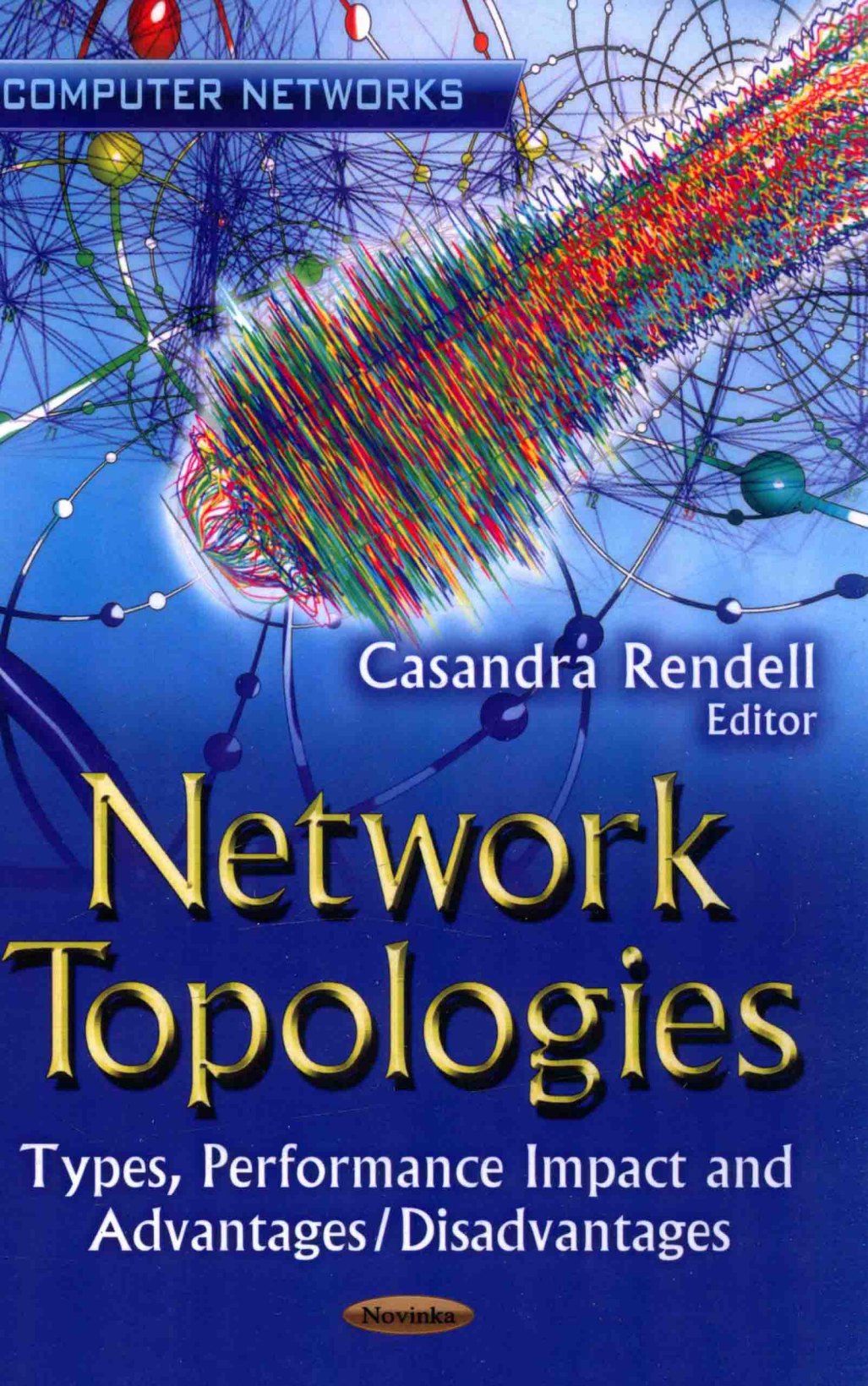
COMPUTER NETWORKS
Casandra Rendell
Editor
Network Topologies
Types, Performance Impact and Advantages / Disadvantages
Novinka

COMPUTER NETWORKS

# NETWORK TOPOLOGIES

## TYPES, PERFORMANCE IMPACT AND ADVANTAGES/DISADVANTAGES



CASANDRA RENDELL
EDITOR


*New York*

**Library of Congress Cataloging-in-Publication Data**

ISBN: 978-1-62618-180-9

*Published by Nova Science Publishers, Inc. † New York*

# NETWORK TOPOLOGIES

## TYPES, PERFORMANCE IMPACT AND ADVANTAGES/DISADVANTAGES

# COMPUTER NETWORKS

Additional books in this series can be found on Nova's website under the Series tab.

Additional e-books in this series can be found on Nova's website under the e-book tab.

# PREFACE

In this book, the authors discuss the types, performance impact and advantages/disadvantages of network topologies. Topics include the analysis of the topological structure of networks and related problems via simple rules; application of artificial neural networks in the diagnosis of coronary heart disease; network topology, signal function pathways and cell function; model-based risk analysis of complex networks; and network analysis of clinical, medical and molecular data in systems science.

Chapter 1 - It has been widely recognized that networks is a good tool for describing various artificial and natural systems. Thus, a natural problem arises: how to construct networks and further properly study its features. At present, it is still lack of basic laws for describing the quantitative relationships in biological networks, social networks, etc. Therefore, the construction method of networks should be different from that used to build up the mathematical models in the past based on the quantitative relationships in various scientific fields, such as physics and mechanics. In this chapter, the authors provide a new method to construct networks via simple rules and further study related problem. The main feature of the method is that simple basic rules for constructing networks are firstly proposed based on the current scientific understanding of the problems discussed, and then the networks is builded up through these rules, and further the rationality of the rules is evaluated by comparing the topological characteristics of the constructed networks with those measured from real networks, related problem is finally analyzed using the obtained reasonable network models. Two examples will be given to illustrate the research approach. One example is about the modeling of protein-protein interaction networks and the other example is about the positioning of wireless sensor networks.

Chapter 2 - The present research is aimed to develop an ANN diagnostic model for the coronary atherosclerosis and ischemia for patients after coronary angiography on the basis of genetic, clinical laboratory and instrumental examination data. The analysis of the correlation between the signs allowed us to choose the factors most closely connected with the diagnosis. Hierarchical clustering and correlation analysis were adapted to allocate typical fields of diagnostic factors. Various types of ANN topologies (MLP, SVM, PCA, and hybrid network) were analyzed; the authors have found that the models based on ANN with principal components analysis, and double-layer perceptron ANN optimized with genetic algorithms achieve the best diagnostic efficacy.

Chapter 3 - The living beings continuously receive, process and respond to a myriad of signals. To this aim several transduction pathways have evolved, giving rise an amazingly complex functional dialogue into and among cells. Modern analytical techniques (such as 2D electrophoresis, DNA microarray, protein chips, etc...) increased dramatically our knowledge about the molecules involved in these signaling events and allowed the discovery of several biochemical determinants of signal transduction. Despite scientific community expectations, the applicative benefit of this always increasing information to date are scarce. In the authors opinion, this is due to the intrinsic complexity of biological beings, from molecules to cells, from tissues to organs, that make necessary the adoption of the computational modeling strategies. In this context, the authors adopted a biological networks-based approach to study the topology of signal transmission pathways. In particular the authors have realized and statistically analyzed the main topological indexes of biological networks representing the molecular events involved in 11 pathways of relevant importance for human biology [smooth and striated muscle contraction, release cycle of six neurotransmitters, visual phototransduction (rods), sperm capacitation, insulin signaling pathway, p53 pathway, regulation of retinoblastoma protein (pRb), mitochondrial ATP metabolism, glucose metabolism, signaling events mediated by stem cell factor receptor c-Kit and the circadian clock]. As a result the authors have found that the number of molecules composing each network is likely due to the balancing of energetic cost of each molecule and the need of system stability. In addition, the authors demonstrated that all the examined networks showed the same scale free topology and small world behavior with a clearly non hierarchical structure, unlike other biological networks. In particular the clustering coefficient (i.e. the measure of how each network node tend to cluster to other ones) was virtually zero and it was unrelated to the number of links per node, the characteristic path length (the measure of how many links it

is necessary to pass through to travel between two random selected nodes) was about 6 and the averaged number of link per node ranged between 2 and 2.8. From these findings, it is possible to conclude that all examined pathways share the same topology that confers them some important biological features such as robustness against random failure, controllability, and specificity and efficiency in signal transmission, which are the specific signature of signal transduction pathways.

Chapter 4 - *Model-based Risk Analysis* (MBRA) is a method and tool for analyzing complex system risk as defined by a network of components and their links to one another. Network risk is defined as the expected consequence across all nodes and links of the network. Network nodes and links are the fundamental assets represented in MBRA. Nodes can be anything – a building, tunnel, computer, water pump, reservoir, power substation, person, group, or telecommunication switching office. Links can be anything: a power line, familial relation, pipe, Internet fiber optic cable, or virtual dependency. Nodes and links have properties: values needed to compute risk as well as network properties such as number of connections, betweeness, and height.

MBRA employs fault trees to model multiple threats arrayed against single or multiple nodes; and expected utility theory (EUT) to compute overall network risk. In addition to calculating EUT risk, MBRA performs various modeling and simulation functions as follows:

1 Network properties: compute degree, betweeness, height, and contagiousness.
2 Objectives: optimize on risk, vulnerability, probability of failure, and consequence.
3 Resource allocation: distribute budgets across nodes and links such that one of the objectives listed in 2 is optimized.
4 Return on Investment Analysis: analyze changes in objectives versus investment in threat, prevention, and response capability.
5 Cascades: simulate cascade failures precipitated by a single failure.
6 Flow: simulate the loss of flows through the network due to a single failure.
7 Displays: calculate and display the exceedence probability generated by simulation of faults in the cascade and flow simulations.
8 Threat analysis: model threats versus nodes and links as a fault tree.
9 Game theory: simulate a 2-person game and find optimal allocations of threat, vulnerability, and response investments using Stackleberg optimization.

10 Exceedence probability analysis and probable maximum loss analysis (an approach favored by the insurance industry).

MBRA has been used since 2006 to analyze dozens of critical infrastructure systems of interest to homeland security. Applications have mainly come from students doing classroom projects. Experience has showed that MBRA can be extremely expressive and capable of analyzing complete systems of all kinds and sizes. Some of MBRA's limitations have become apparent: threat, vulnerability, and consequence decrease as an exponential function of investment – a model that may not represent reality. Additionally, cascade failure probability in some systems, such as a power grid collapse or flood, actually increases as a collapse is taking place. MBRA assumes likelihood of collapse is a constant. Finally, MBRA does not model dynamical systems with feedback loops. It assumes directional or bidirectional flows, but does not model dynamic feedback. MBRA is freely available at www.CHDS.us for Windows and Macintosh operating systems.

Chapter 5 - Cancer is a large class of diseases. Systems level analysis of complex diseases such as cancer requires the analysis of relationships between different types of clinical as well as molecular data. Chemotherapy, in addition to radiotherapy and surgery, is used to treat cancers. However, cancer drugs mostly target general processes, i.e. DNA synthesis, toinhibit cell division. A recent trend has been to use targeted chemotherapy, to more selectively treat specific types of cancer. Systems science is the analysis of the global organization of the relationships between groups, i.e. the different types of diseases and drugs, and is in contrast to traditional approaches of analyzing individual parts of a system. Such analyses have provided an understanding of the combined clinical and molecular aspects of diseases such as cancer. Networks of diseases, drugs, and genes showed unexpected connectivity for particular members, correlations between the connectivity with epidemiological features, etc. Network analysis of clinical, medical and molecular data is a promising branch of systems science.

# CONTENTS

In: Network Topologies
Editor: Casandra Rendell

ISBN: 978-1-62618-180-9


*Chapter 1*

# ANALYZING THE TOPOLOGICAL STRUCTURE OF NETWORKS AND RELATED PROBLEMS VIA SIMPLE RULES

***Shuiming Cai[1], Qinbin He[2] and Zengrong Liu[3,*]***
[1]Faculty of Science, Jiangsu University, China
[2]Department of Mathematics, Taizhou University, China
[3]Institute of Systems Biology, Shanghai University, China

## ABSTRACT

It has been widely recognized that networks is a good tool for describing various artificial and natural systems. Thus, a natural problem arises: how to construct networks and further properly study its features. At present, it is still lack of basic laws for describing the quantitative relationships in biological networks, social networks, etc. Therefore, the construction method of networks should be different from that used to build up the mathematical models in the past based on the quantitative relationships in various scientific fields, such as physics and mechanics. In this chapter, we provide a new method to construct networks via simple rules and further study related problem. The main feature of the method is that simple basic rules for constructing networks are firstly proposed based on the current scientific understanding of the problems discussed, and then the networks is builded up through these rules, and

* Email: zrongliu@126.com.

further the rationality of the rules is evaluated by comparing the topological characteristics of the constructed networks with those measured from real networks, related problem is finally analyzed using the obtained reasonable network models. Two examples will be given to illustrate the research approach. One example is about the modeling of protein-protein interaction networks and the other example is about the positioning of wireless sensor networks.

## I. INTRODUCTION

In recent years, with the rapid development of information technology and science technology as well as the extensive cross-penetration among various disciplines, complex networks have become a hot research field and attracted much attention from the scientific communities. Actually, if using nodes or vertices to represent basic elements with certain dynamical characteristics and information systems, while edges or links to represent the relationship or connection of these basic elements, then a wide range of systems in nature and society can be described by models of complex networks consisting of nodes connected by edges. For example, the cell can be described as a complex network of chemicals connected by chemical reactions; the Internet is a complex network of routers and computers linked by various physical or wireless links; the World Wide Web (WWW) is an enormous virtual network of webpages connected by hyperlinks. Other examples include food webs, social networks, organizational networks, coauthorship and citation networks of scientists, neural networks, cellular and metabolic networks, protein-protein interaction networks, electronic power grids, etc., [1-3]. The ubiquity of complex networks in science and technology has naturally led to a set of common and important research problems. Among them, the most basic issues are how to characterize network anatomy since structure always affects function. For instance, the topology of social networks affects the spread of information and disease, and the topology of the power grid affects the robustness and stability of power transmission [1].

In the early days, the study of complex networks always adopted an implicit assumption that the interaction patterns among the individuals of the networks can be embedded onto a regular and perhaps universal structure such as a Euclidean lattice. However, a number of studies have suggested that most real-life networks have no apparent design principles, and so they cannot be described as regular graphs. In late 1950s, two Hungarian mathematicians Paul Erdos and Alfred Renyi proposed an Erdos-Renyi (ER) random graph model

to describe a network with complex topology [4]. Their work had laid a foundation of the random network theory, followed by intensive studies in the next 40 years and even today. Random graphs have been proposed as the simplest and most straightforward realization of large-scale networks with no apparent design principles. But our intuition clearly indicates that many real-life complex networks are neither completely regular nor completely random. Diverse complex systems, such as the cell and the Internet, must display some organizing principles which should be at some level encoded in their topology as well [2]. Therefore, tools and measures to capture in quantitative terms the underlying organizing principles need to be developed.

In the past few years, the computerization of data acquisition in all fields and the availability of high computing power have led to the emergence of huge databases on the topology of various real networks. The public access to the huge amount of real data has in turn stimulated great interest in trying to uncover the generic properties of different kinds of complex networks [5]. As a result, many quantities and measures have been proposed to characterize the structural properties of networks in recent years [6, 7]. In this endeavor, two significant discoveries are the small-world effect and the scale-free feature of most real-life complex networks [8, 9]. In addition, many other interesting topological properties of complex networks have also been revealed, such as hierarchical modularity and disassortativity [10-12]. These empirical findings naturally lead to a problem: how to construct networks that can reproduce the known structural properties observed in real network, since the models proposed in mathematical graph theory turned out to be very far from the real needs. Moreover, modeling the structure of a complex network would also lead to a better knowledge of its evolutionary mechanisms and to a better cottoning on its dynamical and functional behavior [7].

Unfortunately, it is usually lack of clear quantitative relationships in the complex network modeling, but only some descriptive terms are given therein. This results in the construction method of networks should be different from that used to build up the mathematical models in the past based on the quantitative relationships in various scientific fields, such as physics and mechanics. Consequently, there exists a problem of how to give some simple rules to construct the network model based on these descriptive terms. The main purpose of this chapter is to provide some introduction and insights into this problem, with emphasis on how to extract some simple rules from these descriptive terms, and then to construct the network model via these simple rules for achieving the goal of modeling. Two examples will be given to illustrate the research approach. One example is about the modeling of protein

interaction networks [13] and the other example is about the positioning of wireless sensor networks [14].

## II. SOME BASIC CONCEPTS

In this section, we shall first introduce some definitions and notations, and then provide a brief review of several important quantities used to describe the topology of a network.

### Definitions and Notations

Graph theory is the natural framework for the exact mathematical treatment of complex networks and, formally, a complex network can be represented as a graph. An *undirected* graph $G$ consists of two sets $\mathrm{N}(G)$ and $\mathrm{E}(G)$, such that $\mathrm{N}(G) \neq \varnothing$ and $\mathrm{E}(G)$ is a set of unordered pairs of elements of $\mathrm{N}(G)$. The elements of $\mathrm{N}(G) = \{n_1, n_2, \cdots, n_N\}$ are the *nodes (or vertices, or points)* of the graph $G$, while the elements of $\mathrm{E}(G) = \{e_1, e_2, \ldots, e_K\}$ are its *edges (or links, or lines)*. The number of elements in $\mathrm{N}(G)$ and $\mathrm{E}(G)$ are denoted by $N$ and $K$, respectively. A node is usually referred to by its order $i$ in the set $\mathrm{N}(G)$. In a undirected graph, each of the edges is defined by a pair of nodes $i$ and $j$, and is denoted as $(i, j)$ or $e_{ij}$. The edge is said to be *incident* in nodes $i$ and $j$, or to join the two nodes; the two nodes $i$ and $j$ are referred to as the *end-nodes* of edges $(i, j)$. Two nodes joined by an edge are referred to as *adjacent* or *neighboring*. The *neighborhood* of a node $i$, henceforth represented as $O(i)$, corresponds to the set of nodes adjacent to node $i$. In the complex network literature, it is often assumed that no *loops*, i.e., edges from a node to itself, or *multiple edges*, i.e., couples of nodes connected by more than one edge, exsit. Graphs with either of these elements are called *multigraphs* [15, 16].

In this chapter, the emphasis will be on undirected graphs rather than multigraphs or directed graphs. For multigraphs and directed graphs, the reader is referred to the references [6, 7] and the books [15, 16].

## Average Degree

The simplest and perhaps also the most important characteristic of a single node is its *degree*. The degree $k_i$ of a node $i$ is usually defined to be the total number of its connections. Thus, the larger the degree, the "more important" the node may be in a network [5]. The *average degree* $< k >$ of a network is the average of $k_i$ for all vertices in the network, that is,

$$< k >= \frac{1}{N} \sum_i k_i.$$

In a sparse network, the average degree $< k >$ usually satisfies $< k > \propto \ln N$ [2].

## Average Shortest Path Length

Shortest paths play an important role in the transport and communication within a network. Suppose one needs to send a data packet from one computer to another through the Internet: the geodesic provides an optimal path way, since one would achieve a fast transfer and save system resources [7]. Therefore, shortest paths have played an important role in the characterization of the internal structure of a network [7]. If one represents all the shortest path lengths of a network as a matrix $D$ in which the entry $d_{ij}$ is the length of the geodesic from node $i$ to node $j$. The *average shortest path length* $< L >$ of the network, then, is defined as the mean of geodesic lengths over all pairs of nodes [8]:

$$< L >= \frac{1}{N(N-1)} \sum_{i \neq j} d_{ij},$$

where $N$ is the total number of nodes in the network. Here, $< L >$ determines the effective "size" of a network, the most typical separation between two nodes therein. In a friendship network, for example, $< L >$ is the average number of friends existing in the shortest chain connecting two persons in the

network [5]. It was an interesting discovery that the average path length of most real complex networks is relatively small, even in those cases where these kinds of networks have many fewer edges than a typical globally coupled network with an equal number of nodes. This smallness inferred the small-world effect, hence the name of small-world networks [5].

## Average Clustering Coefficient

A common property of social networks is that cliques form, representing circles of friends or acquaintances in which every member knows every other member. This inherent tendency to clustering is quantified by the clustering coefficient, defined as the average fraction of pairs of neighbors of a node that are also neighbors of each other [5]. Suppose that a node $i$ in the network has $k_i$ edges, which connect this node to $k_i$ other nodes. These nodes are all neighbors of node $i$. Clearly, at most $k_i(k_i-1)/2$ edges can exist among them, and this occurs when every neighbor of node $i$ connected to every other neighbors of node $i$. The *clustering coefficient* $C_i$ of node $i$ is then defined as the ratio between the number $E_i$ of edges that actually exist among these $k_i$ nodes and the total possible number $k_i(k_i-1)/2$, namely,

$$C_i = \frac{2E_i}{k_i(k_i-1)}.$$

The *average clustering coefficient* $<C>$ of the network is then given by the average of $C_i$ over all the nodes in the network:

$$<C>=\frac{1}{N}\sum_i C_i,$$

which characterizes the overall tendency of nodes to form clusters or groups. By definition, $0 \le C_i \le 1$, $0 \le <C> \le 1$; and $<C>=1$ if and only if the network is globally coupled, which means that every node in the network connects to every other node. In a completely random network consisting of $N$ nodes, $<C> \propto 1/N$, which is very small as compared to most real networks. It has

been found that most large-scale real networks have a tendency toward clustering, in the sense that their clustering coefficients are much greater than $O(1/N)$, although they are still significantly less than one (namely, far away from being globally connected) [5]. This, in turn, means that most real complex networks are not completely random. Therefore they should not be treated as completely random and fully coupled lattices alike.

Given the clustering coefficients of the nodes, the clustering coefficient can be expressed as a function of the degree of the nodes [6]:

$$C(k)=\frac{\sum_i C_i \delta(k_i-k)}{\sum_i \delta(k_i-k)},$$

where $\delta(\cdot)$ denotes Kronecker's delta function, that is, the average clustering coefficient of all nodes with $k$ edges. For many real networks, this function has $C(k)\propto k^{-\upsilon}$. This behavior has been associated with a hierarchical structure of the network, with the exponent $\upsilon$ being called its hierarchical exponent [10, 11].

## Degree Distribution

Not all nodes in a network have the same number of edges. The spread in the number of edges a node has, or node degree, over a network is characterized by a distribution function $P(k)$, which is the probability that a randomly selected node has exactly $k$ edges. A regular lattice has a simple degree sequence because all the nodes have the same number of edges; and so a plot of the degree distribution contains a single sharp spike (delta distribution) [5]. Any randomness in the network will broaden the shape of this peak. In the limiting case of a completely random network, the degree sequence obeys the familiar Poisson distribution; and the shape of the Poisson distribution falls off exponentially, away from the peak value $<k>$. Because of this exponential decline, the probability of finding a node with $k$ edges becomes negligibly small for $k>><k>$. In the past few years, many empirical results showed that for most large-scale real networks the degree distribution deviates significantly from the Poisson distribution. In particular,