 1, is a multiple of one. To use a thread to connect every single pair of buttons, and the chances of you doing so are *p*. The end outcome is an ER random network with *N* nodes and roughly *pN(N - 1)/2* edges, which is a physical representation of the network (Fig. 6). Random graph theory's main goal is to figure out at what level of connection probability, represented by probability *p*, a given network attribute will be most commonly seen. The prospect that substantial features of random graphs may emerge in an unanticipated way is one noteworthy example of this kind of finding. We can see that if we are discussing how many more buttons would you be able to take up if you were to pull up one button, the probability *p* is greater than a particular threshold termed $p_c \sim (\ln N)/N$, practically every random graph is linked, according to ER. This discovery implies that you can pick up all of the buttons on the floor by picking up one at a time.

$< k > = p(N - 1) \cong pN$ will be the average degree of the random graph is indicated by *pN*. Let's suppose L_{rand} is the average length of a random network's route. Intuitively, about $< k >^{Lrand}$ nodes of the random network are placed within L_{rand}'s distance or quite near to it. As a direct result of this, $N \sim < k >^{Lrand}$, implying that $L_{rand} \sim \ln N / < k >$. The smallworld effect is shown by the logarithmic growth in average route length with increasing network size. Because *ln N* rises slowly with *N*, even in a large network, the average path length may be fairly short. This is due to the fact that *ln N* climbs slowly as *N* increases. In a completely random network, such as a friendship network which is completely random, the likelihood that two random people from your network are friends is not larger than the probability that

two random people from your network are friends. As a result, $C = p =< k> /N << 1$. is used to derive the clustering coefficient for the ER model. This implies that, in general, a random network with a large size will not show any clustering. In reality, the ER method generates a linked homogeneous network that, for a greater N, comes near to resembling a Poisson distribution.

ER random-graph network can be generated as follows:

1. For initialization, begin with N isolated nodes.
2. Chooses one instance of each potential pairing of nodes from the N given nodes, and then links them with an edge using a probability p $\in$ (0,1).

According to statistical analysis, the predicted number of edges in such a network is provided by the statistic pN (N - 1) / 2.

In general, the density of the resultant network will rise in proportion to the magnitude of the parameter p, as illustrated in Figure 3-2. If p is equal to zero, the previously unconnected nodes will remain disconnected; if p is equal to one, then the network will be fully joined.

Note that the previous approach of generating a random-graph network will not create multiple edges or self-loop at any one node; as a result, all of the outputs are simple graphs.

Also, due to the random nature of the connection probability p, numerous runs of the generating procedure described above, even when using the same probability p for joining nodes, yield networks that are only slightly different from one another. Despite this, each of the resulting networks is similar in various ways and has a common set of properties.

The research on ER random-graph networks has mostly focused on solving the following question: for what values of p will the produced graph have the required features? Above all, Erdos and Rényi discovered that many essential characteristics of such random graphs appear unexpectedly. This may be interpreted in the sense that given a collection of random graphs created by the probability p, practically all of these graphs either have a specific property represented by the letter P or do not have this characteristic labeled by the letter P. However, Erdos and Rényi's most significant finding was that many key features of such random graphs arise.

When p exceeds a specific threshold, such as $p_c \sim ln / N$, practically all random graphs created using the approach described above will become linked, while before, virtually all such graphs were unconnected networks. It contained several parts of separate clusters, therefore p had to be larger than or equal to this criterion.

An outstanding illustration of a Poisson distribution of an ER random-graph network may be seen in Figure 10. As k goes further away from the average value $<k>$, which would be the

predicted value μ in the Poisson distribution function described earlier and the point at which the function achieves its maximum value, it is abundantly evident that the Poisson curve decreases in value in an exponential fashion.

In a manner similar to that described for the delta distribution up above, this indicates that a node in the network with a degree *k* that is either relatively or excessively low will be difficult to identify, if not impossible to discover. To put it another way, every node in such a random network has nearly the same degree, at least in concept. This is true even if the network is not perfectly connected. This particular kind of network is referred to as a *homogenous network*.

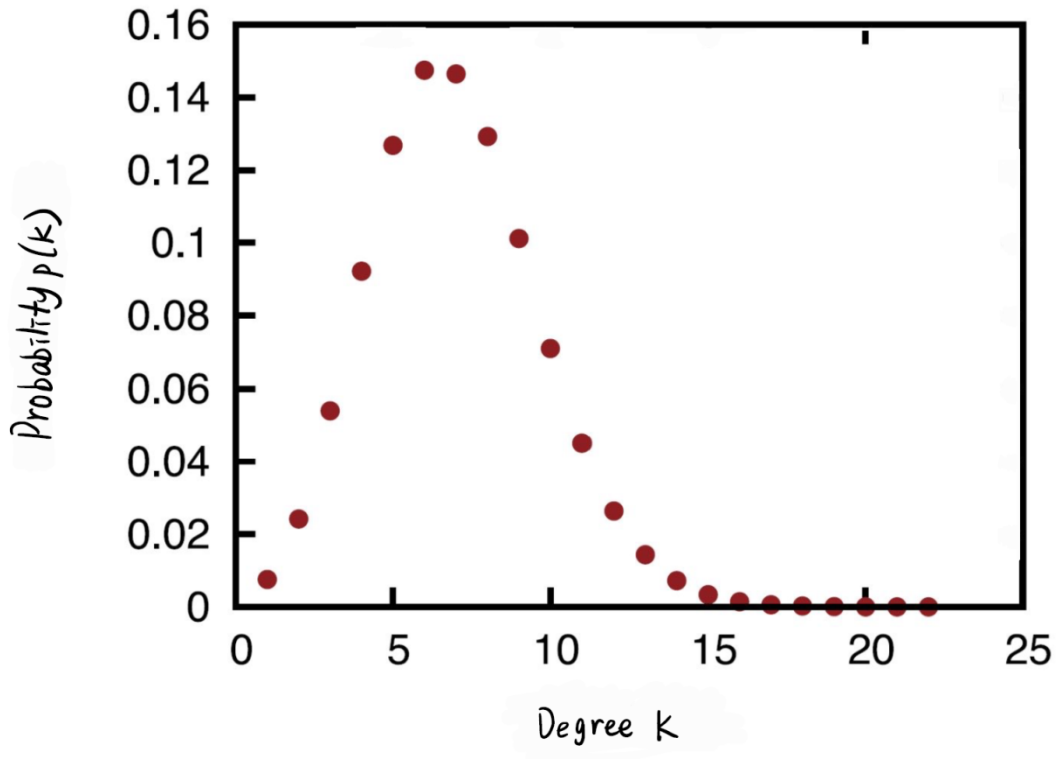

Fig.10 illustration of a Poisson distribution of an ER random-graph network

The traditional ER random-graph network theory has been used, updated, and expanded upon; as a result, it has been utilized to provide a more accurate description of a large number of complicated real-world networks. One can construct multiple degree sequences $\{k_i\}_{i=1}^{n}$, which results in many networks with *N* nodes, all of which have the same node-degree distribution *P(k)*, as an example, for a desired node-degree distribution *P(k)*, which reflects the fraction of nodes in the network with *degree k*. This can be done to achieve a desired node-degree distribution *P(k)*, which reflects the fraction of nodes in the network with *degree k*. A *configuration model* is a collection of many of these networks that may be used to

explain a variety of real-world networks. This collection can be thought of as a model for the actual world.

Small-world networks

Even though regular lattices are clustered, they do not exhibit the small-world effect. On the other hand, random graphs exhibit the small-world effect but not clustering. Both the regular lattice and ER random models, as a consequence, overlook important elements of many real networks. The great majority of real-world networks, on the other hand, are neither entirely random nor perfectly regular. Neighbors are known by many individuals, but the lattice model suggests that they are the only ones they know. Contrary to popular belief, links between websites on the Internet were most likely not created at random, as the ER method would have us believe.

Watts and Strogatz [5] proposed the small-world network model known as WS small-world network first to characterize the shift from a regular lattice to an unexpected graph. The following steps may be used to break down the construction of the WS model.

1. To begin, we will construct a nearest-neighbor connected network consisting of N nodes arranged in a ring. In this network, each node will be located in close proximity to its neighboring nodes, with $I = 1, 2, ..., K/2$, and the value of K will be an even number.
2. We introduce an element of randomness into the network by rewiring each edge with a probability p, varying p in such a way that the shift from order $p = 0$ to randomness $p = 1$ can be carefully monitored.

A random node is picked at random from the whole network to be the new end of the connection in this case. No two nodes may have more than one connection to each other, and there is no way for a node to join to another node. Through the use of this approach, $pN K/2$ lengthy edges are generated. To avoid the need for many neighborhoods, these edges link together nodes that would otherwise be isolated. As variables of rewiring probability p, the clustering coefficient $C (p)$ and average route length $L(p)$ in the WS small world model may both be thought of. This is conceivable because the frequency with which the network is rewired affects both variables. Ring lattices often have lengthy average route lengths, although the structure is crowded. It is extremely fast to lower the average route length if there is a low risk of rewiring, if local attributes are roughly comparable to those of the original regular network, and if the clustering coefficient doesn't fluctuate from its initial value. Not surprisingly, this conclusion has come to light. However, a few random rewiring may reduce the average distance traveled by routes by a substantial amount. Several rewired connections will not significantly impair the network's local clustering, on the other hand.

Alternatively, the small-world model may be seen as a homogeneous network in which each node has about the same amount of links to other nodes. The ER random graph model and the WS small-world network model are similar in this respect. Because of the work on WS small-world networks, researchers have been scrambling to come up with more sophisticated network models, including numerous iterations of the WS model. It was suggested by Newman and Watts [23] and has since been dubbed the NW small-world model after them. Nodes in the NW model are connected with a probability of p rather than two nodes that are the closest neighbors to each other being disconnected. There can only be one connection at a time between nodes, and they can't connect to themselves in any way. A globally coupled network is formed if p is greater than zero, while a locally coupled network is formed when p is less than zero. The NW model is easier to understand than the original WS model because it does not lead to the formation of separate clusters, unlike the WS model, which does. The NW model and the WS model are nearly identical when p and N are sufficiently small and large, respectively. For the purpose of simplicity, these two models are now often referred to as "small-world models." (Fig. 11)

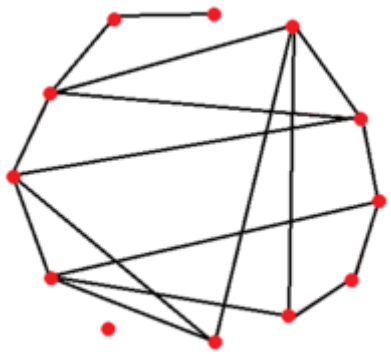

Fig. 11 A Small World Network

According to social network theory, most individuals have close ties with their neighbors or coworkers, a fact that may be traced back to small-world models. According to the NW model and the wire-rewiring model, many individuals have a small number of close friends who live far away, and these friends are represented by long-range edges. Friends who reside on the other side of the world are likely to have long-distance relationships.

Scale-free networks

A homogeneous connection distribution is predicted by both the ER random graph model and the WS small-world model. From this point on, the connection distribution is expected to decline dramatically. An exponential network is one kind of network. One of the most recent findings in the subject of complex networks is that a lot of large-scale complex networks,

such as the World Wide Web, the internet, and metabolic networks, are scale-free and based on a power law structure.

Using a new network model, Barabási and Albert (BA) (Fig. 12) explained the genesis of the power-law degree distribution [7,8]. They argued that the bulk of current models fail to account for two important aspects of real-world networks. Although the number of nodes may be modified, the number of edges remains constant throughout the construction process of certain models. Adding additional nodes to existing networks is a constant process in actual networks. It's impossible to keep track of all the websites and scholarly papers that are constantly being published on the Internet and in academic journals. When establishing new edges, the random graph model and the small-world model both need that the chance of each result be equal, which is unworkable in practice. Accordingly, websites with more links are likely to get greater traffic; a new publication is more likely to reference well-known and frequently-cited works rather than many other works. Wealthy people getting more affluent is a phenomena that is not captured by other models.

Fig. 12 ALBERT-LÁSZLÓ BARABÁSI

According to the BA model, network self-organization in a scale-free structure relies heavily on growth and preferred attachment. Networks often expand over time by adding new nodes, which are typically linked to other existing nodes that have a large number of connections,

according to these findings. The following is an example of a scale-free model's generating strategy:

1. Beginning with a small number of nodes (*m0*), a new node is added at each time step and is connected to the $m \leqslant m0$ which were already connected.
2. The likelihood that a new node will be connected to node I, which is one of the m already-existing nodes, is proportional to node I 's degree ki, using the formula $\Pi_i - k^{i\!\!\!/}/\Sigma_j k_j$. This kind of attachment is referred to as preferential attachment.

The use of this strategy will ultimately result in a network that has $N = t + m_0$ nodes and mt edges once *t* time steps have passed. As the network continues to grow, it will ultimately arrive at a state that is scale-independent if it meets these two criteria: The pattern of degree distribution has not changed throughout the course of time, which is to say that it does not evolve in a distinct manner as the size of the network increases. The likelihood of discovering a node that has *k* edges is proportional to k^{-3} in the appropriate degree distribution. Additionally, this probability rises as the number of edges in the network grows. The suitable degree distribution is represented by a power law with a exponent of -3.

The scale-free model's (Fig. 13) clustering coefficient seems to be much greater, as shown by the numerical data. In comparison, the scale-free model's average route length is just slightly shorter than the average degree of a random graph of the same size. In other words, the presence of a few "big" nodes in the network, those with high degrees and a significant number of connections, is critical to connecting the other nodes. On the other hand, no analytical technique to predicting the average route length and clustering coefficient of the scale-free model exists yet. Simple and easy, the BA model explains how the power-law degree distribution is formed. When compared to the various networks that exist in the real world, this model has a number of obvious flaws.

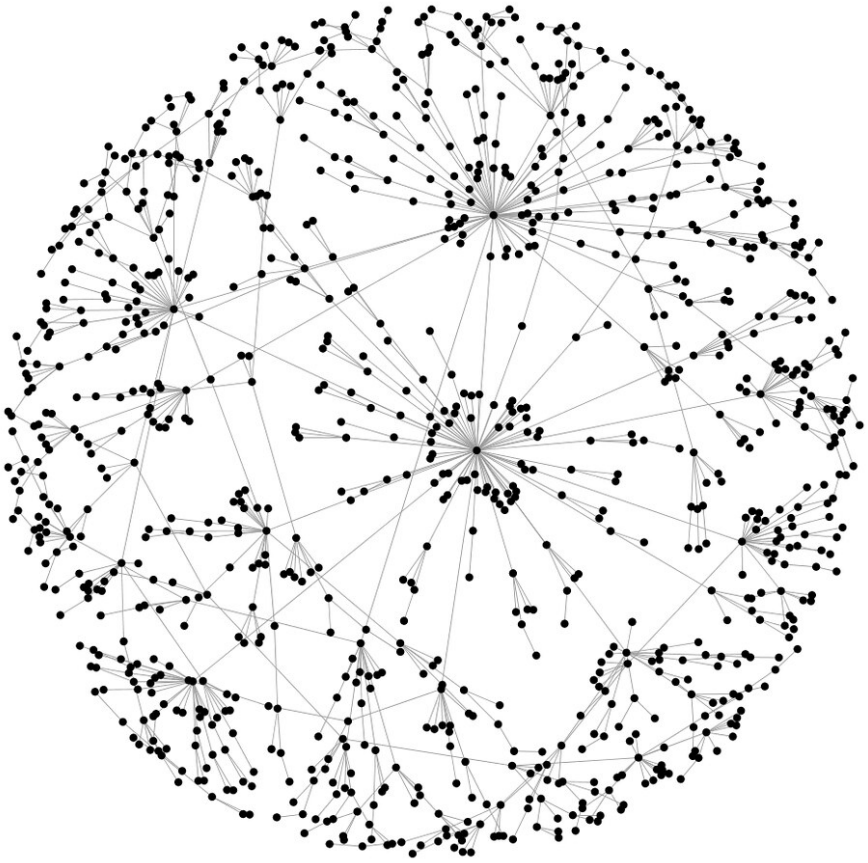

Fig. 13 Scale-Free Network Model

Due to this revelation, there has been an increase in network formation research in an effort to overcome the constraints of the BA model. Albert and Barabási summarize the various models in their essay [2].

Someone recently [24] recognized "network motifs" as patterns of links that occur in intricate networks in considerably higher numbers than in entirely random networks. These patterns may be distinguished from fully random networks by their complexity. The fields of biology, neurology, ecology, and engineering are just few of the many that have yielded results similar to this while searching for patterns. This investigation could result in the identification of the essential building components that are specific to each form of network.

Network Robustness

Every human-created product has the potential to be defective in some way: You may require a tow truck if a component in your vehicle's engine breaks, but a wiring issue in your computer's chip might render it worthless. Many natural and social systems, on the other

hand, have a remarkable ability to maintain their basic functions even if a percentage of its components are damaged or absent.

As a result of the huge number of misfolded proteins and incomplete chemical reactions in our cells, we are seldom aware of their repercussions. Even if a significant section of a big company's employees is not present, the organization may continue to operate smoothly. Finding out how this toughness is cultivated might be useful in a variety of fields:

- The concept of robustness is very important in the fields of biology and medicine since it helps to explain why some mutations may cause diseases while others cannot.
- Researchers in the fields of ecology and environmental science are now struggling with the question of when an ecosystem might collapse due to human activities.
- To create communication systems, automobiles, or aircraft that are able to carry out their fundamental duties despite the loss of individual components is the pinnacle of achievement in the field of engineering.
- Those social scientists and economists who study how well human societies and institutions hold up under adverse conditions like hunger, war, and shifts in the status quo are alarmed by this development.

Many types of systems, including biological, social, and technological ones, rely on networks to function properly. A cell's robustness is encoded in its complex regulatory, signaling, and metabolic networks; a society's resilience is tied to the intricately woven social, professional, and communication web that sustains it; and an ecosystem's survivability cannot be understood without a thorough examination of the food web that sustains each species. Nature creates networks to give resilience wherever it is needed.

Achilles' Heel of Complex Networks

In complex networks, there is an intriguing "Achilles heel" that takes the shape of a conflict between robustness and fragility. Let's take a look at a vast network with a number of distinct connections as an example. Remove one of the nodes at each time step. When a node is removed, all of its connections with other nodes are likewise removed. This may cause some of the network's remaining nodes' paths to become distorted. If one of the routes connecting nodes I and j is disrupted, the distance d_{ij} between them may rise, contributing to an increase in the network's average path length L. If one of the pathways between the nodes is disturbed, this may happen. If the situation is severe enough, the loss of only one route between I and j might result in the two nodes being cut off from one other. Even if just a small fraction of the network's nodes are still up, if the network is part of a large cluster with many nodes, the connection to the network may be deemed resilient or error-tolerant.

A branch of the US Department of Defense's Advanced Research Projects Agency built the ARPANET in the late 1960s, which was the predecessor of the Internet. If any of the ARPANET's subnetworks or gateways were down at any one time, the ARPANET's main duty was to ensure that communication services were always available to users. Nowadays, the Internet has grown to be a massive network that affects almost every area of contemporary life. Wherever individuals go in today's world, their effect may be felt. Whether or whether the network can be maintained in the face of unavoidable faults or continuous cyberattacks is a valid question to ask. A positive outcome of our experiment was that we were able to demonstrate that the Internet would not collapse even if more than 80% of the nodes failed. This is an important discovery. Even a small number of nodes may have the same effect as a large number of nodes being deliberately deleted. If he focused on a few key nodes with a large number of connections, he may achieve this impact. Error tolerance and attack vulnerability have been shown in scale-free networks [25-28].

These characteristics stem from the fact that the underlying structure of degree distributions in scale-free networks is notoriously non-homogeneous. The Achilles' heel refers to a potentially exploitable vulnerability in complex networks named after the mythological hero Achilles, who was magically protected in practically every part of his body save one little place - his heel.

During the Trojan War, Achilles was the most fearsome warrior in Greek mythology. A son of Thetis and Peleus, he was born. When Achilles was born, legend has it that his mother Thetis threw him into the Styx River in an attempt to make him eternal. In the process of drowning him, she failed to soak the other heel she was holding while holding him by one heel (Fig. 14). She was guarded from the Styx's magical water, but the portion of her body that she used to hold him was still exposed. An arrow thrown by an enemy hit Achilles in the heel, killing him in the midst of combat. When a powerful entity has a weakness, the word "Achilles' heel" might be used to describe it. Fig. 3-13 (b) [16] illustrates "The Achilles' heel of the Internet," an article published in the journal Nature on July 27, 2000.

Fig. 14 The Achilles' heel

Fig. 15 The Achilles' heel of the Internet

All the edges connecting one node in a network are erased when one node is removed at a time, and if this is repeated, the network will eventually collapse and become unconnected (Fig. 16). If the network continues to operate correctly even after some nodes are removed, it is considered to be resistant to the removal of nodes. Both networks are judged to have different levels of resilience after a given number of nodes are removed from each of them.

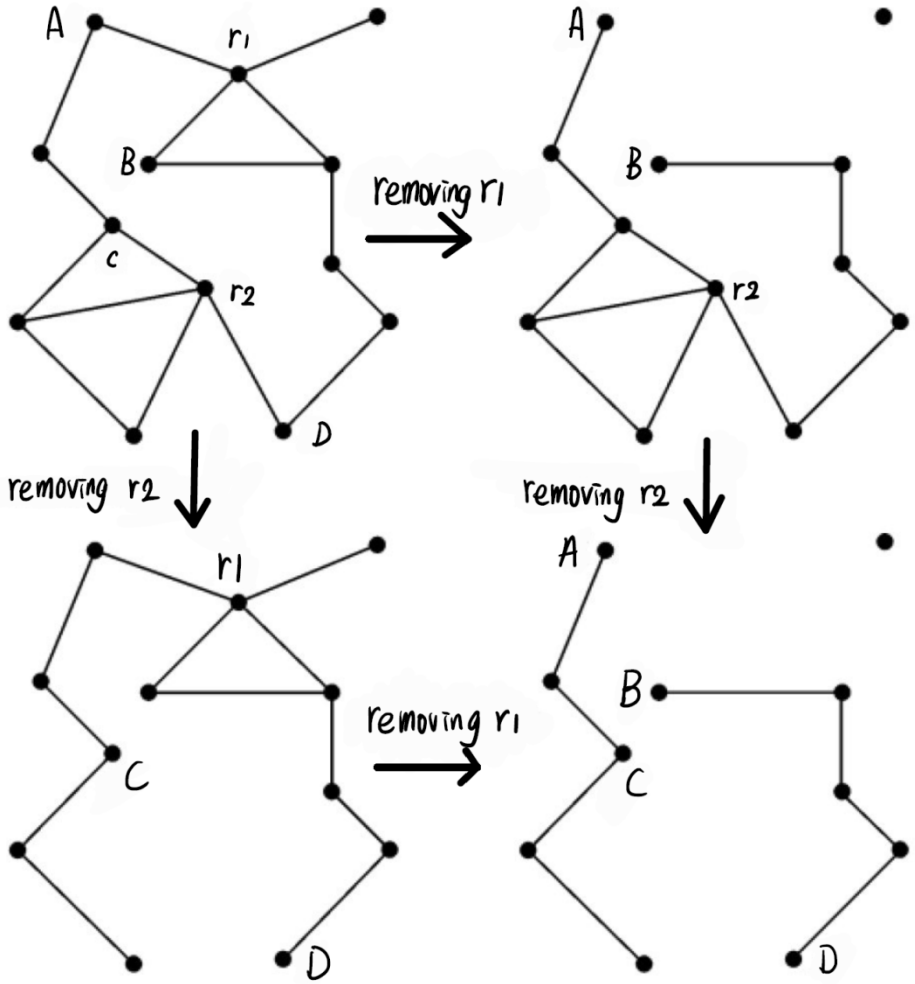

Fig. 16 A connected network becomes unconnected due to removal of nodes

Typical Attacks Method

Following malfunctions and/or attacks on any of these two types of networks, we explore the efficacy of rewiring solutions for both types of networks. When we talk about failures, we're referring to the random elimination of a network node. It represents real-world occurrences such as an accident or shutdown of a peer-to-peer network or in the field of telecommunications. An assault, on the other hand, corresponds to the removal of a node with a higher degree, and it reflects, for instance, a purposeful attack across a network in which the attacker aims to do as much damage as possible. An assault also corresponds to the removal of a node with a higher degree. Discovering such nodes is made much simpler by several stochastic approaches, which eliminate the need of having prior knowledge of the topology of the whole network. In this study, we modified the heuristics suggested by (Cohen et al.,

2003), in which the attacker chooses a random node from the entire network and then a random neighbor of this node in order to target nodes with a higher degree. The original heuristics suggested that the attacker should select a random node from the entire network and then a random neighbor of this node. Assume that the node that was attacked most recently has a probability of kp_k of having a degree of k, where p_k represents the degree distribution of the network. As a direct result of this, higher-degree nodes have a greater risk of being attacked.

There are six attack method widely used, details are showed below in table 1.

		Node-removal	Edge-removal
Targeted	Betweenness	Remove the node with the largest betweenness	Remove the edge with the largest betweenness
	Degree	Remove the node with the largest out-degree	Remove the edge with the largest edge degree
Random		Remove a node randomly	Remove an edge randomly

Table 1 The 6 attack methods

Local Triangle Structure

People's social life benefit tremendously from the usage of social networks, which include the use of online social platforms and communications [1, 2]. Important nodes, also known as sensitive nodes, play a key role in the process of information diffusion. The identification of critical nodes is a significant problem on account of the widespread use of these nodes across a variety of industries. The propagation of information and thoughts, for instance, may be sped up by influential people operating inside networks [3].

Over the course of the last several years, a variety of methodologies for assessing significant nodes have been proposed. Numerous research have provided topological location-based techniques, such as degree centrality (DC) [4], betweenness centrality (BC) [5], closeness centrality [6], eigenvector centrality [7], and structure hole [8]. One example of this is degree centrality (DC). The degree centrality of a node is considered to be one of the most basic topological properties and has a low computational cost. It is generally accepted that a node with a higher degree will have a much bigger influence than a node with a lower degree. On the other hand, the weak link between nodes indicated that a node's high degree of centrality does not inherently suggest that it is important. For example, Granovetter [9] conducted research on the topic of how individuals obtain jobs and found that information regarding positions that lead to employment was much more likely to originate from acquaintances with weak ties than that from closer friends. This was the case even though the information could

have been obtained from either source. The structural hole method, which is analogous to the robustness of weak ties, is used in order to search for the bridge joints, which are defined as streams that are spanned by a group of nodes that are connected indirectly to one another. When looking at the big picture, the betweenness and proximity centralities of a node are what ultimately decide its impact. Both of these approaches, however, lost their usefulness in large-scale networks as a consequence of the immense computing complexity needed in calculating the shortest paths between each pair of nodes in the network.

As a consequence, a plethora of alternative centrality indices have been devised to compensate for those measures' inadequacies. The most significant people, according to Kitsak and his colleagues [10], were found in the center of the networks with the greatest K-shell value. As a consequence, Kitsak's K-shell technique was employed to find the major nodes that were located in key locations. Nodes with the same K-shell values, on the other hand, often have different spreading effects. The researchers either adjusted the K-shell decomposition technique or obtained more neighbor information to achieve a better degree of accuracy. To address the drawbacks of the K-shell, the S-shell decomposition methodology [11], the mixed degree decomposition method [12], and the method-based gravity formula [13] have all been improved. By storing many stages of information about neighbors, the semilocal centrality (LC) [14], the neighborhood coreness centrality [15], the cluster ranking approach [16], and the neighborhood centrality method [17] were proposed to boost the accuracy of the assessment. Another aspect that is taken into account while estimating the spreading ability is route variety [18].

K-shell is an extension of H-index centrality that determines the significance of nodes based on the concept of H-index [19]. According to this concept, the H-index of a node is the greatest value H such that it has at least H neighbors of degree no less than H. K-shell uses this concept to determine the significance of nodes. The local structural centrality (LSC) [20] is a technique that evaluates the significance of nodes based on their first and second neighbors as well as their cluster coefficient. This approach was influenced by the LC method.

It is vital to acquire the natures of a node as often as feasible in order to construct a more accurate meter for evaluating the significance of the node. When determining their worth, we take into account not just the degrees of the nodes but also the nearby triangular structures. The recommended approach, in addition to the other seven indices, has been evaluated using data taken from the actual world as well as data generated by artificial networks. The results of these evaluations show that the suggested method in this study is accurate and effective. It has been shown that the connections of a node's neighbors, in addition to the degree centrality of the node itself, are acceptable factors to include when calculating the effects of nodes.

Methodology

We looked at how well the undirected networks $G = (V, E)$ maintained its connectedness after a certain number, p, of the network's nodes were taken out in accordance with the assault method. For the sake of this research, we used the assumption that attackers only had access to network $G' = (V', E')$, which is a network with undetermined defects. Additionally, we presumed that perfect knowledge on network G was not accessible. We determined which nodes of network G' were vulnerable to assault by using the various attack tactics. After that, we looked at how well the real network G connected despite having the identical nodes taken out.

In order to create network G, we used two distinct kinds of artificial networks. We made use of a concept called the Geometric Random Graph (GRG) in order to generate artificial networks. The GRG approach produces random networks in which the connections between nodes are determined depending on the geographic coordinates of both the nodes in the network. We were able to build 100 networks, each with a total of 10,000 nodes, by making use of the GRG models. The radius for linking nodes in the GRG model was set at 0.015 in order to produce linked networks while keeping the average degree within a reasonable distance from those of other networks. Table 3 displays the properties of the GRC network, and Figure 1 illustrates the degree distributions of each network. The values that are shown in Table 1 for the GRG networks are the averages of the values for each of the 100 networks that were produced. According to Figure 1, the GRG network has a degree distribution that is similar to a normal distribution.

	Geometric Random Graph (GRG)
No. of nodes	10,000
No. of links	34,874.3
Average degree	6.975
Clustering coefficient	0.591
Average shortest path length	51.671

Table 2 Statistics of target networks

Project Description

The project title is that Strengthening a random-graph network against destructive attacks.

Within the context of this simulation-based research, the student will investigate three common types of complex network models. These models include the random-graph network, the small-world network, and the scale-free network. Using such information, in particular random-graph networks, the following research issue is investigated: Assume that the designer of the random-graph network has the ability to strengthen the network's resilience against random and targeted attacks to remove nodes and edges by adding a certain percentage (for example, 10 percent of the total size of the network) of new edges to the network. This project will simulate and verify the robustness of the strategy of randomly adding new edges to form as many triangles as possible in the original network. It will then compare this strategy to the strategy of adding new edges to form as many as possible triangles in the original network, in order to verify that local triangle structure is advantageous in strengthening the robustness of complex networks.

I would like to separate the project into two parts. The first part is to generate a Erdős–Rényi random graph by using the source file. I would like to have only two input, which are the no. of notes and the probability. And the desired output will be the graph as given in the source code file. In addition, it would be perfect that if the program can output the clustering coefficient also.

The next part of the project is to start to generate the node removal and edge removal attacks. So, I would like you to modify the last part of the project. One is to form the local triangle structure and the other one is randomly adding new edges (in my understanding I believe it means it is the same as the old program.). And the desired output is some graphs that verify the network robustness.

Project Design and Implementation

To start the project, the first thing I have done is to do some research about the related topic. For example, about the network theory, random graph network, and some others typical network models like small world network and scale free network. After that is the research about how to generate some attacks, some typical attack method, and then is about the local triangle structure to verify the network robustness.

Implementation

Random Graph Model

When comparing the attributes of real-world networks, a random graph model is often used. It is produced by haphazardly putting edges between a group of n nodes to form a network. Erds and Rényi (1959) presented the very first model of a random graph, which they called G(n, p). This model has n nodes (which are equal to |E|), and the chance that each possible edge would appear in the network is given by the parameter p $\in$ (0, 1).

The Erds-Rényi model may be used to extract certain properties. For example, as the network grows larger, the average nodal degree remains constant as $\bar{d} = \frac{p(n-1)}{2}$, as well as the graph seems to have a Poisson degree distribution: the likelihood that a node has the degree d_i is $pd_i \approx \frac{\bar{d}^{di} e^{-\bar{d}}}{d_i!}$; and the expected number of edges is $\frac{n(n-1)}{2}p$; apart from that, since the probability that two nodes form an edge is p, which is independent as to whether those who share a common neighbor or not, the average clustering coefficient *C(G)* increases as the network grows larger and slowly approaches 0.

We may deduce that the *probability p* of forming a connection in between two nodes inside a large-scale network seems to be equivalent to a network density $p = \frac{2\bar{d}}{n-1} = \frac{2\frac{|E|}{|V|}}{|V|-1} = \frac{2|E|}{|V|(|V|-1)} = \Delta$ since $\bar{d} = \frac{|V|}{|E|}$.

Random Graph Model with clustering

If that is a small network with hundreds or thousands of nodes and thousands or tens of thousands of edges, we can make a comparison other network properties at any and all levels between both the random graph as well as the real network to determine how non-random the real network is. Random graph models are presumed to replicate several characteristics of real networks, including the number of nodes and the number of edges. This allows us to determine the degree to which the real network is not random.

However, because the random graph models maintain the same number of nodes and edges as the actual network, they are only able to completely replicate the character trait of skewed degree distribution. This means that the majority of the nodes in the network have low degrees, while a select few, also known as "hubs," have high degrees. The scenario is quite different when the scale of the network develops to be as huge as it is on Facebook, Twitter, or even in a phone network. In a previous section, we mentioned that as the size of the network increases further and further, the average clustering coefficient inside a random graph moves closer and closer to zero. This indicates that in a large-scale simple random graph, the existence of triangles or higher-order local structures is extremely unlikely due to random chance. Consequently, it is pointless to compare a local structure beyond the phrase level because there is no such thing in a random network on a large scale with a zero clustering coefficient. This is why it is nonsensical to evaluate the local structure. In addition to this, the geodesic distance between the nodes inside the actual network is restricted. On the other hand, the edges of a large-scale simple random graph are dispersed according to random chance, as well as the geodesic distance will indeed be infinite.

Because of this, it is necessary for us to create the random network when clustering the data. Not only do we begin to repair the average clustering coefficient or even the global clustering coefficient just at triadic level, but we also limit the number of nodes just at nodal level as well as the number of edges at the phrase level. This is done at the nodal level. We are able to simulate, just as the actual network does, the properties of having a non-zero clustering coefficient as well as a restricted geodesic distance by using this method. Additionally, the traits of having a non-zero clustering coefficient plus having a restricted geodesic distance are connected with other significant aspects of the network, like the component sizes, the presence of a gigantic component and its size, and the percolation properties. Additionally, we will be able to investigate the resilience of the network, as well as percolation , cascading failure, and diffusion process, as well as the influence of network architecture on dynamical systems, as a result of this. There are at least four different methods that have been developed in recent years to construct random graphs with clustering

Guo and Karines Algorithm

In 2009, Guo and Kraines developed a model that is one of the ones that are used to produce random graphs with clustering. The algorithm is broken down into three distinct stages. In the first step of this process, the power-law degree distribution is used to the construction of a random graph *G* that consists of a collection of *nodes V* and a set of *edges E*. Second, a random selection is used to choose five nodes from the random network, and then two edges that connect those nodes are partially rewired to add another triangle. Third, the procedure will be continued when either the average clustering coefficient of just the rewired graph is larger than or equal towards the target average clustering *coefficient C(G*) or it achieves a

certain predefined number of trials. If neither of these conditions is met, the process will continue until it has reached the maximum number of possible outcomes.

The algorithm states that the five notes should meet the following criteria:

1. x and w are altering of v;
2. y and z are not altering of v;
3. e_{wy} and e_{xz} do exist.
4. e_{wx} and e_{yz} do not exist.

Every one of the rewiring operations begins with a random network with a value of V and E and a clustering coefficient of 0. When two random pairs of edges are rewired, the average clustering coefficient $C(G)$ would be changed each time. This process will continue until it reaches 0.24, which is the value of the average clustering coefficient $C(G)$ in the actual network. And in order to bring the average clustering coefficient $C(G)$ up to date, we need to refresh both the numerator $\Delta(i)$ which is the number of triangles each node i is involved in, as well as the denominator $\tau(i)$ which is the variety of available triangles each node i could be involved in. Only then will we be able to bring the average clustering coefficient $C(G)$ up to date. The average clustering coefficient, denoted by the symbol $C(G)$, is defined as the mean of the sum of the means, as was discussed before. It is feasible for a small network with hundreds or even thousands of nodes (for example, Guo and Kraines generated a test network with 1000 nodes and 3000 edges), however it is not feasible for just a large-scale network of vast numbers of nodes and huge number of edges because it will take an unacceptably lengthy time to update the triangle list as well as 2-path list millions of times. This is because updating the triangle list as well as 2-path list millions of times will require a lot of memory.

Bansal et al. Algorithm

It is possible to think of the method developed by Bansal et al. as an improved version of the one developed by Guo and Kraines. When determining the threshold for rewiring, Bansal et al. used both the transitivity measure $T(G)$ and the average clustering coefficient $C(G)$, but Guo and Kraines simply adopted the average clustering coefficient $C(G)$. In addition to this, they come up with a way to save time when computing by adding two triangles together at the same time.

The researchers Bansal et al. test their approach by contrasting a number of clustered random networks only with actual networks that correspond to them; the highest network size among them is 4,713. If, on the other hand, we would like to create a clustered random network in a decent length of time for a large-scale network with 6.7 million nodes and 15.9 million edges, we should throw out the indicator of average clustering coefficient C(G) because of the same rationale when we discuss the algorithm of Guo and Kraines and instead just try to fit the

transitivity measure T(G), which is the same as fitting the expected number of triangles (G). This is necessary in order to generate We will be able to rewire the edges on bulk using this method, which will save a significant amount of time.

Newman Algorithm

Newman (2009) developed a third approach to build a random network with a clustering coefficient that is not zero. It is possible to comprehend Newman's method as the process of obtaining a clustering random graph *G(V, E)* from a configuration model *G(V, S, T)*. This model requires an input of a triangular vector t_i as well as a single edge vector s_i with each node *i* in the actual network. Because the nodal degree of a node *i* can be written as $d_i = 2t_i + s_i$, its generating function again for joint distribution of triangles with single edges is provided by $g(x,y) = \sum_{s,t} p_{st} x^s y^t$, where p_{st} is the probability how a node is involved in *s* single edges and *t* triangles.

The issue with Newman's approach is that it makes excessive use of the edges in order to generate the very same amount of triangles as the actual network does. This is the difficulty. In the first two methods, the amounts of edges |*E*| incidental to a node *i* and 2-paths $\tau(G)$ are fixed in the rewiring procedures. Furthermore, each time we rewire two edges to create a triangle, we are aware that this represents one step closer to the predicted global clustering coefficient C_{Δ}. However, in Newman's method, our attention is solely focused on the addition of triangles. As a consequence, we lose influence on the number of edges |*E*| incidental to the a node I as well as the number of 2-paths $\tau(G)$, and the triangles end up being dispersed throughout the whole network space. After the addition of the 742,026 edges that connect the 5,597,411 triangles, we are left with just 91,406,099 2-paths. Additionally, the global clustering coefficient C rises to 0.18, as well as the average clustering coefficient *C(G)* rises to 0.44.

Gleeson Algorithm

Gleeson (2009) goes on to argue that now in Newman's algorithm, even though a node I does have nodal degree d_i, then that could be a member of up to $d_i/2$ disjoint triangles, and the upper bound of the local clustering coefficient is $C(i) = \frac{\frac{d_i}{2}}{\binom{d_i}{2}} = \frac{1}{d_i - 1}$, and as a result, the constraint impedes its fit mostly to real systems in the world. Gleeson's argument is based on the fact that Newman's algorithm was developed by Newman.

In the alternative, Gleeson (2009) suggests using -theory networks to produce random graphs with non-zero clustering coefficients. He generalizes Newman's model by starting from another higher-order motif, called a *k-clique*. This is a complete graph among *k* nodes, each

of which is connected to every other node in the graph. Rather than joining triangles and single edges incident to each node *i* he uses this motif to generalize Newman's model.

To connect all of the *k-cliques* into a single entity, Gleeson (2009) makes use of external connections, which are analogous to Newman's single edge and represent the edges that are not engaged in any cliques. For instance, if the average degree of a real-world network is between 3 and 4, it is possible to build a clustered random graph by combining some 3-cliques (triangles) and some 4-cliques, while leaving the other nodes as individuals (i.e., 1-cliques).

Simulations

In conclude of the above four algorithms, I chose to generate my own random graph with a similar version.

According to statistical analysis, the number of edges that should be present in the random graph network is expected to be $E(M) = P_{RG}\frac{N(N-1)}{2}$. A uniformly random process of adding or deleting edges may be carried out in order to precisely manage the number of edges M. When an edge is being added, the direction of the edge can be randomized.

As a sort of assault, node removal as well as edge removal may be carried out. In this instance, the edge-degree is determined by computing the geometric mean of such degrees held by the two nodes that are located at the end of the graph [19]. Therefore, the degree of an edge A_{ij} is denoted by $\sqrt{k_i \cdot k_j}$, where k_i^{in} represents the degree of the node *i* that is connected to the edge, and k_j^{out} represents the degree of the node *j* that is connected to the edge.

After the attack, rewiring strategies is used to strengthen the robustness of the random graph network.

When a node in a network is attacked, its neighbors, which we refer to as impacted nodes, take action to maintain the network's overall connectivity by acting similarly to the attacked node. In this section, we will examine the following two alternative strategies:

A) the random rewiring, in which affected nodes randomly connect to other nodes within the network;

B) the formation of the local triangle structure.

The first method, denoted by the letter A, is the simplest one: the afflicted nodes simply reconnect themselves to a network node that is chosen at random. It is not usually relevant in a real network since doing so would need each node to have knowledge of the whole network

in order to choose a newer node to connect to. This is not the case. The primary objective of this approach is to generate a null hypothesis, which can then be contrasted with other, more plausible hypotheses. To be more explicit, we will utilize simulations to determine if the structure of the network remains unchanged or is disrupted, and if the latter occurs, how it occurs. During the course of the attacks, we are also going to determine whether or not the network still has a gigantic component, and we are going to monitor how large it grows.

For the formation of the local triangle structure, the node's capacity for spreading relies strongly on the degree as well as the structures of its neighbors. The greater the number of triangular structures that are generated between a node and its immediate neighbors and the node itself, the greater the likelihood that the node is located in a dense portion of the network. In networks, the relationship between the number of triangles and the clustering coefficients is not a straightforward one [22]. A higher density of triangles doesn't really automatically mean that the clustering coefficients are high. Therefore, in order to avoid the very same flaw as that of the meaning of the clustering coefficient as well as some indicators, the percentage of triangles rather than the number of triangles is considered an indicator to measure how closely the nodes neighbors are connected to each other. This is done instead of using the number of triangles.

We are employing many of the rewiring procedures, and in order to investigate the topology of the network after it has been subjected to a large number of assaults or failures, we are maintaining the same number of nodes throughout the research. Therefore, as immediately as the network suffers the loss of a node, we immediately replace it with a new one. The structure that has been evaluated will determine how the newly added node is connected to the rest of the network. When a new node is added to a random graph, it will be linked to another node that has been picked at random. A method like this will prevent the graph from collapsing after certain assaults since all of the nodes will have been removed.

So now, lets move on to the result.

Results

After generating the random graph serious measurement had been done to measure the robustness of the graph. In order to compare, two strategies are used, which are the random rewiring strategy and the local triangle structure. Also, six attack methods are introduced. The algorithm to form the local triangle structure will also be introduced.

It has been established that the resilience of controllability [12] and network stability [15, 16] both benefit from the use of the local triangle topology. While this is going on, the triangular structure is becoming more noticeable in daily life. As a direct consequence of this, a directed random triangular network has been constructed in this context for the goal of evaluating the controllability and resilience of networks:

1. Start with N - *3* independent nodes, with the remaining 3 nodes connected in a directed ring.
2. Pick two nodes at random, i and j, with really no edge A_{ij} or A_{ji}. Then, at random, choose a node k from every one of node j's neighbors. If an edge A_{jk} exists, add two edges A_{ij} as well as A_{ki}; if an edge A_{kj} exists, add two edges A_{ji} as well as A_{ik}.
3. Step 2) should be repeated till M edges have indeed been added.

CONTROLLABILITY ROBUSTNESS COMPARISON

Under six distinct attacks, the resilience both for structural controllability [17] as well as precise controllability [18] is investigated. The number of external controllers (also known as driver nodes) required to maintain network controllability after the network has been attacked is measured by the density of control-nodes n_D, defined by $n_D = \frac{N_D}{N}$; in which N_D is indeed the number of external controllers (also known as driver nodes) required to maintain network controllability after the network has been attacked, and N is the network size. Note that although N doesn't really change when in an edge-removal attack, it does decrease after a node is removed. The lower the n_D, the more stable the network controllability will be, according to this metric.

The number of elements within maximum matching E determines the number of driver nodes ND for structural controllability: $N_D = max\{1, N - |E*|\}$; where $|E^*|$ is indeed the cardinal number of elements within maximum matching E^*. In terms of state controllability, is if network's adjacency matrix A is sparse, the amount of driver nodes N_D may be estimated using the formula [18]: $N_D = max\{1, N - rank(A)\}$.

The value of n_D is computed according to the first equation but also recorded after every node or edge is removed to measure the robustness including both state as well as structural controllability.

Robustness Measure

A measurable measurement of the controllability and resilience of the network in the face of various attacks is offered. It is a comparison measure that offers a rating of the various networks after taking into consideration the removal of every node or edge while the operation is being performed.

As the node-removal attack is carried out, the density among driver nodes, denoted by n_D, is presented as a function of the proportion of nodes that have been removed, denoted by p_N. The controllability of both networks is assessed ordinarily at each point of p_N due to the fact that a lower density of driver nodes n_D signifies a greater level of network controllability. After that, one calculates the overall average rank, which is denoted by the letter μ and serves as a quantitative and comparative evaluation of the network's controllability and resilience. This quantitative measure may easily be developed to include scenarios in which there are more than two networks, as well as attacks that involve the removal of network edges.

In addition to the average rank, the number n_W of attaining first position is also taken into consideration. This reveals the instances in which the network has the best controllability while being subjected to an attack.

Simulation Results

The outcomes of a targeted assault are unique, but the results of a random strike are representative of an overall pattern. The outcomes of the simulations involving the removal of nodes are shown in Figures 17-22.

The outcomes of the random node removal are shown in Figures 17 and 18. It is quite evident that LTS performs better than RWS.

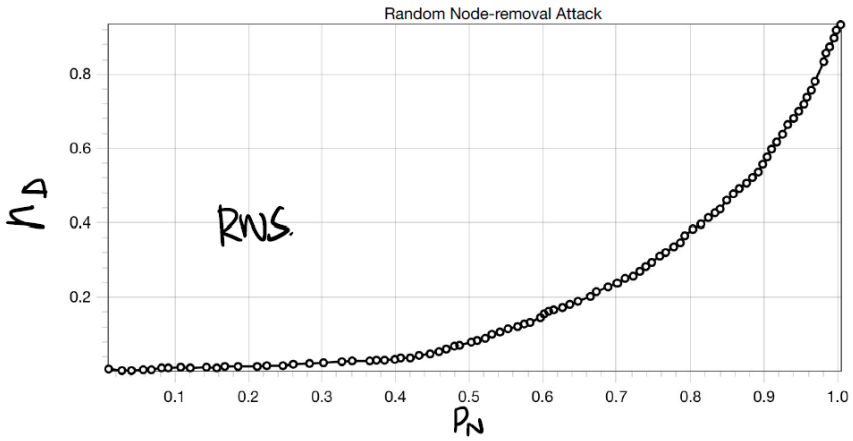

Fig. 17 RWS Random Node-removal attack

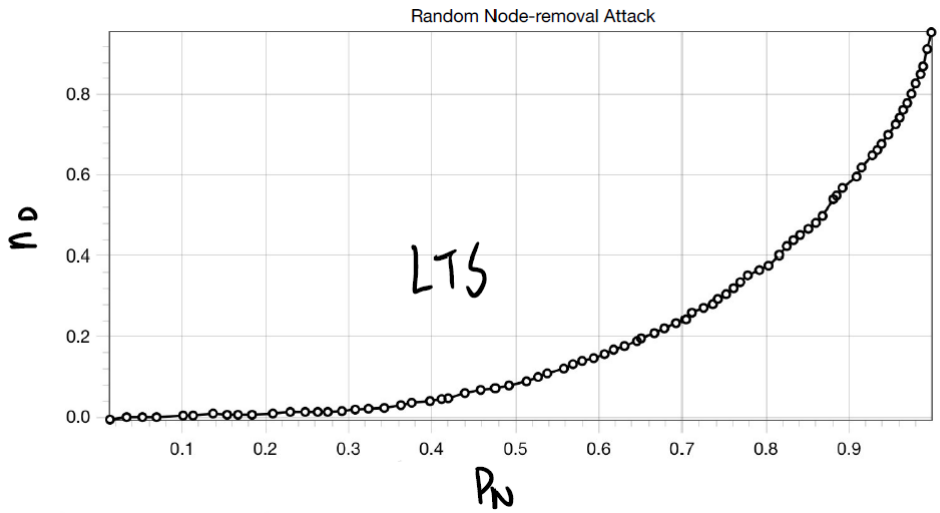

Fig. 18 LTS Random Node-removal attack

The outcomes of betweenness-based targeted node-removal are shown in Figures 19 and 20 respectively. When seen side by side, the performances are tough to compare. It is possible to deduce that LTS maintains a stable performance during the whole process of attacking.

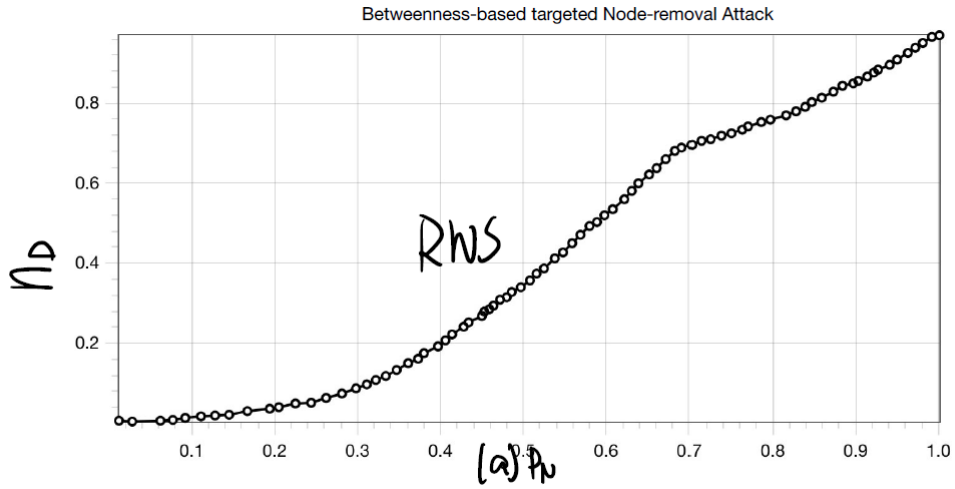

Fig. 19 RWS Betweenness-based targeted node-removal attack

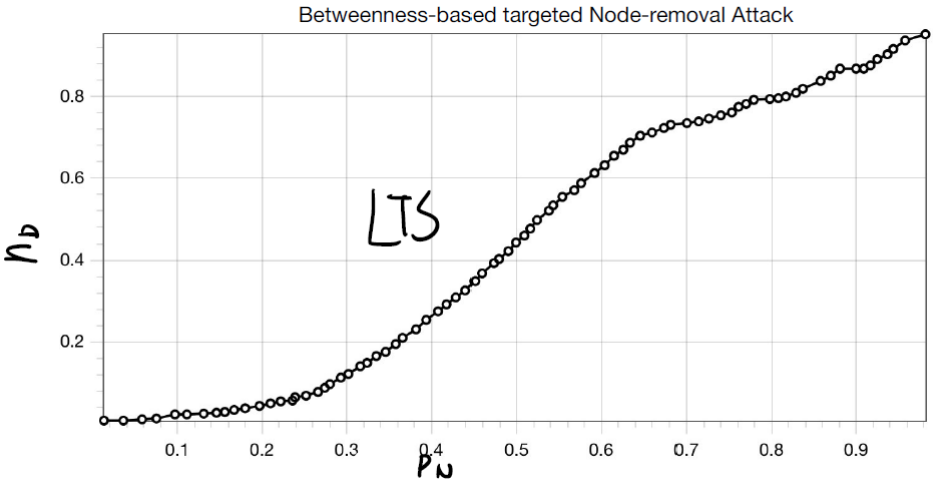

Fig. 20 LTS Betweenness-based targeted node-removal attack

The outcomes of degree-based targeted node removal are shown in Figures 21 and 22 below. Within the context of this scenario, RWS demonstrates the highest level of performance overall. Notable is the fact that LTS is not much worse.

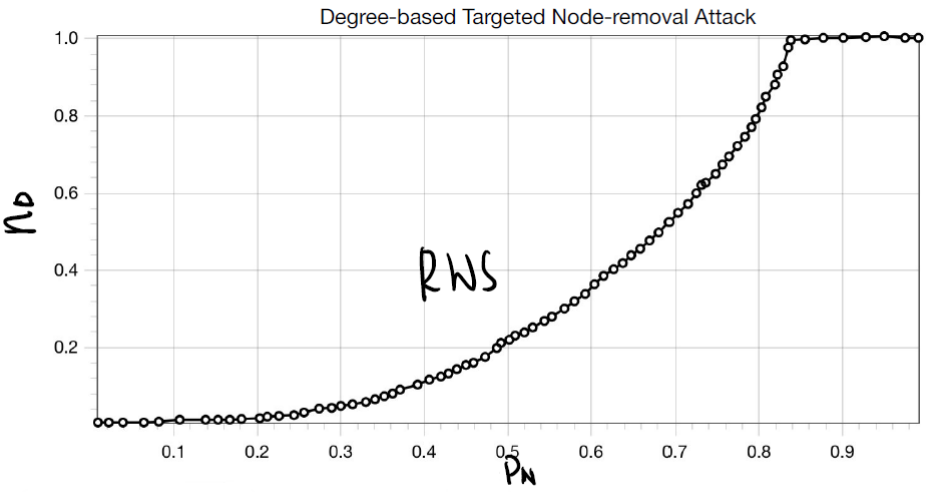

Fig. 21 RWS Degree-based targeted node removal attack

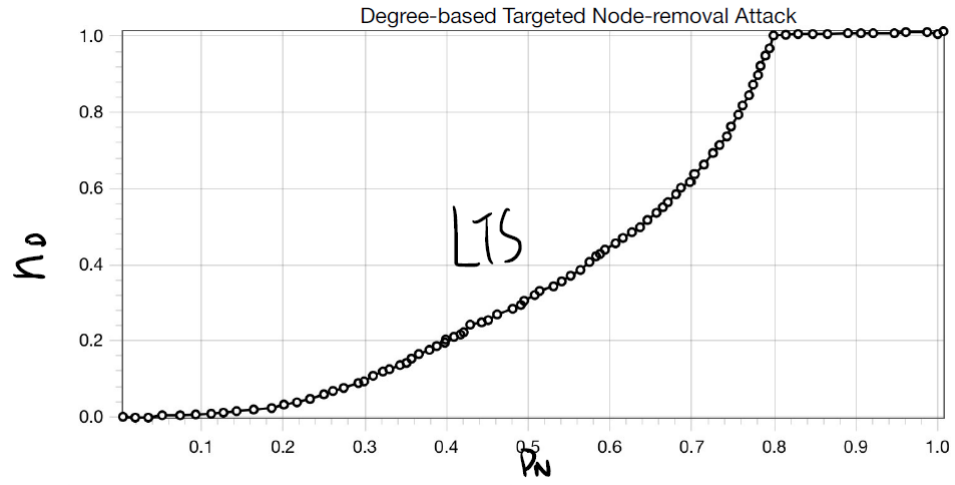

Fig. 22 LTS Degree-based targeted node removal attack

Figures 23–28 show the results of a simulation run that examined edge-removal attacks.

The outcomes of the random edge-removal process are shown in figures 23 and 24. On this size, the results are quite comparable to those obtained by removing nodes at random, as illustrated in 17 and 28, but on smaller scales, there are several distinguishing characteristics between the two sets of findings. LTS performs better than RWS.

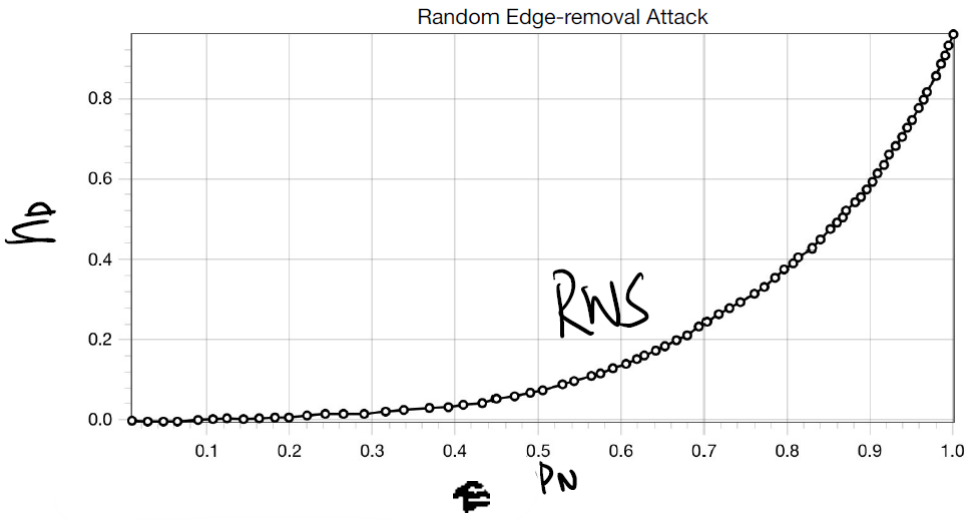

Fig. 23 RWS Random edge-removal attack

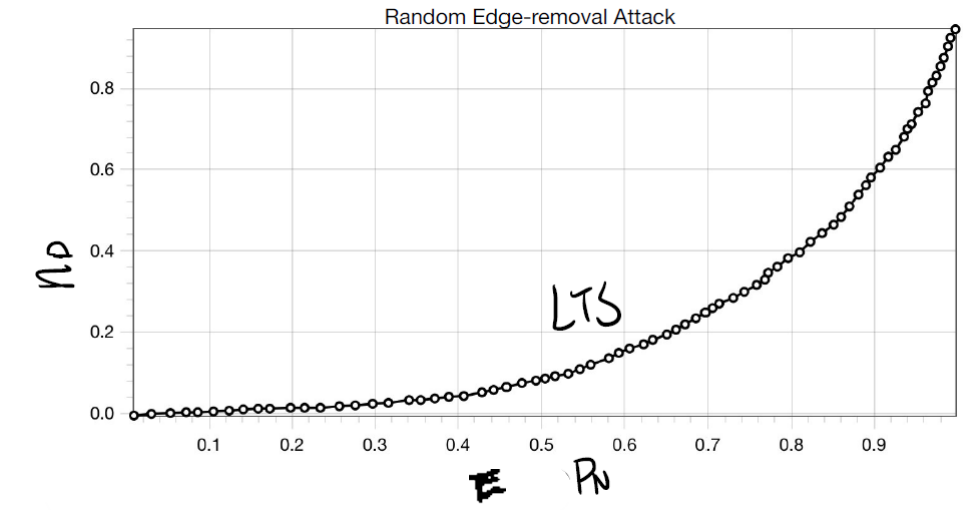

Fig. 24 LTS Random edge-removal attack

The results of betweenness-based targeted edge-removal are shown in Figures 25 and 26, while the results of degree-based targeted edge-removal are displayed in Figures 27 and 28. Additionally, LTS performs better than RWS.

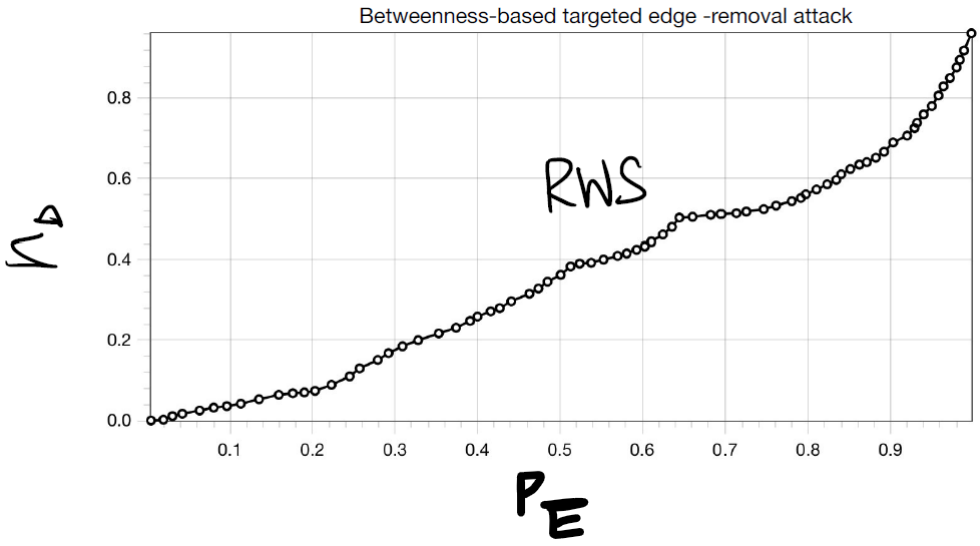

Fig. 25 RWS Betweenness-based targeted edge-removal attack

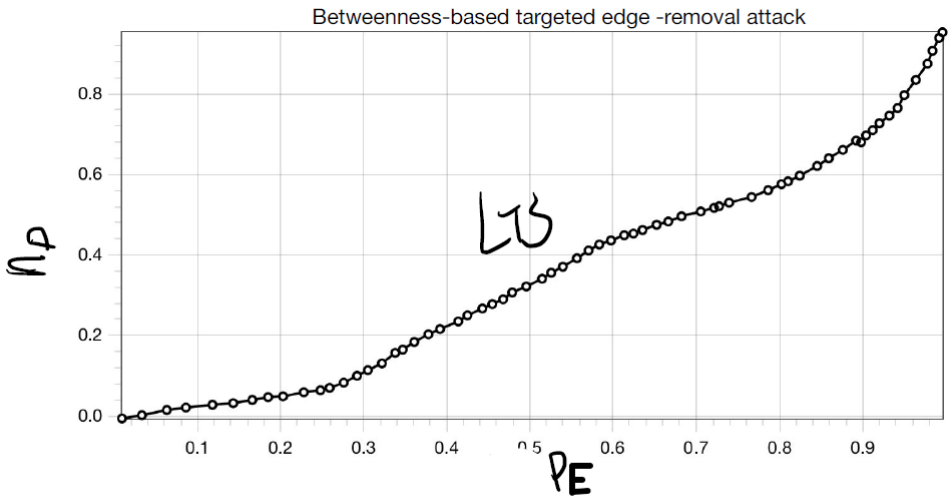

Fig. 26 LTS Betweenness-based targeted edge-removal attack

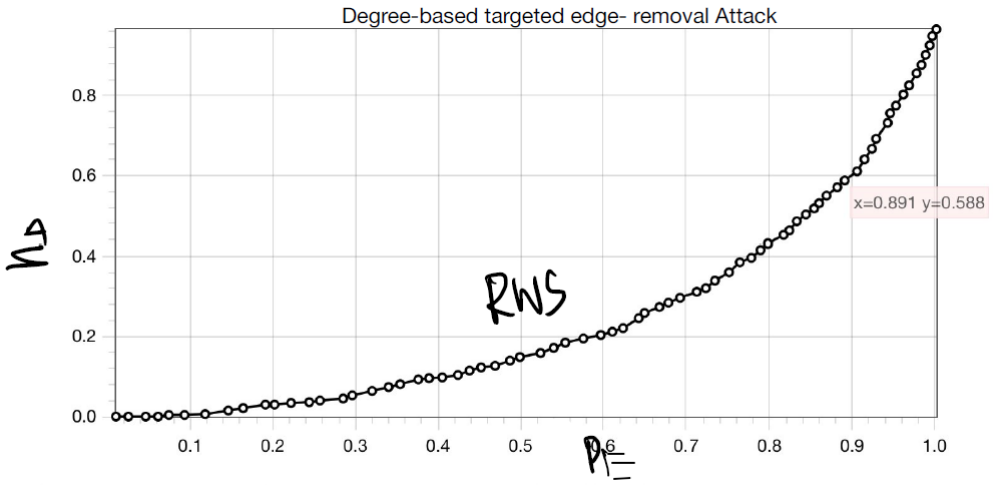

Fig. 27 RWS Degree-based targeted edge-removal attack

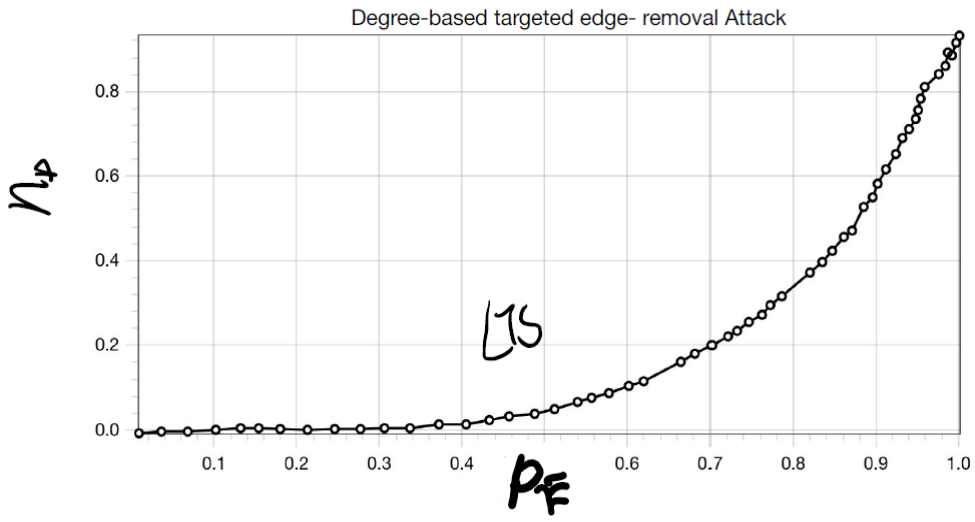

Fig. 28 LTS Degree-based targeted edge-removal attack

Discussion

So, after analyzing all the result, we can conclude that local triangle structure is advantage in strengthening complex network robustness.

In this research, we came to the conclusion that random graphs are relatively resistant to both failures and assaults, which is consistent with what is now a well accepted finding. Furthermore, some unexpected consequences have been seen when the damaged networks do not stay inert and instead react to the suppression of nodes. On the basis of this, we developed two different strategies: A) rewiring at random, and B) the local triangle structure.

The controllability and resilience of random-graph networks against multiple types of malicious assaults, such as node-removal as well as edge-removal-based random or targeted attacks, has been studied in depth in this research. Some examples of these types of attacks include node removal as well as edge removal. It has been shown that the random-graph network performs consistently well against assaults aimed at removing nodes, while it does badly against attacks aimed at removing edges. This reveals that the local triangular structure is a significant component of networks that contributes significantly to the controllability and durability of the network.

In addition to this, we investigated the impact that connection failures and network sampling have on the effectiveness of various network attack strategies. When we did comprehensive simulations, we utilized real data from the networks as well as theories of how networks are created. We presented evidence that network attack strategies are resistant to link failures even in systems with highly skewed degree distributions. In addition, we have shown that the effectiveness of network attack strategies is greatly diminished when the network is generated by the use of random sampling. In spite of this, network sampling may nevertheless have significant repercussions, even when the sample size is quite small (for example, when it is less than 30 percent).

Conclusion

In recent years, developments in complex networks have uncovered some astonishing parallels between seemingly unrelated systems, such as the World Wide Web, cellular brain networks, metabolic systems, and even the society of Hollywood movie stars. These discoveries have been made possible by the fact that there are many similarities between these systems. In particular, there has been a great deal of research done regarding the consequences of network topology upon the dynamical behaviors of networks. However, this has been considered as only the tip of a giant iceberg, since there are enormous problems and technical obstacles that still need to be overcome in the fields of modeling, analysis, and control, in particular the synchronization of large dynamical networks.

We are in the process of gradually integrating and linking networks that serve a variety of functions, including information, energy, transportation, commerce, and others. Because of the critical nature of these networks, concerns have been raised about the potential for harm and the repercussions of system failures. As a result, there is a pressing need for a deeper understanding of the fundamental concepts that underlie these systems. Because of this, a greater investment is required in the creation and operation of all different kinds of large and complex dynamic system networks in order to improve one's ability to assess and predict potential future challenges. The end goal is to make the most of the available net potentials so that our human civilization may profit as much as possible from them. In order to acquire this level of understanding, in-depth study is necessary. This research must offer a solid scientific foundation for further investigations into real-world complex dynamical networks, as well as creative methods for the construction and use of these networks. Innovative strategies and procedures are required in order to solve the severe problems that we encounter on a daily basis, which range from the biological sciences to power grids to network technology. At the beginning of the twenty-first century, it has developed into a significant issue while simultaneously posing a significant opportunity for scientists and engineers.

Fig. xx The End

References