

Characterization of Complex Networks

Application to Robustness Analysis

Characterization of Complex Networks

Application to Robustness Analysis

Proefschrift

ter verkrijging van de graad van doctor
aan de Technische Universiteit Delft,
op gezag van de Rector Magnificus Prof.dr.ir. J.T. Fokkema,
voorzitter van het College voor Promoties,
in het openbaar te verdedigen op donderdag 9 oktober 2008 om 10:00 uur

door

Almerima JAMAKOVIĆ

elektrotechnisch ingenieur
geboren te Sarajevo, Bosnië en Herzegovina.

Dit proefschrift is goedgekeurd door de promotor:
Prof.dr.ir. P.F.A. Van Mieghem

Samenstelling promotiecommissie:

Rector Magnificus	Voorzitter
Prof.dr.ir. P.F.A. Van Mieghem	Technische Universiteit Delft, promotor
Prof.dr. D. Hutchison	Lancaster University
Prof.dr. J.C. Arnbak	Technische Universiteit Delft
Prof.dr.ir. I.G.M.M. Niemegeers	Technische Universiteit Delft
Prof.dr.ir. M.P.C. Weijnen	Technische Universiteit Delft
Prof.dr.ir. N.H.G. Baken	Technische Universiteit Delft
Prof.dr.ir. E.R. van Dam	Tilburg University

ISBN 978-90-79787-01-2

Published and distributed by:

Next Generation Infrastructures Foundation
P.O. Box 5015
2600 GA Delft
The Netherlands
T: +31 15 278 2564
F: +31 15 278 2563
E: info@nginfra.nl
I: www.nginfra.nl

This research was funded by the Next Generation Infrastructures Foundation programme.

Keywords: complex networks, network topology, graph and spectral measures, network robustness

Copyright © 2008 by A. Jamaković

All rights reserved. No part of the material protected by this copyright notice may be reproduced or utilized in any form or by any means, electronic or mechanical, including photocopying, recording or by any information storage and retrieval system, without written permission from the author.

Printed in The Netherlands

I call to witness time, the beginning and end of all things - to witness that every man
always suffers loss.

Contents

1	Introduction	1
1.1	Why studying complex networks	1
1.2	Scope of complex networks research	1
1.3	Scope and contribution of thesis	2
1.4	Thesis outline	3
2	Modeling of Complex Networks: The Set of Graphs	5
2.1	Notation	5
2.2	Random graph of Erdős-Rényi	6
2.3	Small-World graph of Watts-Strogatz	6
2.4	Scale-Free graph of Barabási-Albert	7
3	Characterization of Complex Networks: The Set of Graph Measures	9
3.1	Structural measures	9
3.2	Spectral measures	11
4	Analyzing the Relationship Between Graph Measures	15
4.1	Visual analysis	15
4.2	Correlation analysis	19
4.3	Principal component analysis	22
4.4	Conclusions	24
5	Laplacian and Normalized Laplacian Spectrum of Complex Networks	27
5.1	Laplacian spectrum of complex network models	27
5.1.1	Random graph of Erdős-Rényi	28
5.1.2	Small-world graph of Watts-Strogatz	30
5.1.3	Scale-free graph of Havel-Hakimi	31
5.2	Laplacian spectrum of empirical networks	32
5.3	Normalized Laplacian spectrum of empirical networks	34
5.4	Conclusion	34

6	Algebraic Connectivity of Complex Networks	37
6.1	Algebraic connectivity in random graph of Erdős-Rényi	38
6.1.1	Asymptotic behavior in random graph of Erdős-Rényi	38
6.1.2	Minimum nodal degree in random graph of Erdős-Rényi	39
6.1.3	Analytical approximation for algebraic connectivity in random graph of Erdős-Rényi	40
6.1.4	Verification of analytical approximation	41
6.2	Relationship between algebraic, node and link connectivity in random graph of Erdős-Rényi	42
6.3	Conclusion	46
7	Spectral Radius of Complex Networks	47
7.1	Upper bounds for spectral radius	47
7.2	Spectral radius of empirical networks	49
7.3	Conclusions	53
8	Relationship Between Algebraic Connectivity and Robustness of Complex Networks to Node and Link Failures	55
8.1	Relationship between algebraic connectivity and classical connectivity measures	56
8.2	Relationship between algebraic connectivity and classical connectivity measures in complex network models	57
8.2.1	Random graph of Erdős-Rényi	57
8.2.2	Small-World graph of Watts-Strogatz	59
8.2.3	Scale-Free graph of Barabási-Albert	61
8.2.4	Comparison between complex network models	63
8.3	Conclusions	64
9	Influence of Network Structure on Robustness of Complex Networks	67
9.1	Bounds for algebraic connectivity	68
9.2	Relationship between classical connectivity and probability distribution of algebraic connectivity	68
9.3	Behavior of algebraic connectivity under random failures	70
9.3.1	Random graph of Erdős-Rényi	71
9.3.2	Small-World graph of Watts-Strogatz	72
9.3.3	Scale-Free graph of Barabási-Albert	75
9.4	Conclusion	78
10	Conclusions	81

A Summary Statistics of Graph Measures for Number of Empirical Networks	85
B Dependence of Two Arbitrary Degrees in Random Graph of Erdős-Rényi	89
C Minimum of Normalized i.i.d. Binomially Distributed Sequence of Degrees in Random Graph of Erdős-Rényi	91
D Symbols and Acronyms	95
Bibliography	99
Samenvatting	105
Acknowledgements	107
Curriculum Vitae	109
NGInfra PhD Thesis Series on Infrastructures	113

Summary

Recent advances in the science of complexity have revealed that *complex networks* exist on many scales and in many different fields of science, as well as in nature. Such networks, it is now well established, share common properties in terms of their non-trivial network structure, called the *network topology*. In this thesis we survey a particular line of complex network research, mainly concerned with the characterization of non-trivial topological properties of complex networks. We expand on this research by analyzing those elementary graph measures, further classified into *structural and spectral measures*, that are of interest when quantifying topology-related aspects of the *robustness* of complex networks.

This thesis makes the following contributions to the field of complex networks. In the introductory part we present 1) a general consideration about complex networks research through the mathematical techniques of graph theory, 2) a brief description of the generic theoretical models used for modelling complex networks, and 3) a brief overview of practically important graph measures. In the main part we address the four research questions, which we summarize as follows:

1. Analysis of relations among a variety of proposed graph measures to propose a definite set, capable of expressing the most relevant topological properties of complex networks.
2. Study of the applicability of spectral measures to classify the qualitative topological properties that characterize specific classes of complex networks.
3. Study of the applicability of spectral measures to quantify different topological aspects of robustness of complex networks.
4. Practical application of spectral measures to quantify how the robustness to different types of failures manifests itself in the underlying complex network structure.

At first, we analyze relations among a variety of existing graph measures so as to make a first fundamental step in proposing a definite set, capable of expressing the

most relevant topological properties of complex networks. We extend this study by focusing on the topological properties in relation to the eigenvalues of the network's characteristic matrix. In particular, we adopt spectral graph theory to analyze the spectral properties of theoretical and empirical networks and show how this method aims at qualitative characterization of different classes of networks. We further illustrate the use of spectral measures to the question of quantitative characterization of different topological aspects of the robustness of complex networks: we introduce a particular eigenvalue of the network's characteristic matrix, referred to as the algebraic connectivity, as a quantifier of the robustness to disconnection or component separation in complex networks. Herewith we also introduce another particular eigenvalue of the network's characteristic matrix, referred to as the spectral radius, as a quantifier of the robustness to virus propagation in complex networks. Having suggested possible quantifiers of topology-related aspects of robustness, we analyze the relationship between the algebraic connectivity, and the classical measures of the extent to which a network can cope with its component failures. An immediate consequence of the relationship analysis is presented in the last part of the thesis, where we study the practical application of the algebraic connectivity to quantify how the robustness to different types of failures manifests itself in the underlying complex network structure. At last, we recapitulate the main results and provide a starting point for subsequent further learning in this particular line of complex network research.

Chapter 1

Introduction

1.1 Why studying complex networks

Complex networks emerge in a wide range of disciplines in the natural and social sciences as well as in nature. Frequently cited examples of complex networks include the Internet, the World Wide Web, power grids, transportations systems, food webs, ecosystems, genetic networks, neural networks, social networks, epidemiology networks, etc. Complex networks, in fact, are all systems of either physical (real) and/or logical (virtual) interconnected components where interactions between components give rise to intricate networks. Networks that emerge within such different disciplines are thus considered complex by virtue of their non-trivial network structure, called the network topology. The strength of the complex network study lies in the possibility of representing virtually any natural or social structure, including those undergoing dynamic changes of the network topology. Another strength of complex networks study lies in the fact that a proper knowledge of networks' structure is required to thoroughly understand and predict the overall network performance. For example, in computer networks, performance and scalability of protocols and applications, robustness to different types of perturbations (such as failures and attacks), all strongly depend on the network topology.

1.2 Scope of complex networks research

Complex networks are studied by applying theories and methods originally developed in the field of mathematical graph-theory and statistical physics. While origins of complex network study can be traced back to Leonard Euler's mathematical solution of the Koningsberg bridge problem in 1735 [32], research in complex networks became only recently a focus of attention. Supported by the availability of powerful computational equipment and techniques, the analysis of large-scale networks in the real world has

become possible. Consequently, complex networks research has received a tremendous amount of attention in a wide range of disciplines, resulting in a large number of recent related publications with a truly multidisciplinary nature. An example that was of major importance for the increased interest in the area of complex networks was the discovery of the network structure of the Internet [33]. Although coming from a wide range of disciplines, recent studies altogether have been interested in examining the structure of large-scale networks and have found that networks in real world exhibit specific properties that give rise to the non-trivial topological structure. In particular, three findings have initiated a revival of complex networks research: Watts and Strogatz's observation of small diameter and larger clustering that leads to small-world phenomenon in real-world networks [81], Barabási and Albert's observation of hub nodes that leads to scale-free phenomenon in real-world networks [6] and Newman's identification of community structures present in real-world networks [66].

Complex networks research is primarily divided into two major areas of network modelling and network analysis. The topology of complex networks traditionally has been modelled as the Erdős-Rényi random graph [29, 30, 31]. As mentioned earlier, the growing observation that real-world networks do not follow the prediction of random graphs has prompted many researchers to propose other models, such as small-world [81] and scale-free graphs [6]. One area of complex networks research is thus mainly engaged in presenting and discussing the complex network models and the corresponding analytical tools, covering almost every known aspect of random graphs, small-world and scale-free networks as well as variations of those models. Besides the modelling, much attention has been directed at analysing non-trivial topological properties by means of a definition of a variety of practically important measures, capable to quantitatively characterize topological aspects of the studied large-scale networks. In several papers, among which [2, 7, 22, 68], the authors have presented an extensive survey of such network topology measures. Much of those research efforts, however, are posed within a particular research interest, resulting in a characterization of complex networks from a specific narrow perspective or domain. The outcome of such an approach has a serious drawback: it does not ensure the mutual independence among a variety of proposed measures. In this context, having an increasing number of measures complicates attempts to find those properties that truly characterize complex networks.

1.3 Scope and contribution of thesis

In this thesis we focus on the question of quantitative characterization of non-trivial topological properties and its application to the robustness analysis of different topological aspects of complex networks. In particular, we aim to understand the relationship among a variety of proposed network topology measures so as to make a first fundamental step in proposing a definite set that would form the basis for a concise description of

a given complex network. Furthermore, we aim to analyze a method which allows qualitative characterization of different classes of networks. This classification is performed by adapting the spectral graph theory tool, a method that is used to study properties of a network through the eigenvalues of the network's characteristic matrix. To this part belongs also the application of two specific spectral measures, for which we aim to understand whether and to what extent they can be used as quantifiers of robustness to virus propagation and component failures in complex networks. Finally, we aim to quantify how and to what extent the robustness to different types of failures manifests itself in the underlying complex networks structure.

1.4 Thesis outline

The outline of the thesis is schematically depicted in Figure 1.1. The thesis consists of ten chapters. The first three chapters are the introductory chapters. While Chapter 1 gives the scope of complex networks research and the contribution of this thesis, Chapters 2 and 3 describe the essential tools for understanding the analysis presented in subsequent chapters of the thesis. On the one hand they cover the aspect of complex networks characterization, where the set of graph measures (being structural or spectral) is presented. On the other hand they cover the aspect of complex networks modelling, where various graphs are presented that are used nowadays for modelling the topology of complex networks.

The main body of the thesis comprises seven chapters. Chapter 4 analyses the relationships between graph measures and summarizes our main results on classification of the set of measures that would serve in future characterization studies of complex networks. Chapters 5, 6 and 7 study spectral measures, respectively referred to as the spectrum of the (normalized) Laplacian matrix, the algebraic connectivity and the spectral radius. The spectral measures are analyzed in generic complex network models as well as in a number of real-world networks. This analysis supports the claim that 1) the spectrum of the (normalized) Laplacian matrix is a powerful tool for spectral classification of different topological classes of networks, 2) the algebraic connectivity is an important measure in the analysis of the robustness to disconnection in complex networks and 3) the spectral radius plays an important role in modelling virus propagation in complex networks. Equipped with these results, we choose to focus on the algebraic connectivity to study its application to the robustness analysis. Chapter 8 studies the algebraic connectivity in relation to networks robustness to failures in generic complex networks models. Finally, Chapter 9 relies on the algebraic connectivity to study the influence of network structure on robustness to failures in complex networks.

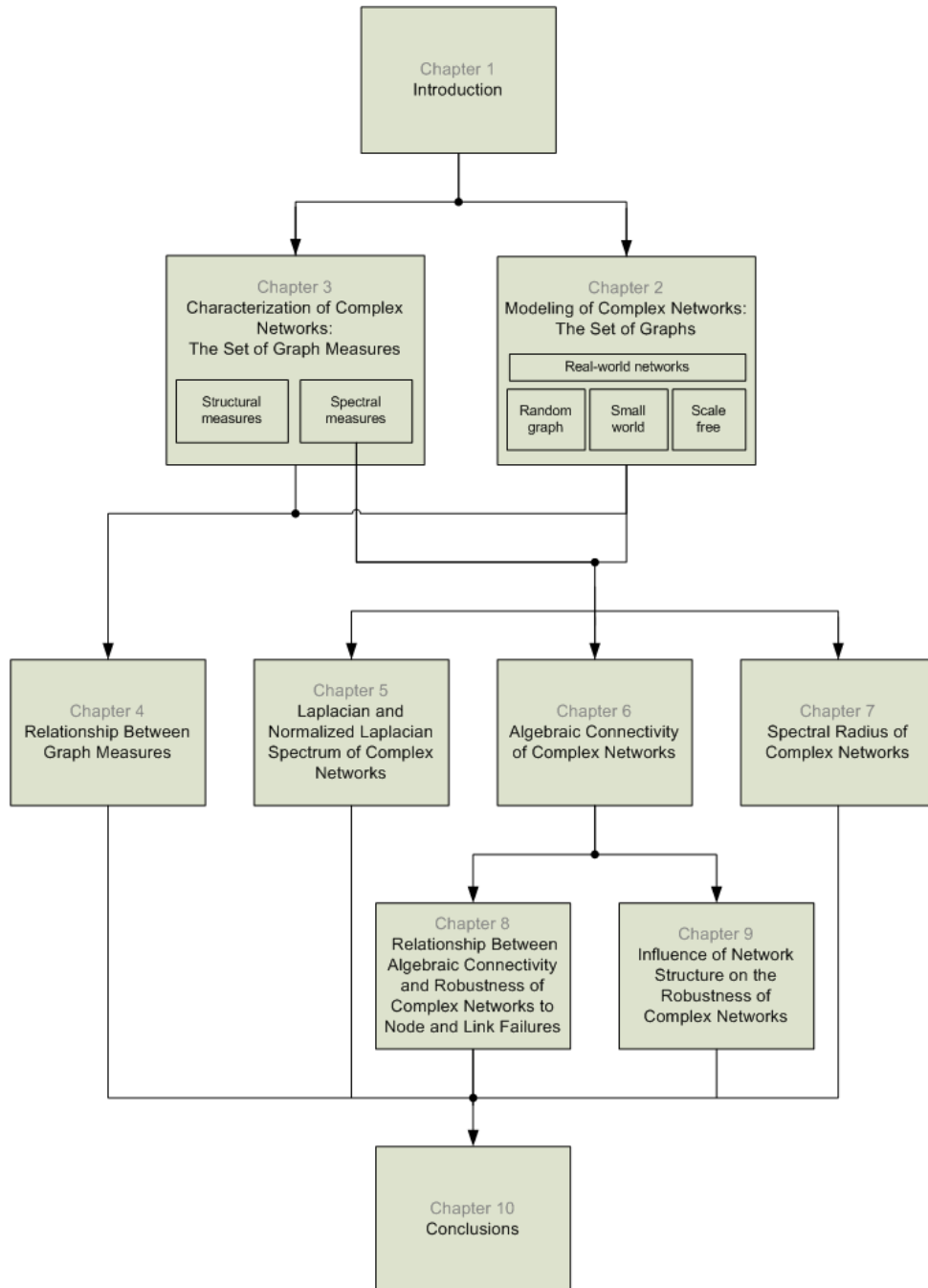


Figure 1.1: Schematic depiction of the outline of this thesis.

Chapter 2

Modeling of Complex Networks: The Set of Graphs

Traditionally, complex networks have been modelled as random graphs. As complex network science has continued to grow in importance and popularity, other models of complex networks have been developed. The two most well-known examples of recently introduced complex network models are those of small-world and scale-free graphs. This chapter introduces those three commonly cited models, which will be used throughout this thesis as a tool to facilitate the topological characterization of complex networks. Before giving the formal description of models, we establish the standard graph theory notation and complex networks terminology.

2.1 Notation

When studying complex networks we apply methods developed in a field of mathematics referred to as the graph-theory. In the context of graph-theory, mathematical structures called graphs are used to model pairwise connections between components of a network. A complex network is thus represented as graph $G = (\mathcal{N}, \mathcal{L})$, where \mathcal{N} denotes the set of components and \mathcal{L} the set of connections among them. The components are called nodes (vertices) and the connections among them links (edges). We only consider unweighted, undirected graphs, meaning that there is no distinction how two components are connected, or in terms of the graph theory, how an link may be directed from one node to another. In addition, a graph cannot contain self-loops or connections beginning and ending at the same component. In this way a complex network, as a system of nontrivial interconnected components, is modelled as a graph $G = (\mathcal{N}, \mathcal{L})$ with respectively $N = |\mathcal{N}|$ nodes and $L = |\mathcal{L}|$ links.

2.2 Random graph of Erdős-Rényi

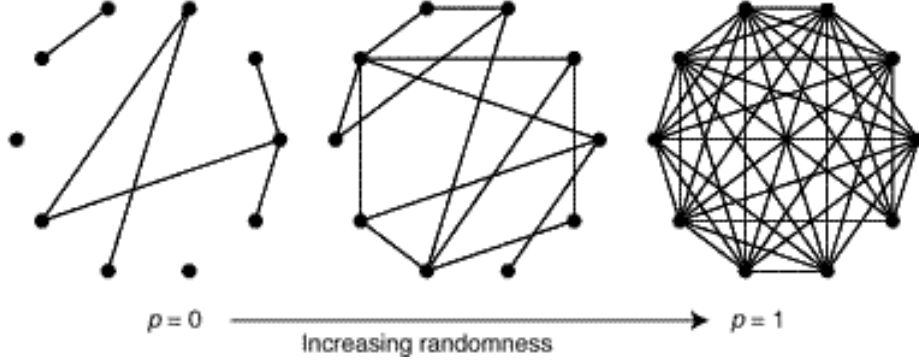


Figure 2.1: The random graph: increasing link probability p implies that we move from a low link density for which there are few links and many small components to a high link density for which an extensive fraction of all nodes are joined together in a single giant component. The figure is taken from [82].

The random graph of Erdős and Rényi is one of the most studied models of a network. We denote the random graph by $G_p(N)$, where N is the number of nodes in the graph and p is the probability of having a link between any two nodes (or shortly the link probability). $G_p(N)$ is the set of all such graphs in which a graph having L links appears with probability $p^L(1-p)^{L_{\max}-L}$, where L_{\max} is the maximum possible number of links. Many properties of the random graph are known analytically in the limit of large graph size N , as was shown by Erdős and Rényi in a series of papers in the 1960s [29, 30, 31] and later by Bollobás [10]. The most important, though, is that it possesses a phase transition: from a low link density or low p value for which there are few links and many small components to a high link density or high p value for which an extensive fraction of all nodes are joined together in a single giant component (see Figure 2.1).

2.3 Small-World graph of Watts-Strogatz

The small-world model describes the fact that despite the large graph size in most real-world networks there is a relatively short path between any two nodes. There are different realizations of the small-world model but the original model as proposed by Watts and Strogatz [81] is by far the most widely studied. It starts by building the ring R_N with N nodes and then joining each node to $2s$ neighbors (s on either side of the ring). This results in the ring lattice $R_{N,s}$ with $L = sN$ links. The small-world

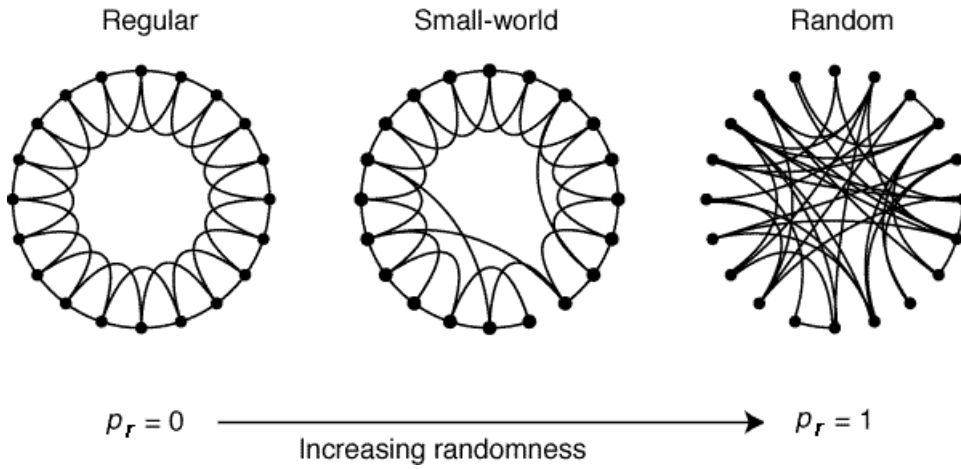


Figure 2.2: The small-world graph: increasing the rewiring probability p_r implies that we move from the completely ordered graph ($p_r = 0$) to the completely random graph ($p_r = 1$). The figure is taken from [81].

graph is then created by moving, with probability p_r , one end of each link (connected to a clockwise neighbor) to a new node chosen uniformly in the ring lattice, except that no double links or loops are allowed. This process introduces $p_r N$ long-range edges and allows the small-world model to interpolate between a regular lattice ($p_r = 0$) and something which is similar, though not identical, to a random graph ($p_r = 1$). For already small p_r the small-world becomes a locally clustered network in which two arbitrary nodes are connected by a small number of intermediate links (see Figure 2.2).

2.4 Scale-Free graph of Barabási-Albert

Interest in scale-free networks began in the late 1990s with the apparent discovery that in many real world networks some nodes act as "highly connected hubs", i.e. a few nodes have a large number of neighbors that is orders of magnitude larger than the average value. This is clearly in contrast to random and small-world graphs where every node has roughly the same number of neighbors (see Figure 2.3). This has allowed the identification of the class of scale-free networks in which a network is viewed as dynamic system that evolves through the subsequent addition and deletion of nodes and links. Barabási [6] offered two important general concepts that create networks with this characteristic feature: growth and preferential attachment of nodes. Consequently, the Barabási-Albert model incorporates these two concepts in the following manner: the model starts with a small number m_0 of fully-meshed nodes, followed at every step by a new node attached to $m \leq m_0 = 2m + 1$ nodes already present in the system. Each new node is connected to m of the existing nodes with a probability that is biased so

that it is proportional to the number of links that the existing node already has. This implies that the nodes with larger degree more likely are candidates for attachment of new nodes. After t steps this procedure results in a graph with $N = t + m_0$ nodes and $L = \frac{m_0(m_0-1)}{2} + mt$ links.

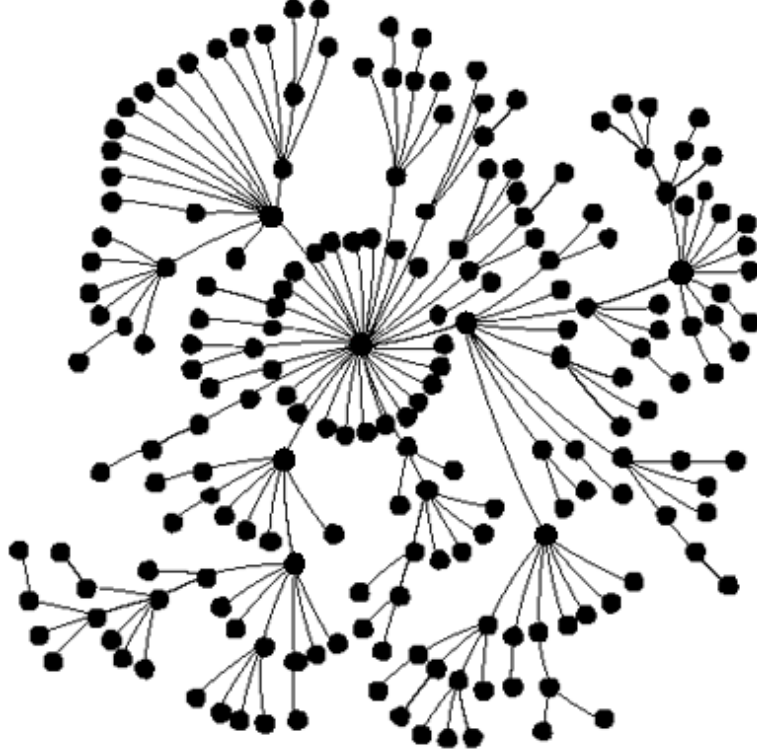


Figure 2.3: Scale-free graph: characteristic feature of these structures is that some nodes have a degree that is orders of magnitude larger than the average value. This figure is taken from [74].

Chapter 3

Characterization of Complex Networks: The Set of Graph Measures

Besides the modelling, the analysis of relevant topological properties is one of the major objectives that guide the research on complex networks. A topological property is any inherently graph-theoretical property, i.e. a property preserved under all possible topological changes of a graph. This means that networks with the same topological property naturally define a certain family of graphs: the ones that have a given specified topological property. Nevertheless, the term property usually refers only to descriptive characterization of networks. For example, the statement "a network does not have nodes with only one neighbor" is a property while "the number of nodes with more than two neighbors in a network equals six" is an topological measure. Consequently, over the past several years a variety of measures have been proposed to quantitatively express relevant topological properties of complex networks. This chapter presents a large set of measures that will be frequently used throughout this thesis more frequently termed as *graph* than as *network topology* measures. As we mentioned in Chapter 2, a complex network is represented as graph $G = (\mathcal{N}, \mathcal{L})$, where \mathcal{N} denotes the set of components and \mathcal{L} the set of connections among them.

3.1 Structural measures

Density

A density measures how complete a group is in terms of the relations among its members. Accordingly, the link density q is defined as the proportion of the maximum possible number of links that actually exist among all nodes.

Degree

The node degree D describes the number of links or neighbors a node has. The node degree distribution is the probability $\Pr[D]$ that the degree D of a randomly selected node has a given value. The number of links that on average connect to a node is called the average node degree. The average node degree can be easily obtained from the degree distribution through $E[D] = \sum_{d=1}^{D_{\max}} d \Pr[D]$, where D_{\max} is the maximum degree in a given graph.

The joint degree distribution $\Pr[D, D']$ is the probability that degrees D and D' of a randomly selected pair of nodes have a given values. A summary measure of the joint degree distribution is the average neighbor degree of nodes with a given degree value [57]. Another summary statistics that quantifies the correlation between pairs of nodes is the assortativity coefficient r : assortative networks have $r > 0$ (disassortative, i.e. $r < 0$ resp.) and tend to have nodes that are connected to nodes with similar (dissimilar resp.) degree [67].

Distance

The shortest path H describes the number of hops between a given pair of nodes. The distance distribution is the probability $\Pr[H]$ that the shortest path H between a random pair of nodes has a given value. From the distance distribution, the average node distance is derived as $E[H] = \sum_{h=1}^{H_{\max}} h \Pr[H]$, where H_{\max} is the longest among the shortest paths between any pair of nodes. H_{\max} is also referred to as the diameter *diam* of a graph. On the other hand, the eccentricity measures the largest distance between a given node and any other node in a graph. The average node eccentricity is the average of eccentricities of all nodes. Obviously, the maximum eccentricity equals the diameter of a graph.

Clustering

The clustering coefficient c_i of a node i is the proportion of the maximum possible number of links that actually exist among nodes within the neighborhood of a node i . For an undirected graph, a node i with degree D_i has at most $\frac{D_i(D_i-1)}{2}$ links among the nodes within its neighborhood. In other words, the clustering coefficient is the ratio between the number of triangles that contain node i and the number of triangles that could possibly exist if all neighbors of i were interconnected [81]. The clustering coefficient for the entire graph c_G is the average of clustering coefficients of all nodes.

The rich-club coefficient [19] is a measure that quantifies how close subgraphs, spawned by the k largest-degree nodes, are to forming a clique. The rich-club coefficient ϕ_x is the ratio of the number of links in the subgraph induced by nodes with degrees larger than a given value x to the maximum possible links between them $\frac{k(k-1)}{2}$.

Centrality

Betweenness is a centrality measure of a node or link within a graph: nodes (links) that occur on many shortest paths between other node pairs have higher node (link) betweenness than those that do not [42]. Average node (link) betweenness is the average value of the node (link) betweenness over all nodes (links).

Coreness

The k -core of a graph is a subgraph obtained from the original graph by the recursive removal of all nodes of degree less than or equal to k [8]. The node coreness of a given node is the maximum k such that this node is still present in the k -core but removed from the $(k + 1)$ -core. Hence, the k -core layer is the collection of all nodes having coreness k . The average node coreness is the average value of the node coreness over all nodes.

Connectivity

A graph is said to be connected if there exist a path between each pair of nodes. Similarly, when there is no path between at least one pair of nodes, a network is said to be disconnected. If in a graph there is an link between every pair of nodes, we say that the graph is complete K_N .

The link κ_L and the node κ_N connectivity are two classical connectivity measures of a graph. The link (edge) connectivity κ_L is the minimal number of links whose removal would disconnect a graph. The node (vertex) connectivity κ_N is defined analogously (nodes together with adjacent links are removed). For $k \geq 1$, a graph is k -connected if either a graph is a complete graph K_{k+1} or it has at least $k + 2$ nodes and no set of $k - 1$ nodes that separates it. Similarly, for $k \geq 1$ a graph is k -link connected if it has at least two nodes and no set of at most $k - 1$ links that separates it. The maximum value of k for which a connected graph is k -connected equals the node connectivity κ_N . The link-connectivity κ_L is defined analogously [9]. The minimum nodal degree D_{\min} of an incomplete graph is an upper bound on both the node and the link connectivity $\kappa_N \leq \kappa_L \leq D_{\min}$. If a graph is a complete graph K_N , then $\kappa_N = \kappa_L = D_{\min}$. If $\kappa_N = \kappa_L = D_{\min}$ we also say that the connectivity of a graph is optimal.

3.2 Spectral measures

Graph matrix

Methods developed in the graph theory also offer another representation of a network, i.e. the adjacency matrix, the Laplacian matrix or any other characteristic matrix of a graph. The adjacency matrix Ω of an undirected graph on N nodes is the $N \times N$

matrix where the nondiagonal entry (i, j) is either zero or one, depending on whether node i and node j are adjacent (connected). All diagonal entries (i, i) are zeros. In this way the adjacency matrix is symmetric. A particular point to note is that the number of nondiagonal entries in a row i equals the degree of i -node. A diagonal matrix, which contains information about the degree of each node, is denoted as the degree matrix Δ . It is used together with the adjacency matrix to construct the Laplacian matrix Λ of a graph $\Lambda = \Delta - \Omega$. The Laplacian matrix of an undirected graph on N vertices is the $N \times N$ matrix where the nondiagonal entry (i, j) is either zero or minus one, depending on whether node i and node j are adjacent. The diagonal entry (i, i) equals the degree of a node D_i . Another matrix representation of a graph is the normalized Laplacian matrix Π . The normalized Laplacian matrix of an undirected graph on N vertices is the $N \times N$ matrix where the nondiagonal entry (i, j) is either zero or minus the inverse square root of the i - and j -node degree, depending on whether node i and node j are adjacent. The diagonal entry (i, i) is either zero or one, depending on whether i -node degree equals zero.

Graph spectrum

The relationship between a graph and the eigenvalues (and eigenvectors) of its characteristic matrix is studied in the spectral graph theory. In the thesis we will consider the adjacency, the Laplacian matrix and the normalized Laplacian matrix. For the adjacency matrix all eigenvalues are real [21], while for the Laplacian [64] and normalized Laplacian matrix [17] all eigenvalues are real but also nonnegative. The ordered set of N eigenvalues of the adjacency matrix is called the spectrum of the adjacency matrix. The same holds for the Laplacian or the normalized Laplacian eigenvalues. There exists a unique adjacency or (normalized) Laplacian matrix for each graph (up to permuting rows and columns). The adjacency or (normalized) Laplacian matrix therefore depends on the node labeling while their set of eigenvalues is almost certainly a graph invariant [77]. Moreover, two graphs may be cospectral, i.e. they share the same eigenvalues but are not isomorphic. However, studies have shown [77] that only a small fraction of graphs are known to be uniquely determined by their spectra and that it is conceivable that almost all graphs have this property. Then, that two graphs with the same set of eigenvalues almost certainly lead to the same graph structure or a graph isomorphism. The set of eigenvalues of a characteristic matrix of a graph is therefore a powerful tool towards a qualitative classification of networks.

Graph eigenvalues

The largest eigenvalue of the adjacency matrix of a graph is denoted as the spectral radius ρ . The spectral radius plays an important role in modelling virus propagation in networks. In fact, the smaller the spectral radius, the larger the robustness of a network against the spread of viruses [79]. For instance, in a complete graph K_N the

spectral radius has the maximum possible while in a path P_N the spectral radius has the minimum possible value.

The second smallest eigenvalue of the Laplacian matrix is denoted as the algebraic connectivity λ_{N-1} . There are many problems in graph theory in which the algebraic connectivity plays a special role [59, 64]. In particular, its importance was emphasized by the fact that the value of the algebraic connectivity equals zero if and only if a graph is disconnected or it has two connected components. Moreover, if $\lambda_{N-i+1} = 0$ and $\lambda_{N-i} \neq 0$ then a graph has exactly i components. This also means that the multiplicity of zero as an eigenvalue of the Laplacian matrix corresponds to the number of components of the graph. Apart from the importance of the algebraic connectivity as the primary source of the number of connected components, the algebraic connectivity measures how well a graph is connected: the larger the algebraic connectivity is, the more difficult it is to cut a graph into independent components [34].

Chapter 4

Analyzing the Relationship Between Graph Measures

Over the past several years, a number of graph measures have been introduced to characterize the topology of complex networks. This chapter presents an analysis of the relationships between graph measures, with the aim of classifying a subset that would serve in future topological studies of complex networks [51, 50]. The classification of measures is based on statistical analysis methods of real data sets, representing the topology of different real-world networks. First, we relate pairs of measures by displaying their values as a collection of points, each having one coordinate on a horizontal and one on a vertical axis. Second, we perform correlation analysis to find out which of the measures are redundant. Finally, we apply principal components analysis (PCA) to support the classification of correlated measures. The presented methods reveal a clear relationship between graph measures, with many of the relationships not previously reported as being trivial. It should be noted that the studied real-world networks stem from as various as possible domains so as to avoid correlations that are due to structural constraints of the systems under study.

4.1 Visual analysis

Many complex networks are characterized by a power-law node degree distribution and a relatively short path between any two nodes. However, some complex networks may lack both, the power-law as well as the small-world character. Among the considered data sets, networks representing the topology of various transportation systems and power-grids are those where the two characteristics were not entirely encountered. In Figure 4.1 we show the node degree distribution of networks that do not obey a power-law behavior.

The average node degree is the coarsest characteristic of node interconnections. In

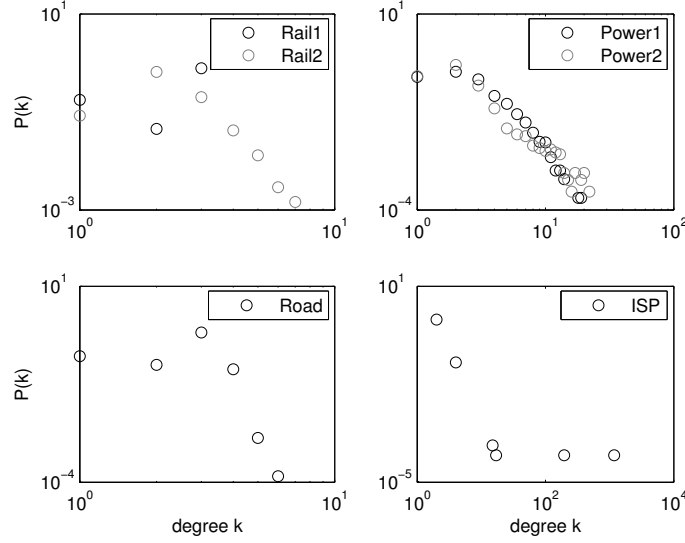


Figure 4.1: Real-world networks that do not obey a power-law degree distribution.

complex networks the average node degree is typically small and independent of the network size N . In Figure 4.2 we show, respectively, the relationship between the link density and the number of nodes, and number of links for various complex networks. As expected, for increasing N , the link density tends to zero and closely follows a power-law with exponent 1 (bottom of Figure 4.2). The link density is thus inversely proportional to the number of nodes while being inversely proportional to the square root of the number of links (top of Figure 4.2). From this it follows that the number of links is proportional to the number of nodes (not shown). Hence, in most complex networks, the classical assumption that $L = O(N)$ holds.

Node correlations play an important role in the characterization of the topology of complex networks. The most general approach to measure correlation among nodes is by means of the assortativity coefficient. On the top left scatter diagram in Figure 4.3 we show that disassortative networks, where high-degree nodes preferentially attach to other low-degree nodes, tend to be more clustered as their disassortativity increases. One should also notice from the ellipse on the top left scatter diagram in Figure 4.3 that those networks are typically assortative while having almost no clustering. The latter group of networks is made of various transportation and power-grid infrastructures. In addition, we observe that assortative networks, on average, have larger distances between pairs of nodes. The relationship between the assortativity coefficient and the average node distance is shown in the upper right scatter diagram of Figure 4.3.

Another interesting result that confirms these conclusions is found in [22]. Here, for the scale-free graph of Barabási-Albert [6], a negative correlation is observed between

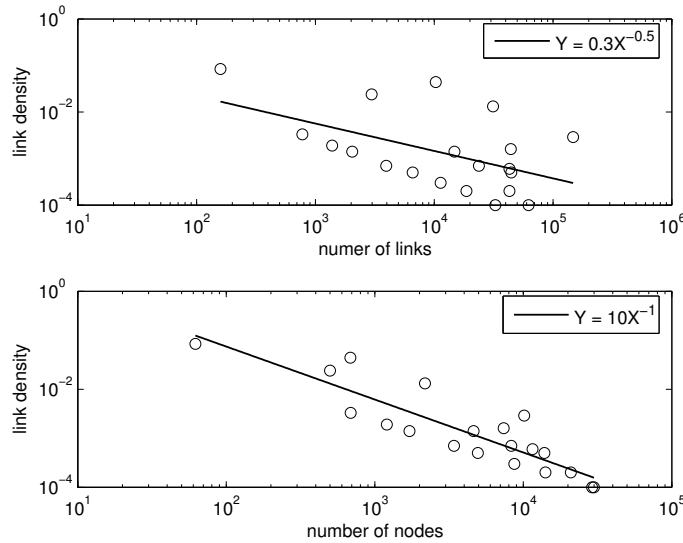


Figure 4.2: The link density as a function of the number of nodes and the number of links in real-world networks.

the assortativity coefficient and the clustering coefficient, and a positive between the assortativity coefficient and the average node distance. Consequently, a strong negative correlation is observed between the clustering coefficient and the average node distance. A possible explanation is that growth and preferential attachment, i.e. the two basis rules upon which the model is based, are responsible for additional links that tend to be established with the hubs, creating a more connected core and therefore contributing to higher clustering and smaller shortest paths.

Recently, it has been shown that complex networks are also characterized by the so called rich-club phenomenon [19]. The average distance between pairs of nodes as a function of the rich-club coefficient (lower left scatter diagram of Figure 4.3) yields that networks with smaller distance are much more likely to have high-degree nodes that form tight and well-interconnected subgraphs. As a result, one might expect that for disassortative networks, having on average smaller distance between pairs of nodes, the rich-club phenomenon would be evident as well. Nevertheless, on the lower right scatter diagram of Figure 4.3, we observe that the rich-club phenomenon is not trivially related to the mixing properties of networks. In other words, the rich-club phenomenon and the mixing properties express different features that are not trivially related or derived from each other.

On the other hand, topological measures associated with a certain feature, such as the shortest path length, are clearly related to each other. For example, average node betweenness increases as a function of average distance between pairs of nodes,

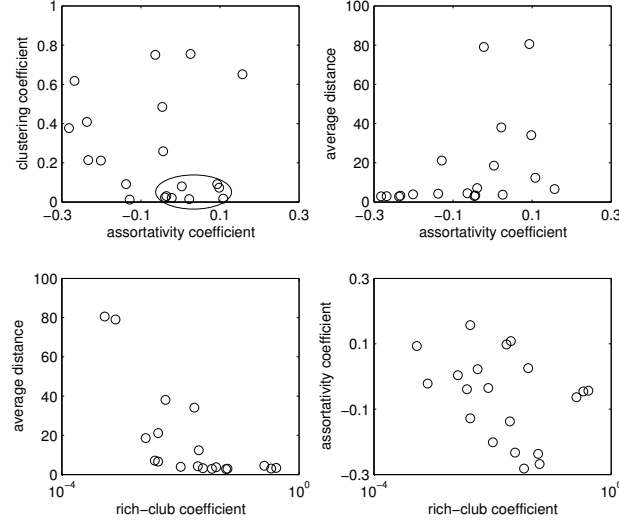


Figure 4.3: The relationship among topological measures for various real-world networks: clustering coefficient, assortativity coefficient, rich-club coefficient and average distance.

verifying that networks that have high average distance will also have nodes that occur on many shortest paths between other node pairs, and consequently, on average, have higher node betweenness. The Internet Service Provider network is a good example of a network for which high average distance between pairs of nodes results in high average node betweenness (see summary statistics presented in the Appendix A).

An important topological property, often ignored in the analysis of complex networks, is coreness. Node coreness refers to the degree of closeness of each node to a core of densely connected nodes, observable in the network [12]. In Figure 4.4 we report the relationship between average node coreness and the previously identified measures. The average node coreness as a function of the assortativity coefficient yields that social networks do not follow the generally observed trend of networks being disassortative but having, on average, higher node coreness. All three social networks are shown within an ellipse on the top left scatter diagram of Figure 4.4. At the same time, networks with higher average node coreness are more likely to have higher rich-club and clustering. Finally, we observe that the average node coreness is directly related to the average node degree. The former relationships are not surprising since on average, higher average node degree means higher rich-club and clustering, both for which we already perceived higher coreness.

Robustness to node and link failures is well captured by the algebraic connectivity. In essence, the algebraic connectivity quantifies the extent to which a network can ac-

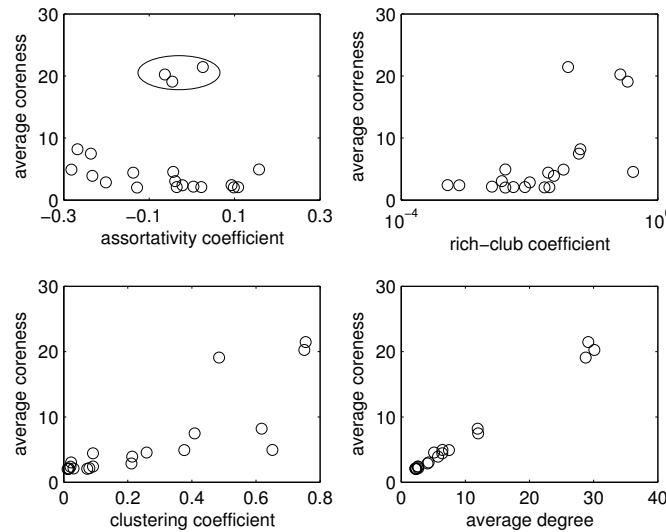


Figure 4.4: The relationship among topological measures for various real-world networks: clustering coefficient, assortativity coefficient, rich-club coefficient, average coreness and average degree.

commodate an increasing number of node- and link-disjoint paths. Figure 4.5 shows the relationships between the algebraic connectivity and the previously identified measures. The algebraic connectivity increases with the average node degree, as networks with higher average degree are better connected and consequently, are likely to be more robust. Note that in the literature [67] it is shown that assortative networks are less vulnerable to both random failures and targeted attacks. Here, we observe that disassortative networks have larger algebraic connectivity. This is not in contradiction with the observed tendency because it is most likely to be related to the hardness to cut the graph into independent components. Moreover, the larger the algebraic connectivity, the more networks seem to have a large rich-club and hierarchical nature. This implies that they have more well-interconnected and centrally-oriented nodes that occur on many shortest paths. Still, the average node betweenness does not seem to be related to the overall connectivity of a graph.

4.2 Correlation analysis

Correlation analysis aims at finding out linear relationships between variables. Variables are in our case the topological measures. From the tables presented in the Appendix A we derive a matrix whose columns are the different measures and the rows are the different real-world networks, denoted by X . We then compute the correlation matrix

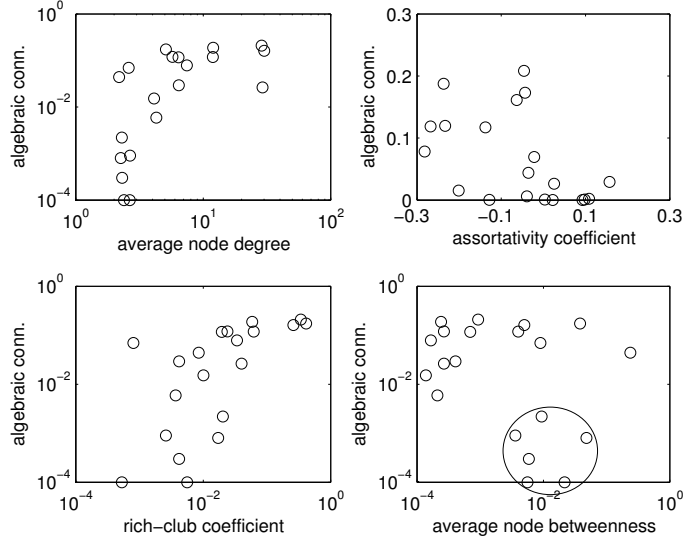


Figure 4.5: The relationship among topological measures for various real-world networks: assortativity coefficient, rich-club coefficient, algebraic connectivity, average node betweenness and average node degree.

of X , denoted by C . Matrix C is symmetric and has 1's elements on the diagonal. Each element (i, j) of C gives the correlation coefficient $c(i, j)$ between measures i and j (columns i and j of X). The correlation coefficient c varies between -1 and 1, and indicates whether the two variables are linearly correlated: positively if $c \sim 1$, negatively if $c \sim -1$, and uncorrelated if $c \sim 0$.

We are not interested in whether the correlation between two measures is positive or negative, but only how strongly two given measures are numerically related to each other. To ease the visualization, we show on Table 4.1 a symbolic encoding version of the correlation matrix. Table 4.1 displays the lower diagonal of the correlation matrix, using the following range of values and coding characters:

- $0 \leq |c| \leq 0.3$: " " (no correlation);
- $0.3 \leq |c| \leq 0.6$: "." (mild correlation);
- $0.6 \leq |c| \leq 0.9$: "+" (significant correlation);
- $0.9 \leq |c| \leq 1$: "#" (strong correlation).

The measures on Table 4.1 are identified by their number at the top of each column, and by the name and number on the left of each row. As the correlation matrix is symmetric, we show only the lower diagonal. First to be noticed is that 58 among the

Topological measures	1	2	3	4	5	6	7	8	9	10	11	12	13	14
Number of nodes (1)	1													
Number of links (2)	.	1												
Link density (3)	.		1											
Average degree (4)		.		1										
Average neighbour degree (5)					1									
Assortativity coefficient (6)					+	1								
Rich-club coefficient (7)	.		+	.			1							
Clustering coefficient (8)		.	.	+			.	1						
Average node distance (9)	.								1					
Average node eccentricity (10)	.								#	1				
Average node coreness (11)		.		#			.	+			1			
Average node betweenness (12)	.								#	#		1		
Average link betweenness (13)	.								#	#		#	1	
Algebraic connectivity (14)	+	.			.			1

Table 4.1: Correlation between topological measures for 20 various real-world networks.

91 lower diagonal elements (not counting the diagonal) have a correlation coefficient less than 0.3 in absolute value. Most measures are thus weakly correlated, indicating that most of them indeed reveal different topological aspects of real-world networks. 21 among the 91 lower diagonal elements correspond to mild correlations, i.e. $0.3 \leq |c| \leq 0.6$. Only 12 among the 91 lower diagonal elements correspond to strong correlations. Based on existing correlations between measures, we can identify the following clusters (see also Figure 4.6):

- **Distance cluster:** average node distance, average node eccentricity, average node and link betweenness.
- **Degree cluster:** average degree, average node coreness and clustering coefficient.
- **Intra-connectedness cluster:** link density, rich-club coefficient and algebraic connectivity.
- **Inter-connectedness cluster:** average neighbor degree and assortativity coefficient.

We labeled different measure clusters according to the type of topological information the group of measures provides. Intra- and inter- connectedness refer to the measures characterizing the observed connectivity, respectively, within and between a (sub)set of nodes in the network. All measures within each cluster are highly or partly topologically redundant. The 14 initial measures can thus be reduced to 6 (including the number of nodes and the number of links) since 8 of them are redundant with those of the same cluster. Besides the strength of the correlations within the groups, the correlation analysis shows to what extent some measures capture several topological properties of a network at once. For example, the number of nodes and the algebraic connectivity,

both exhibit mild correlation to 8 other measures. The number of nodes is related to the number of links and all measures within the distance and the intra-connectedness clusters, while not related to measures within the degree or the inter-connectedness clusters. The algebraic connectivity, on the other hand, is related to all measures within the degree, intra-connectedness, and the inter-connectedness clusters, but not to any measure in the distance cluster.

4.3 Principal component analysis

Correlation analysis measures the strength of correlation between variables. Understanding correlations, however, does not give insight about the number of independent variables, possibly derived from the set of correlated variables. In this context, correlated variables are the topological measures. Principal component analysis (PCA) [54] has proven to be useful for reducing the number of variables (dimensionality) while retaining most of the original variability in the data. The number of transformed, uncorrelated variables are called principal components, which in decreasing order account for as much of the variability in the data as possible.

We denote a given data set as a matrix X whose p columns are the variables to be analyzed $X_i, i = 1, \dots, p$. Each column (variable) has N elements, hence X is a $N \times p$ matrix. PCA performs a rotation of this matrix X such that

$$Y = A'X' \quad (4.1)$$

where A' is an orthogonal matrix¹. Y is the matrix of the rotated data, it is a square matrix of order N . A is found by constraining the covariance matrix of Y , $C_Y = \frac{1}{N-1}YY'$, to be diagonalized. A symmetric matrix can be diagonalized by the orthogonal matrix of its eigenvectors so that

$$C_Y = \frac{1}{N-1}AQA' \quad (4.2)$$

where $Q = XX'$. A is selected so that its columns are the eigenvectors of Q and the principal components of X . The diagonal elements of C_Y give the variance of X along each principal component.

The objective of PCA is to provide information about the minimal dimensionality, necessary to describe the data variability. The percentage of the total data set variance that is captured by a given number of principal components, is presented in Figure 4.7. The first principal component alone captures 76%, the first two components 94% and the first three components more than 99% of the total data set variance. PCA analysis shows that only 3 dimensions are enough to retain most of the original variability in the

¹A matrix is orthogonal if $A'A = I$, where I is the identity matrix.

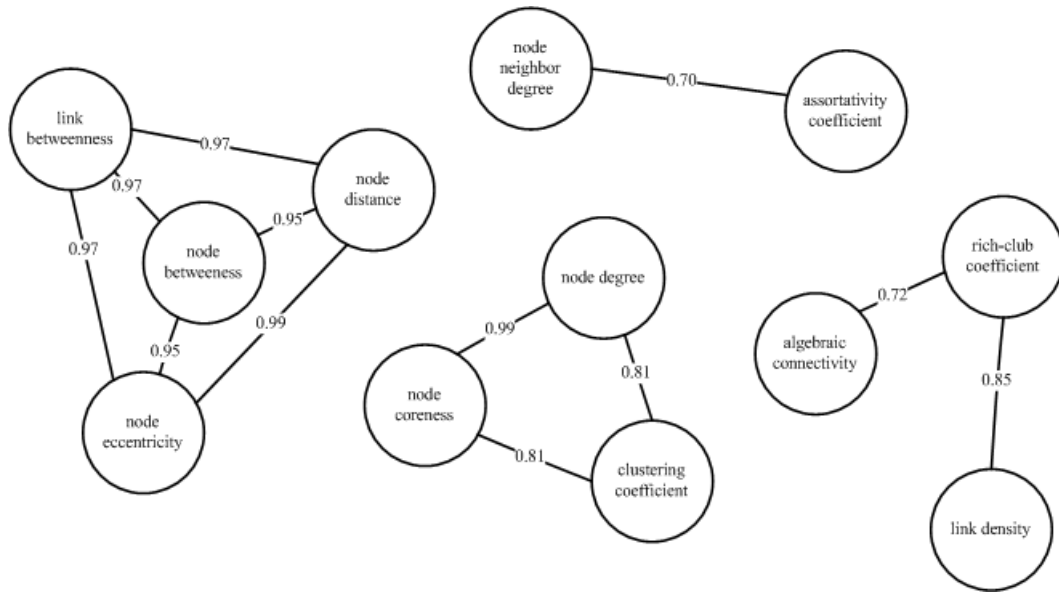


Figure 4.6: A graph in which nodes are topological measures and links the correlations that emerged from the correlation analysis. The corresponding values display the strength of the correlation between pairs of measures.

data. This, however, does not imply that the measures that are not important for the main principal components are unnecessary, but rather that they provide very specific topological information.

The reason why PCA was able to drastically reduce the dimensionality of the data set is because the principal components are a linear combination of all the measures. The first principal component is composed of two measures, i.e. the number of links and the number of nodes. All other measures have a very small weight in the linear combination of this principal component. In fact, the first principal component's measures are those missing from the four clusters we identified in the correlation analysis, presented in Subsection 4.2. The second principal component, besides the average node distance and the average node eccentricity, is also mostly made of the number of links and number of nodes. The third principal component is similar to the second in terms of which measures have the largest weights, but the sign of the weights differs as the principal components form an orthogonal basis. The fourth principal component, that captures a very small fraction of the total variance, is made almost exclusively from the average neighbor degree. PCA reveals that important measures that characterize the variations in the topological measures are the number of nodes and links and the measures within the distance and inter-connectedness clusters. measures within the degree and intra-connectedness clusters are redundant with the number of nodes and the number of links, since both the average degree and the link density can be recovered

from the former measures.

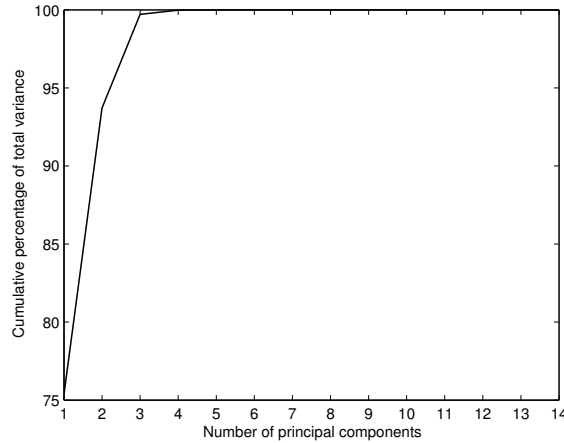


Figure 4.7: Fraction of the variance captured by the principal components.

4.4 Conclusions

In this chapter we studied the correlation between topological measures in real-world networks. The visual analysis revealed the following relationships among pairs of graph measures:

- The clustering coefficient increases with the increasing disassortativity. For assortative networks this relation is not trivial.
- The average node distance increases with the increasing assortativity coefficient and decreases with the increasing rich-club coefficient. Consequently, the assortativity coefficient decreases with the increasing rich-club coefficient.
- The average node coreness increases with the increasing rich-club and clustering coefficient while it decreases with the increasing assortativity coefficient. Furthermore, it is directly related to the average node degree.
- The algebraic connectivity increases with the increasing average node degree and the rich-club coefficient while it decreases with the increasing assortativity coefficient. The algebraic connectivity is not related to the average node betweenness.

The correlation analysis revealed that some measures are highly correlated, resulting in several clearly-defined clusters:

- Distance cluster: 1) the average node distance is strongly related to the average node eccentricity, 2) the average node (link) betweenness to the average node distance and hence 3) the average node (link) betweenness to the average node eccentricity;
- Degree cluster: 1) the average node degree is strongly related to the average node coreness and 2) the average node coreness to the clustering coefficient;
- Intra-connectedness cluster: 1) the rich-club coefficient is strongly related to the link density and 2) the algebraic connectivity to the rich-club coefficient;
- Inter-connectedness cluster: 1) the assortativity coefficient is strongly related to the average neighbor degree.

Overall, the analysis of the relationship between graph measures showed that some measures are either fully related to other topological measures or that they are significantly limited in the range of their possible values. Furthermore, the relationship analysis proved that subsets of graph measures are highly correlated, indicating redundancy among them. These observations further revealed that the set of commonly used measures is too extensive to concisely characterize the topology of complex networks. In this context this work is a first fundamental step towards classification and unification of a definite set of measures that would serve in future topological studies of complex networks.

Chapter 5

Laplacian and Normalized Laplacian Spectrum of Complex Networks

Eigenvalues are closely related to almost all critical properties of graphs [21, 17]. Relationships between graph properties and eigenvalues of the corresponding adjacency matrix have been much more studied in the past than relationships between graph properties and eigenvalues of other characteristic matrices [20, 21, 26]. In this chapter we show how the ordered set of eigenvalues of the Laplacian matrix is to be used as a spectral measure for classifying the qualitative properties that characterize specific classes of networks [46]. We perform an extensive set of experiments on theoretical complex network models so as to identify factors influencing the qualitative characterization. The results are presented in the first part of this chapter. In the second part, we expand on this work and perform a detailed Laplacian spectrum analysis of a number of empirical networks. In addition, we perform a detailed analysis of a number of empirical networks, though using the normalized Laplacian matrix. The empirical networks we consider here include a number of observed Internet topologies, collected using different methodologies from various locations in the world. Our results thus contribute to the recently conducted research on the spectral analysis of the observed Internet topologies [78, 36].

5.1 Laplacian spectrum of complex network models

We aim to identify factors influencing the classification of the qualitative properties that characterize specific classes of networks. In this section, we therefore present simulation results on theoretical complex network models.

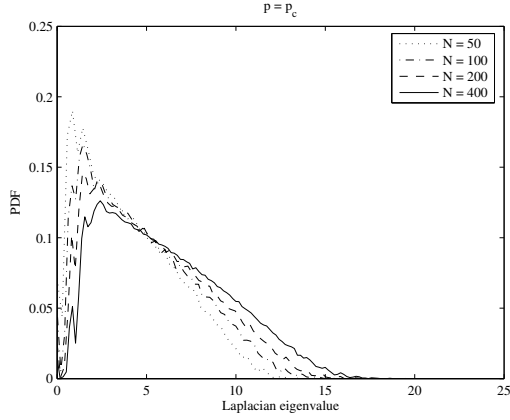


Figure 5.1: The spectrum of the Laplacian matrix of $G_p(N)$ with $N = 50, 100, 200$ and 400 , and $p = p_c$.

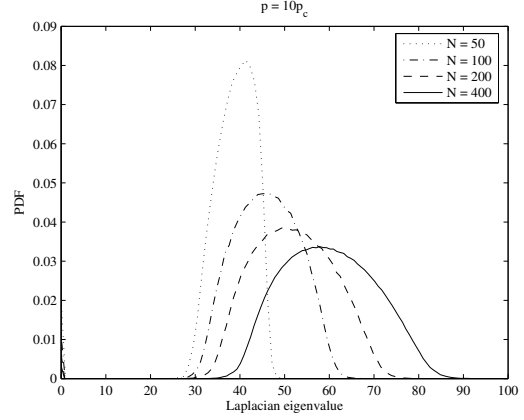


Figure 5.2: The spectrum of the Laplacian matrix of $G_p(N)$ with $N = 50, 100, 200$ and 400 , and $p = 10p_c$.

5.1.1 Random graph of Erdős-Rényi

In Section 2.2 we have seen that in the Erdős-Rényi random graph, denoted as $G_p(N)$, parameters of interest are the number of nodes N and the probability p of having a link between any two nodes. In this set of simulation we consider $G_p(N)$ with $N = 50, 100, 200$ and 400 nodes. The probability p of having a link between any two nodes is equal to $p = p_c \alpha$, where α ranges from 1 to 10 and $p_c = \frac{\log N}{N}$ is the critical threshold probability for which the Erdős-Rényi transforms to a graph with an extensive fraction of nodes joined in a single giant component¹. Then, for each combination of N and p , we have simulated 10^4 independent realizations of a random graph. For each independent realizations, the set of N eigenvalues of the Laplacian matrix has been computed. By picking, for each realization, at random one out of N eigenvalues, we collect 10^4 eigenvalues from which the Laplacian spectrum is plotted. Figures 5.1, 5.2 show the Laplacian spectrum of $G_p(N)$ for the link probability $p = p_c$ and $p = 10p_c$, and increasing number of nodes N . At the critical threshold probability $p = p_c$ there exists random graphs that are not connected, therefore we consider only connected Erdős-Rényi random graphs.

For $p = p_c$ the spectrum is skewed with the main bulk pointing towards the small eigenvalues. Such behavior of a Laplacian spectrum is often found in cases where the topology has a tree-like structure. An extreme case of such type of structure is the star $K_{1,N-1}$, where the eigenvalues are $N, 0$ and 1 (with multiplicity $N - 2$). In order to examine this in more detail, we plot the spectrum of the minimum spanning tree (MST),

¹The value of the link probability p above which a random graph almost surely becomes connected tends, for large N , to $p \sim p_c = \frac{\log N}{N}$ [10].

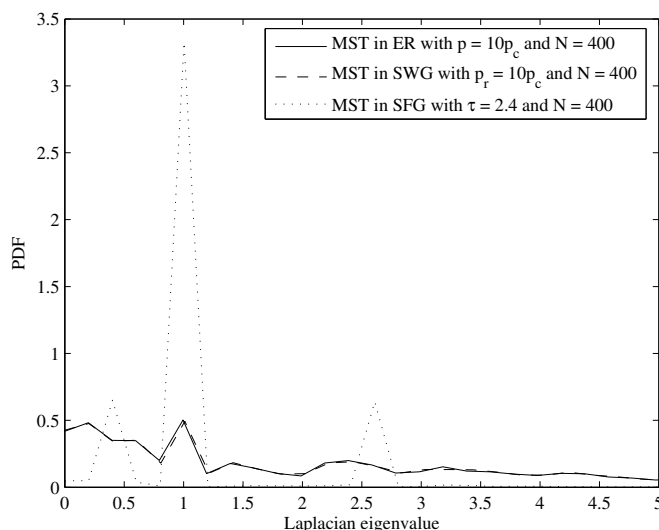


Figure 5.3: The Laplacian spectrum of the MST in the Erdős-Rényi random graph, the Watts-Strogatz small-world graph and the Havel-Hakimi scale-free graph, all with $N = 400$ nodes.

found in each of 10^4 independent configurations of $G_p(N)$. Figure 5.3 shows that the spectrum of the MST in $G_p(N)$ is indeed highly skewed with the main bulk pointing towards the small eigenvalues. In particular, the underlying tree with degree 1 nodes is responsible for the peak at $\lambda = 1$. The spectrum of sparse $G_p(N)$ shows a similar behavior at small eigenvalues (see Figure 5.1), what can be interpreted as the structure that is mainly determined by the underlying tree. Such behavior is more obvious at smaller N , since the larger graph size cause an increase in the link density. More important is that the maximum λ_{N-1} of a tree on $N \geq 3$ is 1 and $\lambda_{N-1} = 1$ if and only if the underlying graph is the star $K_{1,N-1}$. At the other extreme, the minimum λ_{N-1} occurs at the path P_N , namely $\lambda_{N-1}(P_N) = 2 \left[1 - \cos\left(\frac{\pi}{N}\right)\right]$. Thus, roughly speaking λ_{N-1} decreases as the diameter increases [38]: for the MST in $G_{p_c}(N)$, $\lambda_{N-1} \ll 1$ while for sparse $G_p(N)$ $\lambda_{N-1} < 1$, implying that the underlying tree-like structure of a sparse $G_p(N)$ has a small diameter.

For $p = 10p_c$ the spectrum has a bell shape (see Figure 5.2), centered around the mean nodal degree $E[D] = p(N-1)$. Moreover, for fixed $p = 10p_c$, the high peak becomes smaller while the bell shape becomes wider, representing that, for increasing N , the spectrum variance is in agreement with the Wigner's Semicircle law [60]. In fact, the spectrum is pointing to uncorrelated randomness that is a characteristic property of an Erdős-Rényi random graph [60]. Hence, the Laplacian spectra of the Erdős-Rényi random graph are indicating that, for increasing link density, the underlying structure

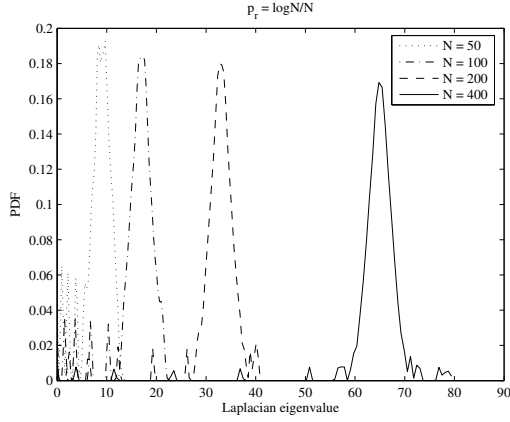


Figure 5.4: The spectrum of the Laplacian matrix of the Watts-Strogatz small-world graph with $N = 50, 100, 200$ and 400 , and $p_r = \frac{\log N}{N}$.

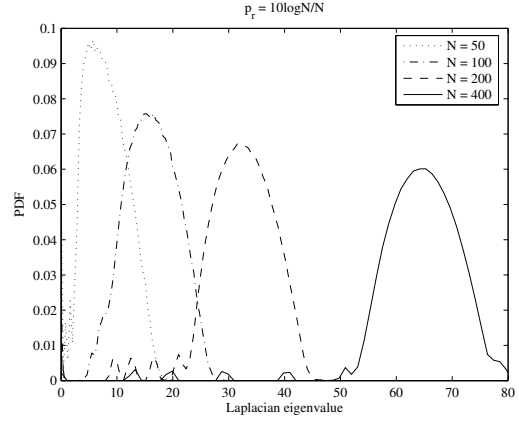


Figure 5.5: The spectrum of the Laplacian matrix of the Watts-Strogatz small-world graph with $N = 50, 100, 200$ and 400 , and $p_r = \frac{10 \log N}{N}$.

of the Erdős-Rényi random graph transforms from a tree-like structure with a small diameter into a more homogeneous graph where the degree is closely centered around the mean degree.

5.1.2 Small-world graph of Watts-Strogatz

In this set of simulations we consider the Watts-Strogatz small-world graph built on the ring lattice $R_{N,s}$ with $N = 50, 100, 200$ and 400 nodes. For each graph size N , every node is connected to its first $2s$ neighbors (s on either side). In order to have a sparse but connected graph we have considered $N \gg 2s \gg \ln N$ in the following ring lattice graphs: $C(50, 4)$, $C(100, 8)$, $C(200, 16)$, $C(400, 32)$. The small-world model is then created by moving, with probability p_r , one end of each link to a new location chosen uniformly in the ring lattice, except that no double links or self-edges are allowed. The rewiring probability p_r equals the link probability in the random graph $G_p(N)$: it starts from $p_r = \frac{\log N}{N}$ and ends with $p_r = \frac{10 \log N}{N}$. Furthermore, for each combination of the graph size N , the neighbor size s and the rewiring probability p_r we have simulated 10^4 independent realizations of the Watts-Strogatz small-work graph. For each independent realization, the set of N eigenvalues of the Laplacian matrix has been computed. By picking, for each realization at random one out of N eigenvalues, we collect 10^4 eigenvalues from which the Laplacian spectrum is plotted.

For the small rewiring probability $p_r = 0$, the Watts-Strogatz small-world graph is regular and also periodical. Because of the highly ordered structure we see in Figure 5.4 that for small p_r the spectrum is highly skewed with the bulk towards the high

eigenvalues, distributed around the mean nodal degree, which irrespective of p_r equals $E[D] = 2s$. The spectrum of the two-dimensional lattice graph with $N \times N$ nodes aims to illuminate this effect. The Laplacian spectrum of the two-dimensional lattice is the sum of two path graphs P_N whose eigenvalues are $\lambda_i(P_N) = 2 - 2\cos(\pi i/N)$, $i = 1, 2, \dots, N$. Consequently, the spectrum of the two-dimensional lattice converges to a pointy shape with a peak centered around the mean nodal degree, which for $N \rightarrow \infty$ converges to 4. The same tendency is observable in the Watts-Strogatz small-world graph: in Figure 5.4, the bulk part centered around the mean nodal degree, alongside the remaining peaks means that the graph is still highly regular and periodical. In fact, the Laplacian spectrum of the ring lattice $R_{N,s}$ with N nodes and $2s$ neighbors comprises the eigenvalues $\lambda_i(R_{N,s}) = 2s - \left(\frac{\sin(\frac{\pi}{N}(i-1)(2s+1))}{\sin(\frac{\pi}{N}(i-1))} - 1 \right)$, $i = 1, 2, \dots, N$. Hence, upon increasing s -regularity, the bulk part of the spectrum shifts towards the mean nodal degree, similar to the Laplacian spectrum of the Erdős-Rényi random graph. In order to examine this in more detail we have calculated the fraction between the largest and the second smallest Laplacian eigenvalue. The fraction in the small-world graph with $p_r = \frac{\log N}{N}$ and $N = 400$ is approximately 4 times larger than the fraction in the small-world with $p_r = \frac{10 \log N}{N}$, indicating that the entire Laplacian spectrum of the small-world graph shifts towards λ_1 . This transition of the bulk spectrum is known as the spectral phase transition phenomenon [69].

5.1.3 Scale-free graph of Havel-Hakimi

In this set of simulations we consider a scale-free graph, which for a given degree sequence constructs a graph with a power-law degree distribution. Havel [41] and Hakimi [40] proposed an algorithm that allows us to determine which sequences of nonnegative integers are degree sequences of graphs. Practically the same as the Barabási-Albert model, this model will have degree distribution with a power-law tail $\Pr[D = k] \approx ck^{-\tau}$, where $c \approx (\zeta(\tau))^{-1}$ and the exponent τ typically lies in the interval between 2 and 3. In order to have a graph which is in agreement with the real-world networks [68], we have used the exponent $\tau = 2.4$. The degree distribution with the exponent $\tau = 2.4$ is shown in Figure 5.6. Then, for each combination of the graph size N and the exponent τ we have simulated 10^4 independent configurations of the power-law graph. For each independent realization the set of N eigenvalues of the Laplacian matrix has been computed. By picking, for each realization at random one out of N eigenvalues, we collect 10^4 eigenvalues from which the Laplacian spectrum is plotted.

As shown in Figure 5.7, the spectrum of the Havel-Hakimi scale-free graph is completely different from the spectra of the other two complex network models. Because of the highly centralized structure the spectrum in Figure 5.7 is skewed with the bulk towards the small eigenvalues. Recall that the Laplacian spectrum of the star $K_{1,N-1}$ is $N, 0$ and 1 (with multiplicity $N - 2$). Consequently, the spectrum is indicating that an

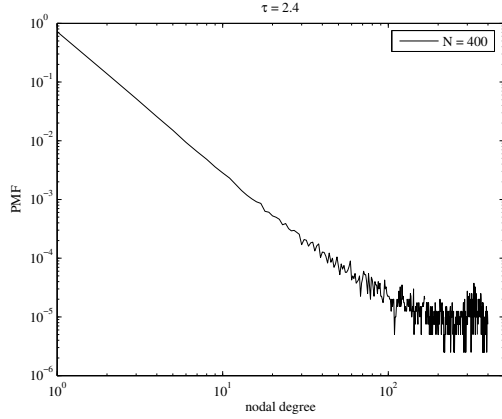


Figure 5.6: The degree distribution of the Havel-Hakimi scale-free graph with $N = 50, 100, 200$ and 400 , and $\tau = 2.4$.

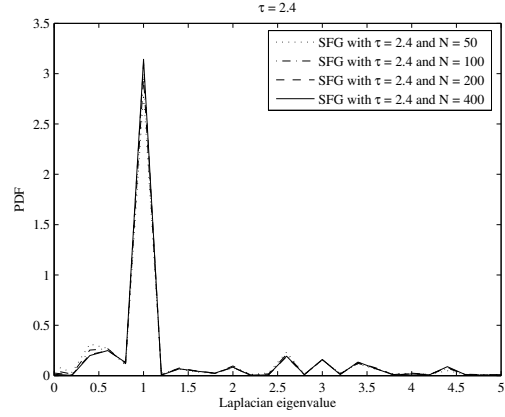


Figure 5.7: The spectrum of the Havel-Hakimi scale-free graph with $N = 50, 100, 200$ and 400 , and $\tau = 2.4$.

underlying structure of this type of the scale-free graph is a star-like structure with few highly connected nodes: although peaks at $\lambda = 2$ and $\lambda = 3$ have vanished, the MST found in the Havel-Hakimi scale-free exhibits a visually similar spectrum (see Figure 5.3). This means that most likely peaks in a spectrum, exemplified here with the peak at $\lambda = 1$, are due to the majority of nodes with the corresponding degrees. Moreover, for the connected graph the product of the non-zero Laplacian eigenvalues equals N times the number of Spanning Trees (ST) found in the corresponding graph [64]. From the simulation results we have found that the number of ST in sparse $G_p(N)$ is much higher than the number found in the Havel-Hakimi scale-free graph. In addition, we have found that the sum of the eigenvalues in $G_p(N)$ that equals the sum of the degrees, i.e. $\sum_{i=1}^N \lambda_i = \sum_i D_i$, is about double the sum of the eigenvalues found in the scale-free graph. Also, the largest Laplacian eigenvalue [64], which is bounded by $[\frac{N}{N-1}D_{\max}, 2d_{\max}]$, grows approximately with N . Hence, the structure of this type of a scale-free graph is highly concentrated around nodes with very large nodal degrees.

5.2 Laplacian spectrum of empirical networks

We have calculated the spectrum of the Laplacian matrix of an observable part of the Internet graph, extracted from the traceroute measurements performed both via RIPE NCC [72] and PlanetLab [71]. The resulting graphs are an observed Internet graphs at the IP-level because the traceroute utility returns the list of IP-addresses of routers along the path from a source to a destination. In fact, a graph obtained from traceroute measurements is an approximation of the Internet graph at the router-

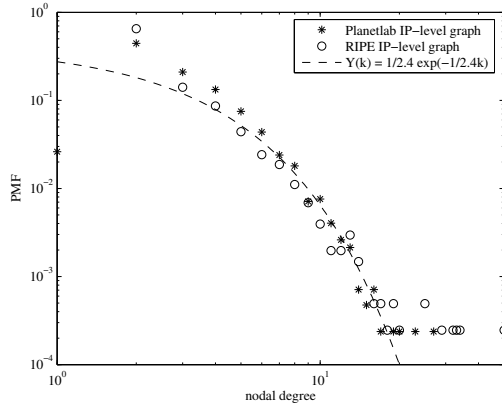


Figure 5.8: The degree distribution of an observable part of the IP-level Internet graph, performed via RIPE and Planetlab.

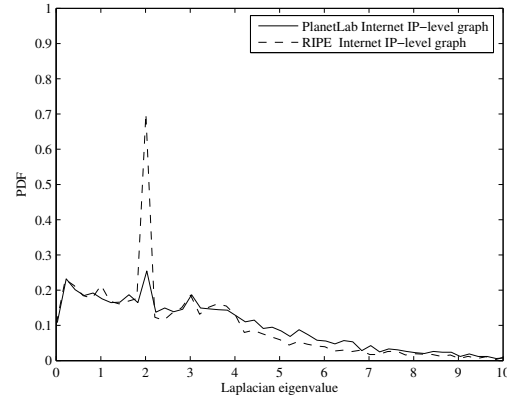


Figure 5.9: The Laplacian spectrum of an observable part of the IP-level Internet graph, performed via RIPE and Planetlab.

level, which again is the union of shortest paths between each pair of a small group of routers. This explains why such graph is denoted as the overlay graph on top of the actual Internet topology. The RIPE NCC measurements, executed on September 18th 2004, have resulted in a graph consisting of 4058 nodes and 6151 links and the PlanetLab experiments, executed on November 10th 2004, in a graph with 4214 nodes and 6998 links.

Figure 5.8 shows the degree distribution and Figure 5.9 the Laplacian spectrum of the observed graphs. In spite of two different sources of traceroute measurements, the Laplacian spectrum stays almost the same: both Laplacian spectra contain a peak at $\lambda = 2$, which most likely is due to the majority of nodes with degree 2, or repeated duplication of nodes underlying the evolution of the observed graph. Besides, the Laplacian spectra contain smaller peaks at $\lambda = 1$ and $\lambda = 3$, although the first one only appears in the spectrum of the graph observed via RIPE. The peak at $\lambda = 3$ possibly also originates from a significant amount of degrees, which were created through repeated duplication of nodes, while the peak at $\lambda = 1$ surely does not, since the graph observed via RIPE does not contain nodes with degree 1. Furthermore, we observe that the two observed graphs differ substantially from the spectra of generic complex networks models. Also, we find that the Erdős-Rényi random and the Watts-Strogatz small-world graph show a similar spectral behavior, which differs considerably from that of the scale-free graph derived from a Havel-Hakimi power-law degree sequence. Despite this discrepancy, the spectrum of MST in the Erdős-Rényi random graph with uniformly distributed link weights does bear resemblance to the spectra of the two observed graphs (see Figure 5.3). The difference in spectra between empirical and synthetic graphs could

be due to the fact that the observed part is a subgraph of the complete Internet graph, where the observed part is mainly tree-like structured just as the MST is.

5.3 Normalized Laplacian spectrum of empirical networks

The adjacency or the Laplacian matrix both employ different normalization and therefore lead to different graph spectra. In this section we focus on the spectrum of the normalized Laplacian matrix [17], where all eigenvalues lie between 0 and 2 allowing easy comparison of networks of different sizes. Figure 5.10 displays the cumulative distribution function of the eigenvalues computed from the normalized Laplacian matrix of four empirical networks. The empirical networks are an observed Internet graphs at AS-level, extracted from four datasets: Chinese [83], Skitter [14], RouteViews [73] and UCLA [70] dataset. The difference between the empirical networks is most easily observed around the eigenvalues equal to 1. These eigenvalues play a special role as they indicate repeated duplications of structural patterns within the network [4]. By duplication, we mean different nodes having the same set of neighbors giving their induced subgraphs the same structure. Through repeated duplication one can create networks with eigenvalue 1 of very high multiplicity. In addition, we observe that the spectra have a high degree of symmetry around the eigenvalue 1. If a network is bipartite $K_{N,N}$, i.e. it consists of two connected parts, each with N nodes and no links between nodes of the same part, then its spectrum will be symmetric about eigenvalue 1. Consequently, the observed Internet graphs at AS-level appear close in spectral terms to a bipartite graph, another phenomenon that arises through repeated structure duplication. In the AS-level Internet graph many ASes share a similar set of upstream ASes without being directly connected to each other. In particular, although small ASes may tend to connect to large upstream providers, they might not connect preferentially to the largest ones, connecting instead to national or regional providers. In summary, these results provide further evidence that the spectrum is a powerful tool express quantitatively important network properties.

5.4 Conclusion

This chapter has evaluated the spectrum of the Laplacian matrix and the normalized Laplacian matrix of theoretical complex network models and a number of empirical networks. This study was motivated by the fact that eigenvalues are closely related to almost all critical network properties. Both the empirical and theoretical network analysis have indicated that the (normalized) Laplacian spectrum yields a set of characteristic spectral properties that on one hand captures what it is specific about the

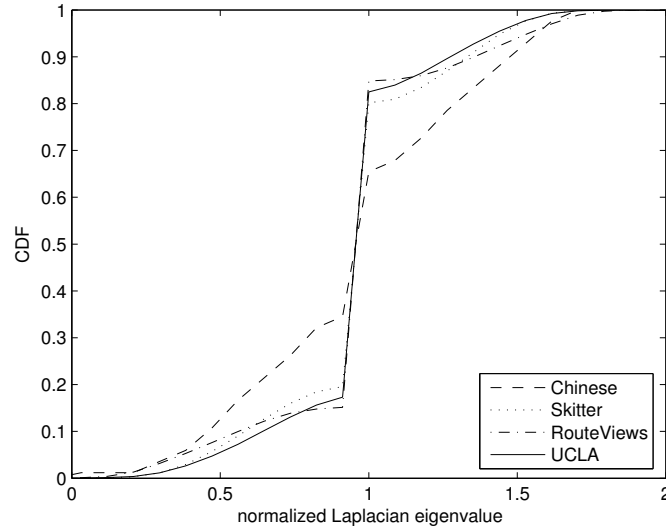


Figure 5.10: The probability density function of the normalized Laplacian spectrum of four samples of the observable part of the AS-level Internet graph: Chinese, Skitter, RouteViews and UCLA dataset.

topology of a network, and on the other hand simultaneously accounts for all its important properties. The spectrum is therefore a powerful tool for classifying the qualitative spectral properties that characterize specific classes of networks. The analysis of the graph spectrum establishes a first step towards a systematic characterization of complex networks.

Chapter 6

Algebraic Connectivity of Complex Networks

In this chapter we rely on the second smallest Laplacian eigenvalue to quantitatively characterize the robustness of complex networks [47]. This particular eigenvalue is often referred to as the algebraic connectivity [34] for the following reasons: 1) a graph is disconnected if and only if the algebraic connectivity is zero, 2) the multiplicity of zero as an eigenvalue of a graph is equal to the number of disconnected components. There is a vast literature on the algebraic connectivity: see e.g. [58, 59, 62] for surveys and e.g. [64, 63] for applications to several difficult problems in graph theory. For the purpose of this work the most important is though its application to the robustness: 1) the larger the algebraic connectivity is, the more difficult it is to cut a graph into independent components, 2) its classical upper bound in terms of the node and the link connectivity provides a worst case robustness to node and link failures [34].

This chapter is organized as follows. In the first part we study the behavior of algebraic connectivity in the Erdős-Rényi random graph. The Erdős-Rényi random graph has been traditionally used to model the topology of complex networks. Besides, for the Erdős-Rényi random graph most of the interesting properties can be analytically expressed. This is in contrast to most other complex network graph models where computations are hardly possible. In the first part, by using the basic approximation that the algebraic connectivity equals the minimum nodal degree, we analytically derive the estimation of the mean and the variance of the algebraic connectivity for the Erdős-Rényi model. Hereby we improve an already existing theorem concerning its behavior [55]. In the second part, we study the relationship between the algebraic connectivity and the node, and the link connectivity. Then, the existing relations between the three measures for the Erdős-Rényi graph are also refined. Finally, we conclude by summarizing our main results on the role the algebraic connectivity has for robustness of complex networks.

6.1 Algebraic connectivity in random graph of Erdős-Rényi

In this section we give an analytical estimate of the algebraic connectivity in the Erdős-Rényi random graph. The analytical estimate relies on the equality with the minimum nodal degree. This approximation is verified by a comprehensive set of simulations, presented in Subsection 6.1.4. Prior to analyzing the minimum nodal degree in Subsection 6.1.2, we give some details on the asymptotic behavior in the Erdős-Rényi random graph.

6.1.1 Asymptotic behavior in random graph of Erdős-Rényi

Many properties of the random graph can be determined asymptotically, as was shown by Erdős-Rényi in a series of papers in the 1960s [29, 30, 31] and later by Bollobas in [10]. Typically, for $N \rightarrow \infty$ and constant mean nodal degree $E[D]$, the binomial degree distribution

$$\Pr[D_i = j] = \binom{N-1}{j} p^j (1-p)^{N-1-j} \quad (6.1)$$

tends to a Poisson degree distribution [60]. On the other hand, for $N \rightarrow \infty$ and a constant link density p , the Central Limit Theorem states that the normalized i.i.d. binomially distributed sequence $\{D_i^*\}_{1 \leq i \leq N}$ of all degrees in $G_p(N)$ tends to be Gaussian distributed [60]

$$D_i^* = \frac{D_i - E[D]}{\sigma[D]} \xrightarrow{d} \frac{e^{-\frac{x^2}{2}}}{\sqrt{2\pi}} \quad (6.2)$$

where $E[D]$ is the mean and $\sigma[D] = \sqrt{\text{Var}[D]}$ is the standard deviation of the nodal degree. In summary, the limit distribution of the degree D_i of an arbitrary node i in $G_p(N)$ depends on how the probability p varies with the number of nodes N when $N \rightarrow \infty$. Then, for a random graph to be connected there must hold that for large N , $p \geq \frac{\log N}{N} \equiv p_c$. Moreover, in their classical paper [29], Erdős-Rényi have proved that for large graph size N the probability that a random graph $G(N, L)$ with $L_x = \left[\frac{1}{2}N \log N + xN\right]$ is connected equals $e^{-e^{-2x}}$ or rewritten [60] in terms of $L = \frac{1}{2}N \log N + xN$ and then p ,

$$\Pr[G_p(N) = \text{connected}] \simeq e^{-Ne^{-p(N-1)}}. \quad (6.3)$$

Finally, the probability that the node connectivity κ_N equals the link κ_L connectivity, which in turn equals the minimum nodal degree D_{\min} , approaches 1 as N approaches infinity or that

$$\Pr[\kappa_N = \kappa_L = D_{\min}] \rightarrow 1 \text{ as } N \rightarrow \infty \quad (6.4)$$

is also proved in [10] and holds without any restriction on p . This was also shown by Bollobás and Thomason in [11]. On the other hand, the asymptotic behavior of the

algebraic connectivity in the Erdős-Rényi random graph $G_p(N)$ is proved by Juhász in [55]: For any $\varepsilon > 0$,

$$\lambda_{N-1} = pN + o\left(N^{\frac{1}{2}+\varepsilon}\right) \quad (6.5)$$

where the algebraic connectivity converges in probability as $N \rightarrow \infty$.

6.1.2 Minimum nodal degree in random graph of Erdős-Rényi

In $G_p(N)$ each node i has a degree D_i that is binomially distributed (6.1). Before proceeding, we first need to show that degrees in the sequence $\{D_i\}_{1 \leq i \leq N}$ are almost independent random variables. In any graph $\sum_{i=1}^N D_i = 2L$ holds, thus degrees in the sequence $\{D_i\}_{1 \leq i \leq N}$ are not independent. However, if N is large enough, D_i and D_j are almost independent for $i \neq j$ and we can assume that all D_i are almost i.i.d. binomially distributed (see also [10, p. 60]). The following Lemma quantifies this weak dependence

Lemma 1. The correlation coefficient of the degree D_i and D_j of two random nodes i and j in $G_p(N)$ for $0 < p < 1$ is

$$c(D_i, D_j) = \frac{\text{Cov}[D_i, D_j]}{\sqrt{\text{Var}[D_i]}\sqrt{\text{Var}[D_j]}} = \frac{1}{N-1}.$$

Proof: see Appendix B. For large N and constant p , independent of N , the normalized i.i.d. binomially distributed sequence $\{D_i^*\}_{1 \leq i \leq N}$ of all degrees in $G_p(N)$ tends to be Gaussian distributed. The minimum of the sequence $\{D_i^*\}_{1 \leq i \leq N}$ possesses the distribution

$$\Pr\left[\min_{1 \leq i \leq N} D_i^* \leq x\right] = 1 - \prod_{i=1}^N \Pr[D_i^* > x] = 1 - (\Pr[D_i^* > x])^N.$$

After considering the limiting process of the minimum of a set $\{D_i^*\}_{1 \leq i \leq N}$ when $N \rightarrow \infty$, we derive (see Appendix C) the appropriate solution

$$D_{\min}^* = \frac{-Y - 2 \log N + \log \left(\sqrt{2\pi \log \frac{N^2}{2\pi}} \right)}{\sqrt{2 \log N}}$$

where Y is a Gumbel random variable [60]. With $D_{\min} = \sigma[D].D_{\min}^* + E[D] = \sqrt{(N-1)p(1-p)}.D_{\min}^* + p(N-1)$, we obtain

$$D_{\min} = p(N-1) - \sqrt{(N-1)p(1-p)} \left(\frac{Y + 2 \log N - \log \left(\sqrt{2\pi \log \frac{N^2}{2\pi}} \right)}{\sqrt{2 \log N}} \right).$$

Finally, let $D_{\min}(p)$ denote the minimum degree in $G_p(N)$. Since the complement of $G_p(N)$ is $G_{1-p}(N)$, there holds that

$$D_{\min}(p) = N - 1 - D_{\max}(1 - p).$$

The law of D_{\max} has been derived by Bollobas [10, Corollary 3.4 (p. 65)] via another method. Using the above relation, Bollobas' results precisely agrees with ours.

6.1.3 Analytical approximation for algebraic connectivity in random graph of Erdős-Rényi

In [34], Fiedler proved the upper bound on the algebraic connectivity in terms of the minimum nodal degree, i.e. for $G \neq K_N$, $\lambda_{N-1} \leq D_{\min}$. From this bound, our basic approximation that $\lambda_{N-1} \simeq D_{\min}$, for large N , follows. A comprehensive set of simulation results, presented in Subsection 6.1.4, supports the validity of this assumption. With this approximation we arrive, for large N , at

$$\begin{aligned} \lambda_{N-1} \simeq & p(N-1) - \sqrt{2p(1-p)(N-1)\log N} \\ & + \sqrt{\frac{(N-1)p(1-p)}{2\log N}} \log \left(\sqrt{2\pi \log \frac{N^2}{2\pi}} \right) \\ & - \sqrt{\frac{(N-1)p(1-p)}{2\log N}} Y. \end{aligned} \quad (6.6)$$

By taking the expectation on both sides and taking into account that the mean of a Gumbel random variable $E[Y] = \gamma = 0.5772\dots$, our estimate of the mean of the algebraic connectivity in $G_p(N)$ becomes, for large N and constant p ,

$$\begin{aligned} E[\lambda_{N-1}] \simeq & p(N-1) - \sqrt{2p(1-p)(N-1)\log N} \\ & + \sqrt{\frac{(N-1)p(1-p)}{2\log N}} \log \left(\sqrt{2\pi \log \frac{N^2}{2\pi}} \right) \\ & - \sqrt{\frac{(N-1)p(1-p)}{2\log N}} \gamma. \end{aligned} \quad (6.7)$$

Similarly, by taking into account that $Var[Y] = \frac{\pi^2}{6}$, the estimate of the variance of the algebraic connectivity in $G_p(N)$ is

$$Var[\lambda_{N-1}] = \frac{(N-1)p(1-p)}{2\log N} \frac{\pi^2}{6}. \quad (6.8)$$

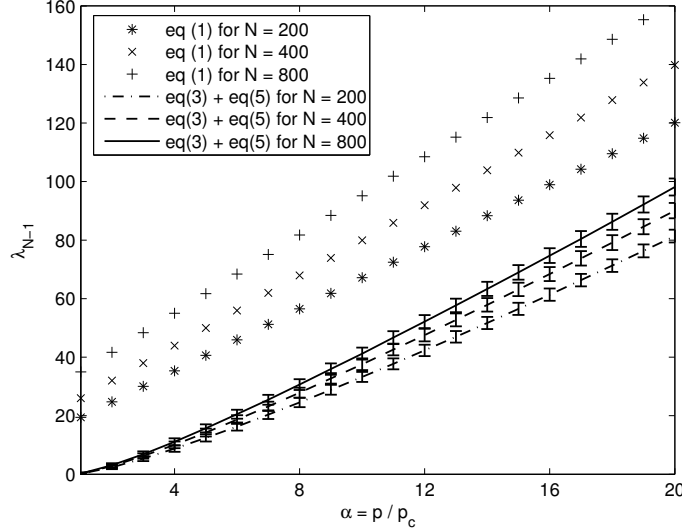


Figure 6.1: A comparison between the estimation of the mean (6.7) as well as the standard deviation (6.9), plotted in lines and error bars, and the theorem of Juhāz (6.5), plotted in markers, for the algebraic connectivity λ_{N-1} as a function of $\alpha = p/p_c$ in the Erdős-Rényi random graph $G_p(N)$ with $N = 200, 400$ and 800 nodes.

An interesting observation is that the standard deviation

$$\sigma[\lambda_{N-1}] = \sqrt{\text{Var}[\lambda_{N-1}]} = O\left(\sqrt{\frac{N}{\log N}}\right) \quad (6.9)$$

is much smaller than the mean (6.7). This implies that λ_{N-1} tends to the mean rapidly, or that, for large N , λ_{N-1} behaves almost deterministically and is closely approximated by the first three terms in (6.6). Hence, the relation (6.6) is more accurate than (6.5) (see also Figure 6.1).

6.1.4 Verification of analytical approximation

In all simulations we consider exclusively the Erdős-Rényi random graph $G_p(N)$ with various combinations of the number of nodes N and the link probabilities p . N takes the following values: 200, 400 and 800. The link probability is $p = \alpha p_c = \alpha \frac{\log N}{N}$, where α varies from 1 to 20. From each combination of N and p , we compute the algebraic connectivity λ_{N-1} and the minimum nodal degree D_{\min} . Then, we classify the simulated graphs according to their value of α , as shown in Figures 6.2 and 6.3. Subsequently, from generated graphs with a given α , we are interested in the extreme values, i.e. $\min \lambda_{N-1}$ and $\max \lambda_{N-1}$, as shown in Figures 6.4 and 6.5.

In Figures 6.2 and 6.3, we have plotted the simulated mean $E[\lambda_{N-1}]$, the corresponding standard deviation $\sigma[\lambda_{N-1}]$, and our estimate for the mean, Eq. (6.7), and the standard deviation, Eq. (6.9), of the algebraic connectivity as a function of α . As illustrated in Figures in Figures 6.2 and 6.3 there is a remarkable correspondence between the simulations and our estimate: the standard deviation is much smaller than the mean, implying that for $N \rightarrow \infty$, λ_{N-1} will rapidly approach $E[\lambda_{N-1}]$. Moreover, our basic approximation that, for large N , $\lambda_{N-1} \simeq D_{\min}$ is verified by the simulations shown in Figures 6.4 and 6.5. We found that $\min \lambda_{N-1}$ or $\max \lambda_{N-1}$ grows linearly with D_{\min} . Note in Figure 6.4 that, in the probability range around the connectivity threshold p_c , the minimum algebraic connectivity is always equal to zero, indicating a non-connected random graph (for details see Section 6.2).

From Figures 6.4 and 6.5 it is clear that if the value of the algebraic connectivity is larger than zero, the random graph has nodes of minimum degree always larger than zero too, referring to $\{\lambda_{N-1} > 0\} \iff \{G_p(N) \text{ is connected}\}$. However, by scrutinizing only degree-related simulation results, we see that the implication $\{D_{\min} \geq 1\} \implies \{G_p(N) \text{ is connected}\}$ is not always true, i.e. for large N and certain p which depends on N , the implication is almost surely (a.s.) correct [60]. For example, the percentage of graphs with $D_{\min} \geq 1$ that leads to a connected $G_p(50)$ increases from 98% for $p = p_c$ to 100% for $2p_c$, while the percentage for $G_p(400)$ increases from 99% for $p = p_c$ to 100% for $2p_c$. Hence, the simulation results confirm that, for large N and rather small $p = \frac{\log N}{N}$, the latter implication a.s. is equivalent.

Simulations demonstrate also that, for a particular fixed $\alpha = \frac{p}{p_c}$, the mean of the algebraic connectivity increases with the size of the random graph: a higher value of the graph size N implies a higher mean of the algebraic connectivity, which in turn indicates that the probability of having a more robust graph is approaching 1 as $N \rightarrow \infty$. The theorem given in Subsection 6.1.1, stating that $\Pr[\kappa_N = \kappa_L = D_{\min}] \rightarrow 1$ as $N \rightarrow \infty$, clarifies this observation in a slightly different way: given that N is approaching ∞ , the node and the link connectivity will become as high as possible, i.e. equal to the minimum nodal degree, and therefore the graph will become optimally connected.

6.2 Relationship between algebraic, node and link connectivity in random graph of Erdős-Rényi

It is well-known that, for any graph, the algebraic connectivity is at most equal to the node connectivity. In this section we therefore analyze the relationship among the three connectivity measures: the algebraic connectivity, the node connectivity and the link connectivity.

We have used the polynomial time algorithm, explained in [35], to find the node and the link connectivity by solving the maximum-flow problem. The maximum-flow

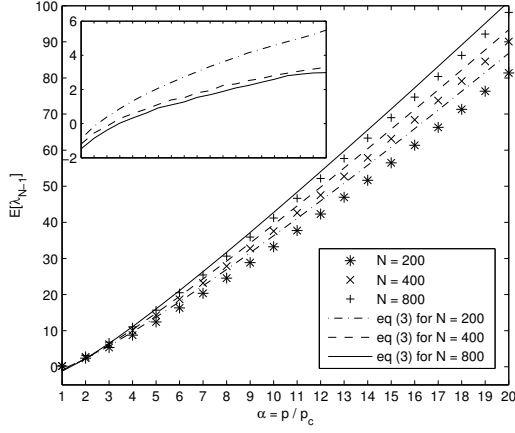


Figure 6.2: A comparison between the estimation (3), plotted in lines, and the simulation results, plotted in markers, for the mean of the algebraic connectivity $E[\lambda_{N-1}]$ as a function of α . In the upper left corner of the figure, we show the difference between the estimation and the simulation results as a function of α .

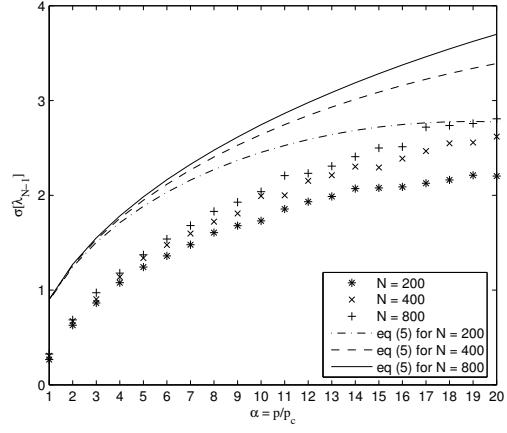


Figure 6.3: A comparison between the estimation (5), plotted in lines, and the simulation results, plotted in markers, for the standard deviation of the algebraic connectivity $\sigma[\lambda_{N-1}]$ as a function of α .

problem can be solved with several algorithms, e.g. Dinic, Edmonds & Karp, Goldberg, etc. If Goldberg's push-relabel algorithm is utilized, as performed in our simulations, the link connectivity algorithm has $O(N^3\sqrt{L})$ -complexity, while the node connectivity algorithm has $O(N^2L\sqrt{L})$ -complexity. We have used the LAPACK (Linear Algebra PACKage) implementation of the QR-algorithm for computing all the eigenvalues of the Laplacian matrix. For linear algebra problems involving the computation of a few extreme eigenvalues of large symmetric matrices, algorithms (e.g. Lanczos) whose run-time and storage cost is lower compared to the algorithms for calculation of all eigenvalues (QR algorithm has $O(N^3)$ -complexity) are known [3].

We simulate for each combination of N and p , 10^4 independent $G_p(N)$ graphs. N is 50, 100, 200 and 400 nodes and the link probability $p = \alpha p_c$, where α varies from 1 to 10. From each combination of N and p , we compute the minimum nodal degree D_{\min} , the algebraic, the node and the link connectivity, denoted respectively by λ_{N-1} , κ_N and κ_L . Then, we order graphs according to their value of α .

Figure 6.6 shows the mean value of the algebraic connectivity $E[\lambda_{N-1}]$ as a function of $\alpha = \frac{p}{p_c}$. In addition, Figure 6.6 shows the mean of the node connectivity $E[\kappa_N]$, the link connectivity $E[\kappa_L]$ and the minimum nodal degree $E[D_{\min}]$.

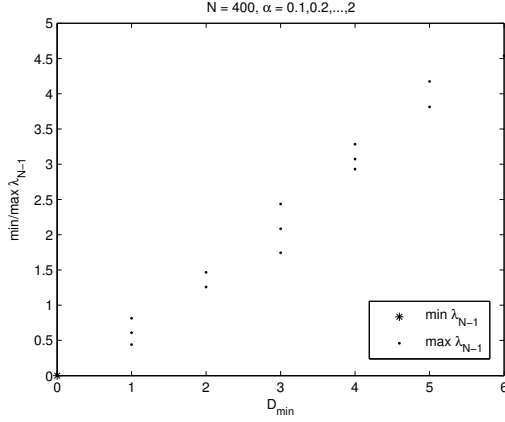


Figure 6.4: The relationship between the minimum algebraic connectivity $\min \lambda_{N-1}$ (maximum algebraic connectivity $\max \lambda_{N-1}$) and the minimum nodal degree D_{\min} in the Erdős-Rényi random graph $G_p(N)$. For each combination of N and $p = \alpha p_c$, $\alpha = 0.1, 0.2, 0.3, \dots, 2$, we generate 10^4 random graphs. Then, from the generated graphs, having a given value of α , we take $\min \lambda_{N-1}$ ($\max \lambda_{N-1}$) and the corresponding D_{\min} , resulting in one point for each considered α .

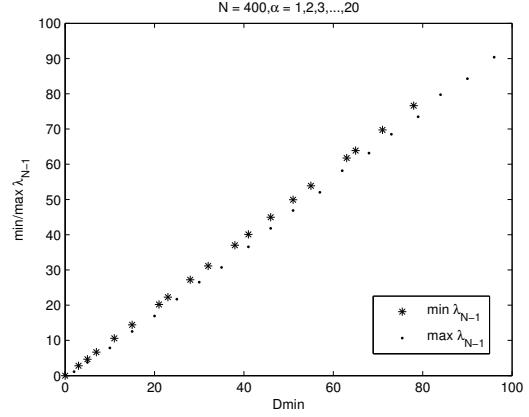


Figure 6.5: The relationship between the minimum algebraic connectivity $\min \lambda_{N-1}$ (maximum algebraic connectivity $\max \lambda_{N-1}$) and the minimum nodal degree D_{\min} in the Erdős-Rényi random graph $G_p(N)$. For each combination of N and $p = \alpha p_c$, $\alpha = 1, 2, 3, \dots, 20$, we generate 10^4 random graphs. Then, from the generated graphs, having a given value of α , we take $\min \lambda_{N-1}$ ($\max \lambda_{N-1}$) and the corresponding D_{\min} , resulting in one point for each considered α .

The first conclusion we can draw after analyzing simulation data is that for all generated random graphs from $p = p_c$ to $p = 10p_c$ the convergence to a surely connected random graph, i.e. $\lambda_{N-1} > 0$, is surprisingly rapid. Results concerning connectivity percentages are plotted in Figure 6.7. For example, the percentage of connected random graphs with 50 nodes increases from about 39% and 98% for p_c and $p = 2p_c$, respectively, to 99% for $p = 3p_c$, where for $p = 4p_c$ the graph is connected. These results are consistent with the Erdős-Rényi asymptotic expression. For $N \rightarrow \infty$, as observable in Figure 6.7, the simulated data as well as the Erdős-Rényi formula confirm a well known result [53] that the random graph $G_p(N)$ is a.s. disconnected if the link probability p is below the connectivity threshold $p_c \sim \frac{\log N}{N}$ and connected for $p > p_c$.

The second conclusion is that our results, regarding the distribution range of the algebraic connectivity and the minimum nodal degree for $G \neq K_N$, indeed comply with the bounds $0 \leq \lambda_{N-1} \leq D_{\min}$: the distribution of the algebraic connectivity λ_{N-1} is contained in the closed interval $[0, N]$, or to be more precise λ_{N-1} is 0 for a disconnected

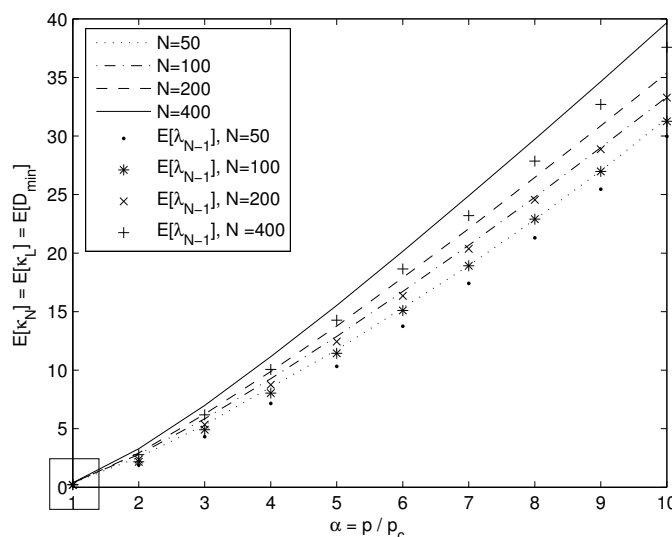


Figure 6.6: Simulated results on Erdős-Rényi random graph $G_p(N)$ for $N = 50, 100, 200$ and 400 and the link probability $p = \alpha p_c$, showing the mean of the algebraic connectivity $E[\lambda_{N-1}]$, the mean of the node $E[\kappa_N]$ and the link $E[\kappa_L]$ connectivity and the mean of the minimum nodal degree $E[D_{\min}]$ as a function of α , where $\alpha = 1, 2, \dots, 10$. Note that for $\alpha = 1$, $E[\kappa_N] = E[\kappa_L]$ but $E[\kappa_N] \neq E[D_{\min}]$.

graph and above bounded by $\frac{N}{N-1} D_{\min}$ for all those link probabilities p for which the graph is connected but not complete¹. Then, obviously $E[\lambda_{N-1}] \leq E[D_{\min}]$.

The third conclusion is that the distribution range of the algebraic connectivity also complies with the bounds $\lambda_{N-1} \leq \kappa_N$. Moreover, in Figure 6.6, for $p > p_c$ and all simulated N , the distributions of the node κ_N and the link κ_L connectivity are equal to the distribution of the minimum nodal degree D_{\min} (recall that in Figure 6.6 for $p = p_c$, the distributions of κ_N , κ_L and D_{\min} are almost equal but not the same). Convergence here to a graph where $\kappa_N = \kappa_L = D_{\min}$ is surprisingly rapid. For example, from the simulation results plotted in Figure 6.8 with $p = p_c$ and size of the random graph ranging from $N = 5$ to $N = 400$, we found that with probability approaching 1, the random graph becomes optimally connected at rather small graph sizes. For all other link probabilities, $p > p_c$, the convergence to $\kappa_N = \kappa_L = D_{\min}$ is faster (see Figure 6.8 for $p = 2p_c$).

Overall, the simulation results show that the random graph $G_p(N)$ a.s. is constructed in such a way that deleting all the neighbors (or the links to its neighbors) of a minimum nodal degree node will lead to the minimum number of nodes (links) whose deletion from a graph will result into a disconnected random graph. Hence, the

¹If a graph G is a complete graph K_N then $\lambda_{N-1} = N > D_{\min} = N - 1$.

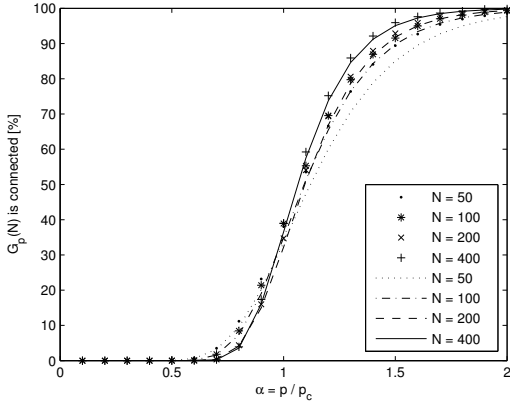


Figure 6.7: Percentage of the connected Erdős-Rényi random graphs $G_p(N)$, i.e. $\lambda_{N-1} > 0$: a comparison between the simulation results, plotted in markers, and the Erdos' asymptotic formula $\Pr[G_p(N) = \text{connected}] \simeq e^{-Ne^{-p(N-1)}}$, plotted in lines, in the probability range $\alpha = p/p_c = 0.1, 0.2, \dots, 2$.

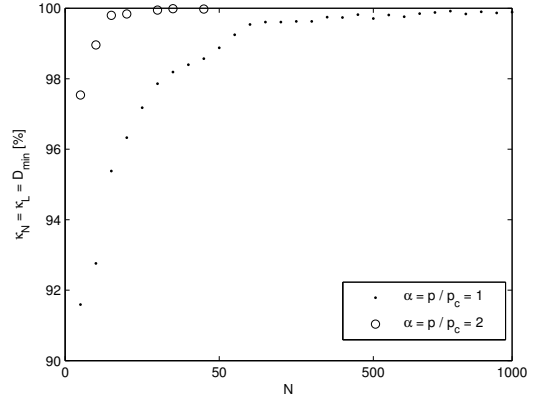


Figure 6.8: Percentage of the Erdős-Rényi random graphs $G_p(N)$ with $p = p_c$ and $p = 2p_c$ in which the node connectivity κ_N , the link connectivity κ_L and the minimum nodal degree D_{\min} converge to $\kappa_N = \kappa_L = D_{\min}$ for small graph sizes N .

minimum nodal degree is a valuable estimate of the number of nodes or links whose deletion results into a disconnected graph.

6.3 Conclusion

In this chapter we have improved an already existing theorem concerning the behavior of the algebraic connectivity in the Erdős-Rényi random graph. Through extensive simulations we have verified the accuracy of the analytical estimation, also for networks having small sizes. Simulations have also shown that for large graph size, the distribution of the algebraic connectivity grows linearly with the minimum nodal degree. In fact, the larger the graph size, the more the Erdős-Rényi random graph is constructed in such a way that deleting all the neighbors (or the links to its neighbors) of a minimum nodal degree node leads to the minimum number of nodes (links) whose deletion disconnects the graph. Moreover, simulation results have shown that the equality between the minimum nodal degree, on one hand, and the node and the link connectivity, on the other hand, occurs regardless of the link probability at already small graph sizes. This, so called optimal connectivity, directly implies that a larger value of the algebraic connectivity quantifies a higher robustness to node and link failures.

Chapter 7

Spectral Radius of Complex Networks

In this chapter we rely on the largest eigenvalue of the adjacency matrix to quantitatively characterize the robustness to virus propagation in complex networks [45]. The largest eigenvalue of the adjacency matrix is often referred to as the spectral radius. Recently it has been shown [79] that the spectral radius plays an important role in modelling virus propagation in networks. In fact, in [79] it was shown that the existence of an epidemic threshold [23] is inversely proportional to the spectral radius of the adjacency matrix. It follows from this result that the smaller the spectral radius, the higher the robustness of a network against the spread of viruses. The contribution of the study presented in this chapter is twofold. First, we study how well-known upper bounds for the spectral radius of graphs match the spectral radii of a number of real-world networks. Second, we compare the spectral radius of these real-world networks with those of commonly used network models.

7.1 Upper bounds for spectral radius

An upper bound for the spectral radius of a graph [79] gives a lower bound for the epidemic threshold for virus propagation in networks. If the effective spreading rate is below this lower bound, then the virus contamination dies out. The sharper the upper bound for the spectral radius, the less effort we need to spend in reducing the effective spreading rate below the lower bound. The effective spreading rate can be lowered by either decreasing the spreading rate β (e.g. by implementing more or better intrusion detection/prevention software) or by increasing the cure rate δ (e.g. by installing more virus scanning software).

The most common graphs for which an explicit expression for the spectral radius is known, are [21]: the complete graph K_N , the path P_N , the ring graph R_N , the k -regular

graph, the k -dimensional lattice and the complete bipartite graph $K_{N,N}$. Since no closed expression is known for the spectral radius of a general graph, we will discuss a number of upper bounds. Although many more bounds are known, the most important ones are presented. The bounds differ both in form as in the parameters that are used. The more information about a graph is used, the better the bounds can be. Since different parameters of graphs are used in the bounds, it seems computationally intensive to compare them in general.

Now, let G be a graph on N nodes and L links, with minimum degree D_{\min} , maximum degree D_{\max} , and spectral radius $\rho(G)$. The oldest and simplest bound, that can be found in any book on spectral graph theory, is

$$\rho(G) \leq D_{\max} . \quad (7.1)$$

A bound in terms of the numbers of nodes and links only is found by Hong [43]: if G is connected, then

$$\rho(G) \leq \sqrt{2L - N + 1} . \quad (7.2)$$

Cao [16] improved this bound at the cost of using more parameters: if $D_{\min} \geq 1$, then

$$\rho(G) \leq \sqrt{2L - (N - 1)D_{\min} + (D_{\min} - 1)D_{\max}} . \quad (7.3)$$

Hong, Shu, and Fang [44] obtained a bound that indicates the relation of the spectral radius to the minimal degree: if G is connected, then

$$\rho(G) \leq \frac{1}{2} \left[D_{\min} - 1 + \sqrt{(D_{\min} + 1)^2 + 4(2L - Nd_{\min})} \right] . \quad (7.4)$$

Das and Kumar [24] obtained a bound that uses very local information of the graph: if G is connected, and D_i is the average degree of the nodes adjacent to node i , then

$$\rho(G) \leq \max \left\{ \sqrt{D_i D_j} : i \sim j \right\} . \quad (7.5)$$

Here $i \sim j$ indicates that nodes i and j are linked. Finally, Cioabă, Gregory, and Nikiforov [18] obtained an upper bound that also uses the diameter $diam$ of the graph: if G is connected and nonregular, then

$$\rho(G) < D_{\max} - \frac{Nd_{\max} - 2L}{N(diam(Nd_{\max} - 2L) + 1)} . \quad (7.6)$$

<i>topological characteristics</i>	A	B	C	D
N	685	14098	4058	18121
L	10271	18689	6151	59507
D_{\max}	117	6	107	2404
D_{\min}	9	1	2	1
$E[D]$	30	2.7	3	6.2
$diam$	11	255	34	8
ρ	50.7	3.5	14.2	110.8

Table 7.1: Topological characteristics of the real-world network of Dutch soccer team players (**A**), Dutch roadmap (**B**), IP-level (**C**) and AS-level (**D**) Internet graph. The notation was introduced in the previous section, with the exception of $E[D]$ which denotes average nodal degree.

<i>upper bound</i>	A	B	C	D
ρ	50.7	3.5	14.2	110.8
7.1	117	6	107	2404
7.2	140.9	152.6	90.8	317.6
7.3	123.8	152.6	65.5	317.6
7.2	120.0	152.6	65.2	317.6
7.5	66.6	4.2	17.1	1252.5
7.6	117	6	107	2404

Table 7.2: Upper bounds on the spectral radius of the real-world network of Dutch soccer team players (**A**), Dutch roadmap network (**B**), IP-level (**C**) and AS-level (**D**) Internet graph.

7.2 Spectral radius of empirical networks

In this section, we give the spectral radius ρ for the following real-world networks: the social network that is formed by all soccer players that have played an international match for the Dutch soccer team (**A**) [27], the Dutch roadmap network (**B**) [52], the network of the observable part of the Internet graph at the IP-level (**C**) [46] and the Autonomous System level (**D**) [65]. In this social network of Dutch soccer team players every node corresponds to a soccer player that has played a game for the Dutch national team. A node is connected with another node if both players have appeared in the same match.

For the considered real-world networks, we show in Table 7.1, a set of graph measures. Furthermore, Table 7.2 illustrates the tightness of the upper bounds introduced

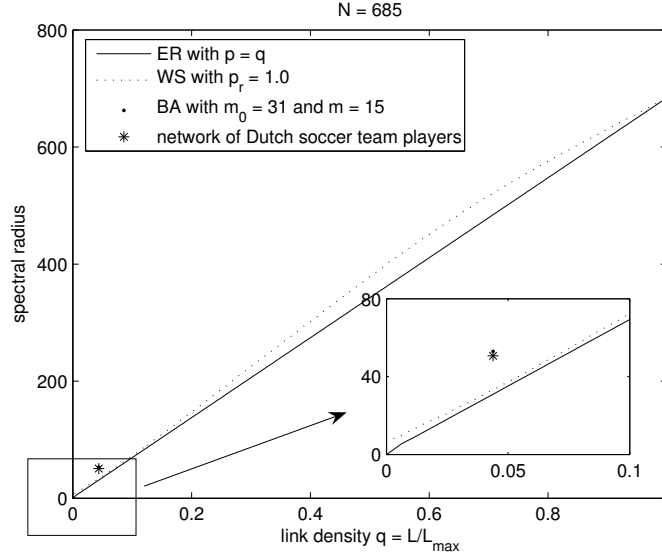


Figure 7.1: Comparison between the spectral radius of the real-world network of Dutch soccer team players and the spectral radius of networks models: the random graph of Erdős-Rényi, the small-world graph of Watts-Strogatz and the scale-free graph of Barabási-Albert.

in Section 7.1. We conclude from Table 7.2 that for networks **A**, **B** and **C** the upper bound 7.5 of Das and Kumar exhibits the best match with the real spectral radius. The overestimation of this upper bound for networks **A**, **B** and **C** is 31%, 20% and 20%, respectively. For network **D**, the upper bounds 7.2, 7.3 and 7.4 provide the best match. However, in this case, the overestimation is as high as 187%. In addition, the following observations can be made from Table 7.2:

- The upper bounds 7.2, 7.3 and 7.4 are of the same order.
- The upper bound 7.6 does not give an improvement of upper bound 7.1.
- For networks **A** and **B**, the simple upper bound 7.1 outperforms the upper bounds 7.2, 7.3 and 7.4.

Next, we explore the spectral radius of generic models (see e.g. [2] and [68]), used for modelling the evolution and the topology of real-world networks, i.e. the random graph of Erdős-Rényi (**ER**) [10], the small-world graph of Watts-Strogatz (**WS**) [80] and the scale-free graph of Barabási-Albert (**BA**) [5].

The main parameters of the Erdős-Rényi graph $G_p(N)$, as explained in Section 2.2, are the number of nodes N and the existence of a link with the probability p , which is independent from the existence of other links. Then, the total number of links in $G_p(N)$

<i>spectral radius</i>	A	B	C	D
ρ	50.7	3.5	14.2	110.8
ρ_{ER}	30.9	4.1	4.4	7.7
ρ_{WS}	32.9	5.8	5.2	10.3
ρ_{BA}	52.9	14.2	14.3	15.1

Table 7.3: The spectral radius of network models: the random graph of Erdős-Rényi, the small-world of Watts-Strogatz and the scale-free graph of Barabási-Albert. For the simulations of the spectral radius the same number of nodes and links is used as in the real-world networks under consideration.

is on average equal to pL_{\max} , where $L_{\max} = \binom{N}{2}$ is the maximum possible number of links. Hence, the link density $q = \frac{L}{L_{\max}}$ equals p .

The Watts-Strogatz graph captures the fact that, despite the large size of the topology, in most real-world networks, there is a relatively short path between any two nodes. The diameter $diam$, presented in Table 7.1, aims to illustrate this effect. The main parameters of the small-world graph of Watts-Strogatz, as seen in Section 2.3, are the number of nodes N , the number of clockwise neighbors a node has s , and the probability p_r a link is rewired to a new location chosen uniformly in the ring lattice $R_{N,s}$. The number of links L in the Watts-Strogatz graph, irrespective of p_r , is always equal to $L = Nk$. Hence, the link density is $q = \frac{2k}{N-1}$.

The Barabási-Albert graph gives rise to a class of graphs with a power-law degree distribution. The Barabási-Albert graph is based on the following parameters: the number of fully-meshed nodes m_0 at the beginning of the construction of a graph, the number of new attached nodes t and the number of links m a newly attached node has. Then, the number of nodes N in the Barabási-Albert graph is $N = t + m_0$ and the number of links L is $L = \frac{m_0(m_0-1)}{2} + mt$. Hence, the link density is $q = \frac{m_0(m_0-1)+2mt}{N(N-1)}$.

Figure 7.1 compares $\rho(\mathbf{A})$ and the average value of ρ for generic network models. Furthermore, in Table 7.3 we consider ρ , calculated for a graph of identical link density as the one in real-world networks under consideration. Figure 7.1 illustrates that the value of ρ_{BA} is closest to $\rho(\mathbf{A})$. Moreover, as shown in Table 7.3, the same tendency is observed for network **C**, whereas for network **D** the value of $\rho(\mathbf{D})$ is not consistent with any of the examined models. Finally, the Dutch road infrastructure (network **B**) is most likely a subgraph of a two-dimensional¹ lattice graph, as found in [52].

¹The spectral radius of the 2D-lattice with sizes z_1 and z_2 such that $N = (z_1 + 1)(z_2 + 1)$ and $L = 2z_1z_2 + (z_1 + z_2)$ is

$$\rho_{2D\text{-Lattice}} = 2 \cos\left(\frac{\pi}{z_1 + 2}\right) + 2 \cos\left(\frac{\pi}{z_2 + 2}\right) < 4$$

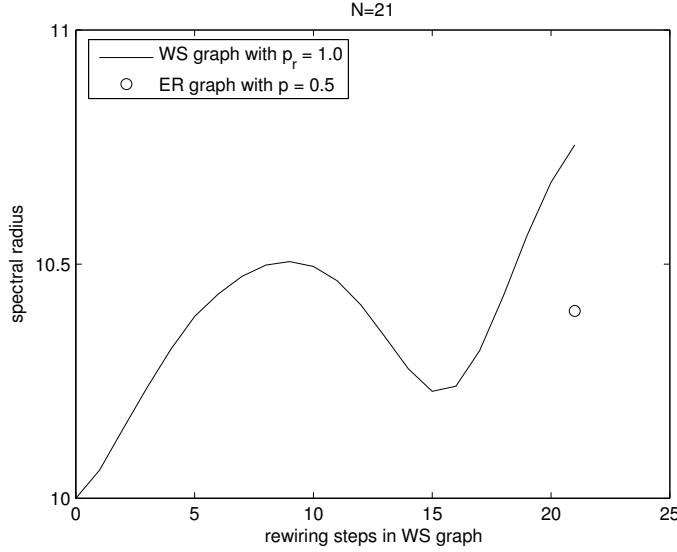


Figure 7.2: The spectral radius of the small-world graph of Watts-Strogatz after, for every node, each link connected to a clockwise neighbor (i.e. here $k = 5$ links) is rewired to a randomly chosen node with the probability $p_r = 1.0$. At the end of N steps, which correspond to the N nodes, the spectral radius is compared with the one of the random graph of Erdős-Rényi, both for the link density $q = 0.5$.

From the simulation results, we saw for small and medium rewiring probabilities p_r that the spectral radius ρ_{WS} virtually corresponds to ρ_{ER} . For example, the difference in ρ between the Watts-Strogatz graph with $p_r = 0.5$ and the Erdős-Rényi graph, in the complete q -range, is hardly noticeable (not shown). However, for high rewiring probabilities p_r (close to 1), ρ_{WS} has the tendency not to converge to ρ_{ER} , see Figure 7.1. Thus, the spectral properties of the Watts-Strogatz graph with $p_r = 1$ are not identical to those of the Erdős-Rényi graph. In order to examine this unexpected behavior in more detail, we have conducted some additional simulations on graphs with a small number of nodes, i.e. $N = 21$ nodes.

To obtain the Watts-Strogatz graph with the link density $q = 0.5$ and $p_r = 1$, we move $k = 5$ links of a given node to a new location chosen randomly in the ring lattice. The resulting spectral radius, after each rewiring step, is depicted in Figure 7.2. In addition, Figure 7.2 contains the spectral radius of the Erdős-Rényi graph with the same link density. The depicted values are obtained by averaging over 100 simulation runs. It is obvious that ρ_{WS} of the Watts-Strogatz graph (obtained after 21 rewiring steps) is larger than ρ_{ER} of the Erdős-Rényi graph. A possible explanation lies in the fact that for $p_r = 1.0$, the Watts-Strogatz graph is an approximation of the random graph with the constraint that each node has a minimum of $D_{\min} = k$ links. This

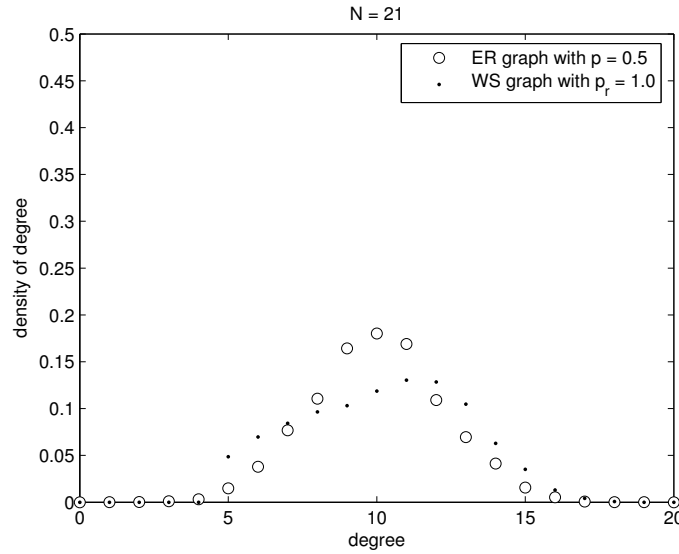


Figure 7.3: The degree distribution of the small-world graph of Watts-Strogatz and the random graph of Erdős-Rényi, both for the link density $q = 0.5$.

implies that the degree distribution of the Watts-Strogatz graph is skewed. Figure 7.3 shows clearly that the degree distribution of the Erdős-Rényi graph is not identical to the one of the Watts-Strogatz graph, where indeed $D_{\min} = k$.

7.3 Conclusions

In this chapter we have introduced the spectral radius as a quantifier of the robustness to virus propagation in complex networks. We have modelled the topology of a complex network as a graph and studied how well-known upper bounds for the spectral radius match the spectral radii of a number of empirical networks. In this context, our main conclusions are:

- For the Dutch soccer team network, the Dutch road map network and the Internet graph at router level, the upper bound given in [24] is reasonably tight.
- For the Internet graph on AS-level, all considered upper bounds seriously overestimate the spectral radius.

Furthermore, we have compared the spectral radius of a number of empirical networks to commonly used theoretical graph models. Conclusion regarding this research question are the following:

- The spectral radii of the Dutch soccer team network and the Internet graph at the router level match well those obtained from scale-free Barabási-Albert graphs with the same link density.
- The spectral radius of the Dutch road map network matches well that of a two dimensional lattice.
- All considered network models, i.e. the random graph, small-world and scale-free graph give spectral radii that are much smaller than that of the Internet graph at AS-level.

Chapter 8

Relationship Between Algebraic Connectivity and Robustness of Complex Networks to Node and Link Failures

In this chapter we study the algebraic connectivity in relation to the network's robustness to node and link failures [49]. Network's robustness to failures is quantified with the node and the link connectivity, two graph measures that give the number of nodes and links that have to be removed in order to disconnect a graph. From [34] it is known that the algebraic connectivity is a lower bound on both the node and the link connectivity. As mentioned in [10], for every node or link connectivity there are infinitely many graphs for which the algebraic connectivity is not a sharp lower bound. It is thus worth looking into the relationship between the proposed measures. The chapter is organized as follows. The first section gives the theoretical background of Fiedler's algebraic connectivity and the two connectivity measures, the node and the link connectivity. The second section provides a comprehensive set of simulation results on the relation between the algebraic connectivity and the node and link connectivity in three well-known complex network models: the random graph of Erdős-Rényi, the small-world graph of Watts-Strogats, the scale-free graph of Barabási-Albert. Through extensive simulations with the three complex network models, we show that the algebraic connectivity is not trivially connected to networks's robustness to node and link failures. Furthermore, we show that the tightness of this lower bound is very dependent on the considered complex network model.

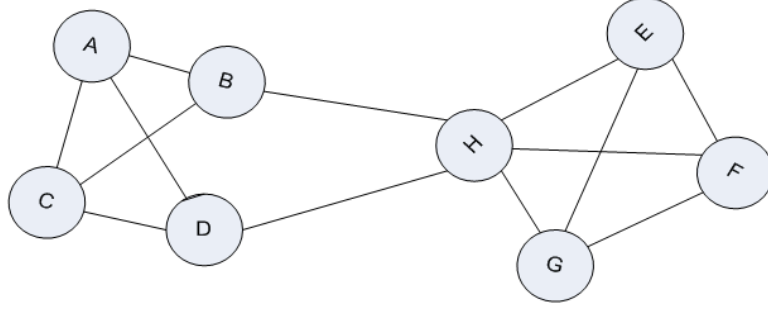


Figure 8.1: A graph with $N = 8$ nodes and $L = 13$ links. The graph's connectivity characteristics are: the node connectivity is 1 (removal of node H), the link connectivity is 2 (removal of links connecting node H to nodes B and D), the algebraic connectivity is 0.6277 and the minimum nodal degree is 3 (minimum number of links a node has).

8.1 Relationship between algebraic connectivity and classical connectivity measures

The second smallest Laplacian eigenvalue λ_{N-1} , as mentioned earlier, is an important measure in the analysis of the robustness of complex networks: 1) the algebraic connectivity is only equal to zero if a graph G is disconnected, 2) the multiplicity of zero as a second smallest Laplacian eigenvalue is equal to the number of disconnected components of G . Then, the node connectivity of an incomplete graph G is at least as large as the algebraic connectivity $\lambda_{N-1} \leq \kappa_N$ [34]. If $G = K_N$ then $\lambda_{N-1}(K_N) = N > \kappa_N(K_N) = N - 1$. The minimum nodal degree D_{\min} of an incomplete graph G is an upper bound on both the node and the link connectivity $\kappa_N \leq \kappa_L \leq D_{\min}$. If G is a complete graph K_N then $\kappa_N = \kappa_L = D_{\min}$.

As shown in Figure 8.1, the relationship between the introduced measures is not trivial: $\lambda_{N-1} = 0.6277 \leq \kappa_N = 1 \leq \kappa_L = 2 \leq D_{\min} = 3$. Accordingly, the minimal number of nodes κ_N and the minimal number of links κ_L to be removed such that no path between any two pairs of nodes remains, in this graph is respectively 1 and 2. Hence, the graph has 1 node-disjoint and 2 link-disjoint paths. This also means that the depicted graph is 1-node and 2-link connected.

We have used the polynomial time algorithm, explained in [35], to find the node and the link connectivity by solving the maximum-flow problem. The maximum-flow problem can be solved with several algorithms, e.g. Edmonds & Karp [28], Dinic [25], Goldberg [37], etc. If Goldberg's push-relabel algorithm is utilized, as performed in our simulations, the link connectivity algorithm has $O(N^3\sqrt{L})$ -complexity, while the node connectivity algorithm has $O(N^2L\sqrt{L})$ -complexity. We have used the LAPACK implementation of the QR-algorithm for computing all the eigenvalues of the Laplacian matrix. For linear algebra problems involving computation of few extreme eigenvalues

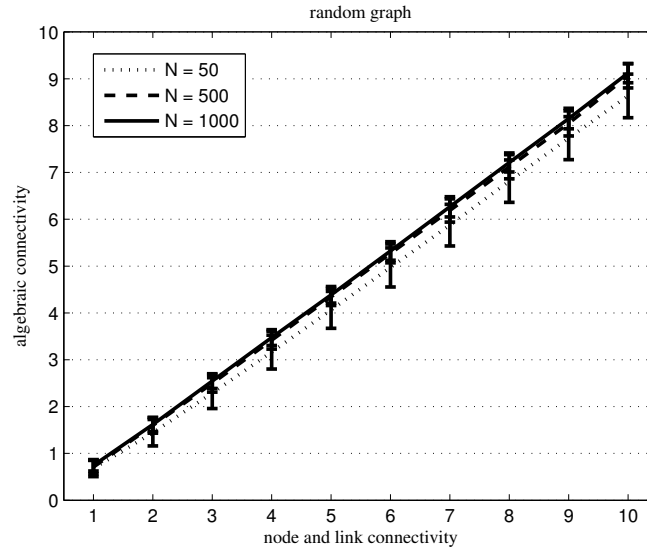


Figure 8.2: The mean as well as the standard deviation (error bars) of the algebraic connectivity μ_{N-1} as a function of the node connectivity κ_N and the link connectivity κ_N in the random graph of Erdős-Rényi with $N = 50, 500$ and 1000 nodes.

of large symmetric matrices, algorithms (e.g. Lanczos) whose run-time and storage cost is lower compared to the algorithms for calculation of all eigenvalues (QR algorithm has $O(N^3)$ -complexity) are known [3].

8.2 Relationship between algebraic connectivity and classical connectivity measures in complex network models

In this section, we present a comprehensive set of simulation results on the relation between the algebraic connectivity and the two connectivity measures in generic complex network models: the random graph of Erdős-Rényi, the small-world of Watts-Strogatz and scale-free graph of Barabási-Albert. Prior to relationship analysis, we define and briefly discuss the models.

8.2.1 Random graph of Erdős-Rényi

In Section 2.2 we have seen that in the Erdős-Rényi random graph $G_p(N)$ the parameters of interest are the number of nodes N and the probability p of having a link between

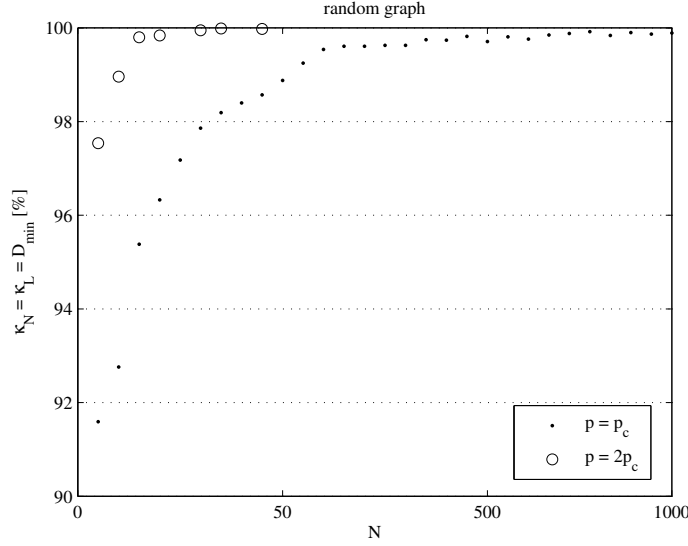


Figure 8.3: Percentage of the Erdős-Rényi random graphs with $p = p_c$ and $p = 2p_c$ for which the node connectivity κ_N , the link connectivity κ_L and the minimum nodal degree D_{\min} converge to $\kappa_N = \kappa_L = D_{\min}$.

any two nodes. We simulate for each combination of N and p , 10^3 independent $G_p(N)$ graphs. N is 50, 500 and 1000 nodes and the link probability $p = \alpha p_c$, where $p_c = \frac{\log N}{N}$ and α varies from 1 to 10. From each combination of N and p , we compute the node connectivity κ_N , the link connectivity κ_L and the algebraic connectivity λ_{N-1} . Then, we classify the simulated graphs according to their value of the node and the link connectivity, which for most graphs have the same value $\kappa_N = \kappa_L$. Thus, in Figure 8.2, the mean value (and standard deviation) of the algebraic connectivity is given as a function of both the node and the link connectivity.

The first observation from Figure 8.2 is that there seems to be a linear relationship between the mean of the algebraic connectivity and the node and the link connectivity. However, from this linear behavior alone it is not clear whether and how fast the algebraic connectivity converges towards the node and the link connectivity.

In [9] Bollobás proved that, irrespective of the link probability p , the probability that $\kappa_N = \kappa_L = D_{\min}$ approaches 1 as $N \rightarrow \infty$. Recall that D_{\min} is an upper bound on both κ_N and κ_L . From Figure 8.3 we observe that the convergence of $G_p(N)$ to a graph where $\kappa_N = \kappa_L = D_{\min}$ is fast. For example, from the simulation results plotted in Figure 8.3 with $p = p_c$ and a size of the random graph ranging from $N = 50$ to $N = 1000$, we observe that with probability approaching 1, $G_p(N)$ has $\kappa_N = \kappa_L = D_{\min}$ for rather small graph sizes. For all other link probabilities, i.e. $p > p_c$, the convergence to

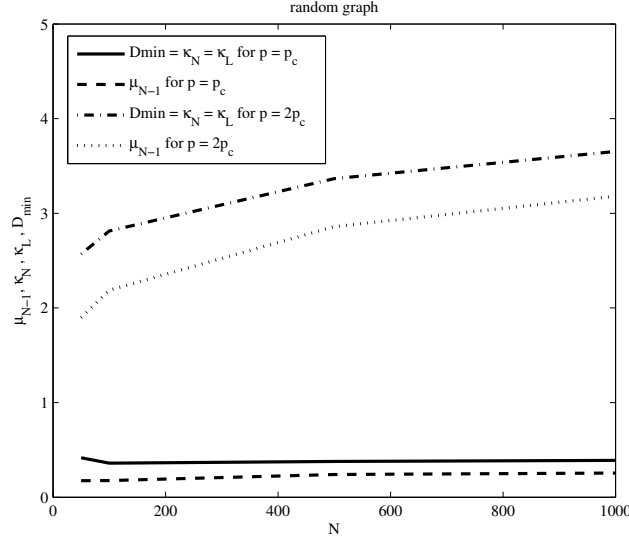


Figure 8.4: The mean of the node connectivity κ_N , the link connectivity κ_L and the algebraic connectivity μ_{N-1} as a function of the number of nodes N in the Erdős-Rényi random graph with a given link density $q = \frac{L}{L_{\max}} = p$, i.e. $p = p_c$ and $p = 2p_c$. Also, the minimum nodal degree D_{\min} is depicted as a function of N .

$\kappa_N = \kappa_L = D_{\min}$ occurs for even smaller values of N (see Figure 8.3 for $p = 2p_c$). This makes D_{\min} a valuable estimate of the minimum number of nodes or links whose deletion results into a disconnected Erdős-Rényi random graph.

Contrary to the convergence of D_{\min} , Figure 8.4 shows that as N tends to large values, the value of λ_{N-1} does not converge towards κ_N or κ_L , and obviously not to D_{\min} . Furthermore, for a given link density q , the difference between λ_{N-1} and κ_N or κ_L is considerable and becomes even more evident if we consider higher values¹ of the link density q (see Figure 8.4 for $q = p_c$ and $q = 2p_c$). This behavior is at odds with the one of D_{\min} .

8.2.2 Small-World graph of Watts-Strogatz

In Section 2.3 we have seen that in the small-world graph of Watts-Strogatz the parameters of interest are the number of nodes N , the number of clockwise neighbors a node has s and the probability p_r a link is rewired to a new location chosen uniformly in the ring lattice $R_{N,s}$. We simulate for each combination of N and s , 10^3 independent Watts-Strogatz small-world graphs. N is 50, 100 and 500 nodes and s varies from 1 to

¹In the Erdős-Rényi random graph, the link density q equals the link probability p .

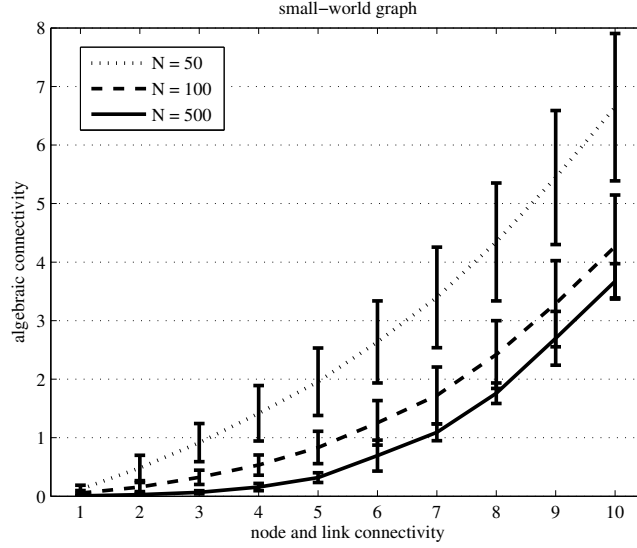


Figure 8.5: The mean as well as the standard deviation (error bars) of the algebraic connectivity as a function of the node and the link connectivity in the small-world graph of Watts-Strogatz with $N = 50, 100$ and 500 nodes.

10. The rewiring probability is set to be $p_r = 1$. From each combination of N and s , we compute the node connectivity κ_N , the link connectivity κ_L and the algebraic connectivity λ_{N-1} . Then, we classify the simulated graphs according to their value of κ_N and κ_L . Similarly to Figure 8.2, we plot in Figure 8.5 the mean (and standard deviation) of the algebraic connectivity as a function of the node and the link connectivity. In most simulated small-world graphs, we observe that $\kappa_N = \kappa_L$. Hence, the curve depicting the mean (and standard deviation) as a function of the node connectivity turns out to be indistinguishable from the curve for the link connectivity.

From Figure 8.5 we observe that the algebraic connectivity is a very loose lower bound on the node or the link connectivity. Moreover, the larger the graph size N , the looser the bound becomes. This means that for a given value of the node or the link connectivity, the mean value of the algebraic connectivity is a decreasing function of the graph size N , opposite to Erdős-Rényi random graph (see Figure 8.2). Furthermore, the larger the graph size N , the smaller the standard deviation (see error bars in Figure 8.5).

In Figure 8.5, for a given value of the node or the link connectivity, the algebraic connectivity λ_{N-1} seemed to be a decreasing function of N . However, Figure 8.6 shows that for small-world graphs with a given link density² q , λ_{N-1} is an increasing function

²In the Watts-Strogatz small-world graph the link density is $q = \frac{L}{L_{\max}} = \frac{sN}{L_{\max}} = \frac{2s}{(N-1)}$.

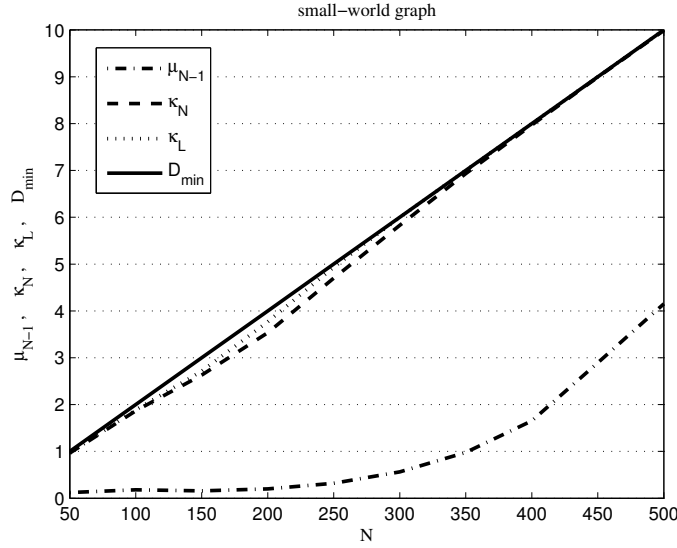


Figure 8.6: The mean of the node connectivity κ_N , the link connectivity κ_L and the algebraic connectivity μ_{N-1} as function of the number of nodes N in the Watts-Strogatz small-world graph with a given value of the link density $q = \frac{L}{L_{\max}} = \frac{sN}{L_{\max}} = 0.04$. The minimum nodal degree D_{\min} equals the number of neighbors s in the ring lattice.

of N . Figure 8.6 also shows that κ_N and κ_L approach D_{\min} for already small N . Thus, similarly to Erdős-Rényi random graph, the minimum number of nodes is a valuable estimate of the minimum number of nodes or links whose deletion results into a disconnected Watts-Strogatz small-world graph. Moreover, the larger the number of neighbors s in the ring lattice (on which the small-world graph is built), the larger the difference between λ_{N-1} and κ_N or κ_L . Recall that in the Watts-Strogatz small-world, each node has a minimum of $D_{\min} = s$ links. Hence, the algebraic connectivity is indicating that as the graph size N increases, the underlying topology of this small-world graph converges to a more robust structure: by expanding N and reducing the link density q , it might be possible to increase the number of nodes or link failures and still get the same value of the algebraic connectivity.

8.2.3 Scale-Free graph of Barabási-Albert

In Section 2.4 we have seen that in the scale-free graph of Barabási-Albert the parameters of interest are the number of nodes N and the number of links a newly attached node has m . We simulate for each combination of N and m , 10^3 independent Barabási-Albert scale-free graphs. N is 50, 100 and 500 nodes and m varying from 1 to 10. In the same way as in the simulations for the Erdős-Rényi random graph and the Watts-

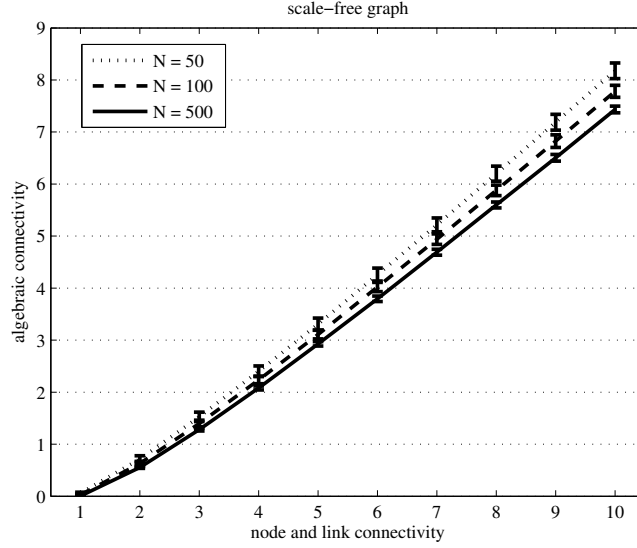


Figure 8.7: The mean as well as the standard deviation (error bars) of the algebraic connectivity as a function of the node and the link connectivity in the scale-free graph of Barabási-Albert with $N = 50, 100$ and 500 nodes.

Strogatz small-world, we compute for each combination of N and m , the algebraic connectivity and the two connectivity metrics. Figure 8.7 shows the mean of the algebraic connectivity λ_{N-1} , obtained by classifying the generated graphs according to their value of the node connectivity κ_N and the link connectivity κ_L . The scale-free graph is constructed in such a way that deleting m links or m nodes to which a new node (in the last time step) is attached, leads to $m = \kappa_N = \kappa_L$. The convergence to a graph where $\kappa_N = D_{\min}$ is observed for all combinations of N and m . Hence, the mean (and standard deviation) of the algebraic connectivity as a function of the node connectivity is identical to the mean obtained for the link connectivity.

Similarly to the Watts-Strogatz small-world graph, Figure 8.7 shows that the mean value of the algebraic connectivity is a decreasing function of the graph size N . However, the algebraic connectivity of scale-free graphs with a given link density q , is an increasing function of N . Recall that for all combinations of N and m , $m = \kappa_N = D_{\min}$. Hence, the link density q , for each N and m , equals $q = \frac{L}{L_{\max}} \approx \frac{m_0^2 + 2mt}{N^2}$. For example, in Figure 8.7, the following combinations of N and m have approximately the same value of q while the algebraic connectivity is an increasing function of N : for $N = 50$ and $m = 9$, $q = 0.37$ and for $N = 500$ and $m = 10$, $q = 0.4$. Thus, the algebraic connectivity indicates that as the link density q increases, the underlying topology of this scale-free graph converges to a more robust structure. It also indicates that by expanding N and

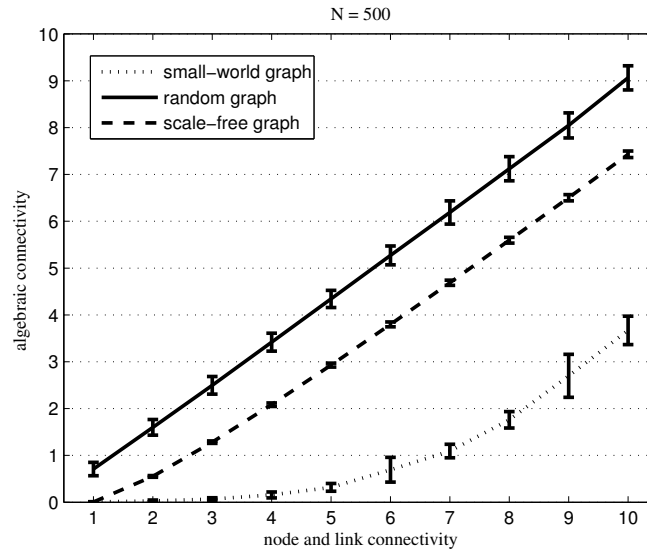


Figure 8.8: The mean as well as the standard deviation (error bars) of the algebraic connectivity as a function of the node and the link connectivity in the random graph of Erdős-Rényi, the small-world of Watts-Strogatz and scale-free graph of Barabási-Albert. All graphs have $N = 500$ nodes.

reducing the link density q , it might be possible to increase the number of nodes or link failures and still get the same value of the algebraic connectivity.

8.2.4 Comparison between complex network models

Figure 8.8 shows the mean as well as the standard deviation (error bars) of the algebraic connectivity as a function of the node and the link connectivity in the considered complex network models. Although the scale-free graph of Barabási-Albert has different topological properties, at least in terms of the degree distribution, Figure 8.8 shows that the relation between the algebraic connectivity and graph's robustness to node and link failures is similar to that in the Erdős-Rényi random graph. This similarity most probably comes from the fact that for both complex network models, the minimum nodal degree is a tight upper bound on the algebraic connectivity, explaining the almost linear relationship between the two connectivity metrics. Recall that we choose the network model parameters so as to perform the simulations within a link density range, which on average results in graphs with a comparable number of links. As shown in Figure 8.9 the small-world and the scale-free graph, both with a given node or link connectivity, on average have the same number of links. However, having the same number of nodes and links the Barabási-Albert scale-free graph seems to be more robust

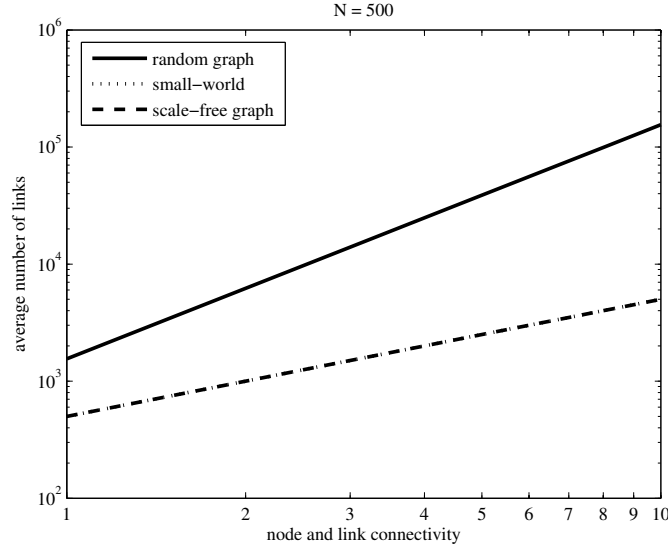


Figure 8.9: The average number of links L a considered complex network model has as a function of the node and the link connectivity. For the random graph of Erdős-Rényi, the number of links in a graph with a given node or link connectivity is on average equal to $L = pL_{\max}$, for the small-world graph of Watts-Strogatz $L = \frac{2s}{(N-1)}L_{\max}$ and for the scale-free graph of Barabási-Albert $L = \frac{m_0^2 + 2mt}{N^2}L_{\max}$.

than the Watts-Strogatz small-world graph. For the Erdős-Rényi random graph, the simulations (within a higher link density range than what is used for the other two models) give rise to the larger number of links and therefore possibly the larger value of the algebraic connectivity. Consequently, from the viewpoint of the node and the link connectivity, the robustness of the Erdős-Rényi random graph is worse than the one of the other two complex network models.

8.3 Conclusions

In this chapter we have studied the algebraic connectivity in relation to the graph's robustness to node and link failures in the main complex network models: random graph of Erdős-Rényi, the small-world of Watts-Strogatz and scale-free graph of Barabási-Albert. Based on a comprehensive set of simulations the following conclusions are drawn:

- The algebraic connectivity increases with the increasing node and the link connectivity. This means that the larger the algebraic connectivity, the larger the

number of node- or link-disjoint paths. The algebraic connectivity measures the extent to which it is difficult to cut the network into independent components and is therefore a quantifier of the robustness in complex networks.

- In the three complex network models, the minimum nodal degree is a tight upper bound on both the node and the link connectivity. Hence, the minimum nodal degree is a valuable estimate of the minimum number of nodes or links whose deletion results into a disconnected graph.
- We observe that the relationship between the algebraic connectivity and network's robustness to node and link failures is not trivial. Even with the network models that have the same density of links and where the same number of nodes (or links) have to be removed in order to disconnect them, the relationship between analyzed measures is not trivial. This points to explicit influence of the network structure on the robustness of complex networks.

Chapter 9

Influence of Network Structure on Robustness of Complex Networks

The problem of constructing networks whose robustness to failures is as high as possible has attracted a lot of attention. For example, in a recent paper [76] it has been shown that the optimal network configuration, under a classical measure of robustness, has a homogenous degree distribution or a degree distribution with no more than three distinct node connectivities. In this chapter we rely on the algebraic connectivity [48], a spectral measures that has been proven to be a distinguishable parameter in many robustness related problems [63]. The first part of this chapter presents the theoretical background on the algebraic connectivity, followed by an analytical result linking the classical connectivity to the algebraic connectivity. Subsequently, by considering the algebraic connectivity as a measure of the network robustness, we show that for a thorough understanding of robustness a proper knowledge of the topological structure of certain classes of networks is extremely important. Furthermore, we show that the type of random failure is highly predictable since the robustness to random node and link failures differs significantly between certain classes of networks. To prove this statement we consider network classes that have two structurally opposite underlying mechanisms: the homogeneous structure of the random graph of Erdős-Rényi versus the heterogeneous structure of the scale-free graph of Barabási-Albert. In addition, we consider the small-world graph of Watts-Strogatz as it exhibits the properties of both the random graph and the scale-free model. The second part of this chapter presents the simulation results on the distribution of the algebraic connectivity in the three complex network models subject to random failures. The homogeneous structure of the random graph of Erdős-Rényi implies an invariant robustness under random node and link failures. The heterogeneous structure of the scale-free graph of Barabási-Albert, on the other hand, implies a non trivial robustness to random node and link failures. Such a theoretically explained behavior was also confirmed for the small-world graph of Watts-Strogatz.

9.1 Bounds for algebraic connectivity

The second smallest eigenvalue of the standard Laplacian matrix of an undirected graph G , as proposed by Fiedler [34], is called the algebraic connectivity. There are several bounds for the algebraic connectivity related to other parameters and metrics of a graph. The following are important for our analysis [34]:

1. If $G \neq K_N$, then $0 \leq \lambda_{N-1} \leq \kappa_N \leq \kappa_L \leq D_{\min}$. If $G = K_N$, then $\lambda_{N-1} = N > \kappa_N = N - 1$.
2. If $G_1 = (\mathcal{N}, \mathcal{L}_1)$ is a subgraph of $G = (\mathcal{N}, \mathcal{L})$, then $\lambda_{N-1}(G_1) \leq \lambda_{N-1}(G)$.

9.2 Relationship between classical connectivity and probability distribution of algebraic connectivity

In this section, we calculate for the random graph of Erdős-Rényi, the values of network model parameters for which the identical distribution of the algebraic connectivity is attained. As a result, we succeed to explain the relationship between the classical connectivity and the algebraic connectivity.

The random graph model is denoted by $G_p(N)$, where N is the number of nodes in the graph and p is the probability of having a link between any two nodes. For large graph size N , the degree distribution of the random graph model, which is a binomial distribution, can be replaced by a Poisson distribution, i.e.

$$\Pr[D = k] = \binom{N-1}{k} p^k (1-p)^{N-1-k} \simeq \frac{E[D]^k e^{-E[D]}}{k!} \quad (9.1)$$

where $E[D] = p(N-1)$ equals the mean nodal degree. For large graph size N , the fundamental result [29] can be deduced that the probability of a random graph being connected is about the probability that it has no node of degree zero. In other words, for large graph size N , the following relation between the k -connectivity to the minimum nodal degree D_{\min} holds [10]

$$\Pr[G_p(N) \text{ is } k\text{-connected}] = \Pr[D_{\min} \geq k]. \quad (9.2)$$

From the above theorems, we have that the probability of k -connectivity in $G_p(N)$ equals

$$\Pr[G_p(N) \text{ is } k\text{-connected}] = \left(1 - \sum_{l=0}^{k-1} \frac{(p(N-1))^l e^{-(p(N-1))}}{l!}\right)^N. \quad (9.3)$$

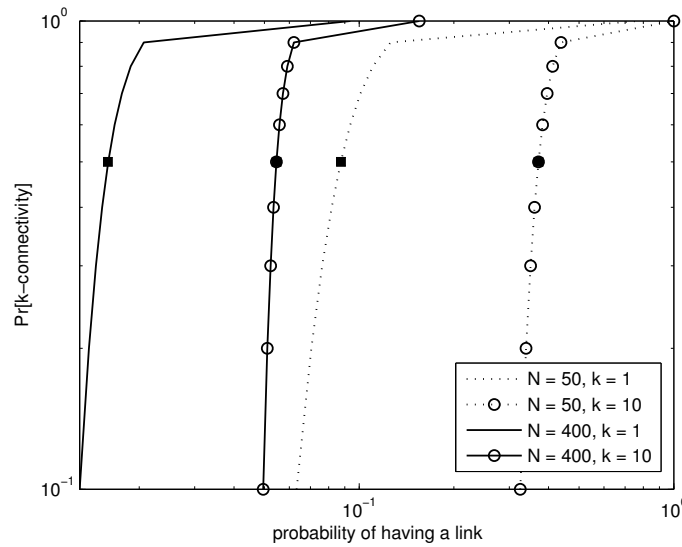


Figure 9.1: The probability of being 1-connected and 10-connected as a function of the link probability p in $G_p(N)$ with $N = 50$ and 400 . Black markers indicate the values of the link probability p for which the probability of being 1-connected and 10-connected equals 0.5. These values are also presented in 9.1.

Solving the above equation for a given probability of being k -connected and a given number of nodes N , one can easily find the probability p of the presence of a link between any two nodes in $G_p(N)$. We use here two arbitrary probabilities, i.e. 0.5 and 0.9, of a graph being k -connected to deduce p . The deduced probabilities of being k -connected are listed in Table 9.1. Figure 9.1 relates, for different values of N , the probabilities of 1-connectivity and 10-connectivity to the link probability p . We observe on Figure 9.1 that larger graph sizes require smaller link probabilities p to have a given probability of being k -connected.

Pr[k -connectivity]	$N = 50$	$N = 100$	$N = 200$	$N = 400$
Pr[1-connectivity] = 0.5	$p = 0.0875$	$p = 0.0503$	$p = 0.0285$	$p = 0.0159$
Pr[1-connectivity] = 0.9	$p = 0.1258$	$p = 0.0693$	$p = 0.0379$	$p = 0.0207$
Pr[10-connectivity] = 0.5	$p = 0.3715$	$p = 0.1963$	$p = 0.1036$	$p = 0.0546$
Pr[10-connectivity] = 0.9	$p = 0.4378$	$p = 0.2280$	$p = 0.1189$	$p = 0.0620$

Table 9.1: The link probability p in $G_p(N)$ with $N = 50, 100, 200$ and 400 , for values of the probability of being 1-connected and 10-connected of 0.5 and 0.9.

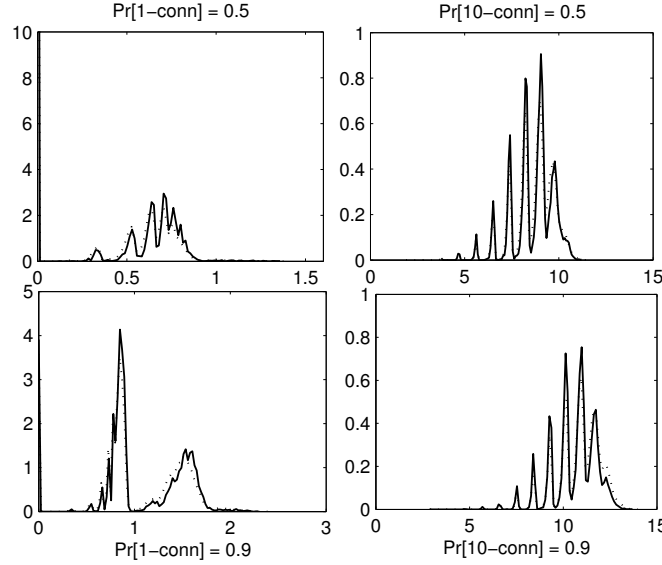


Figure 9.2: The probability density function of the algebraic connectivity λ_{N-1} in $G_p(N)$ with $N = 50$ (dotted black line), 400 (full black line) and the link probability p , which is deduced from the probability of $G_p(N)$ being k -connected.

Figure 9.2 includes the simulation results on $G_p(N)$ with $N = 50$ and 400 and the corresponding link probabilities p (equation 4). From this Figure, we can observe that, for a given probability of the graph's k -connectivity, the algebraic connectivity λ_{N-1} has the same probability distribution for different graph sizes N . Figure 9.2 suggests that the probability of being k -connected might define robustness classes, as they seem to correspond to a particular density of the algebraic connectivity. Although in previous chapter we have shown that the relationship between the algebraic connectivity and network's node and link connectivity is not trivial, here we see that, at least in distributional terms, the algebraic connectivity provides a signature of the probability of being k -connected. This result supports a direct relationship between the algebraic connectivity and graph's robustness to node and link failures and points towards the possibility of defining robustness classes based on the algebraic connectivity.

9.3 Behavior of algebraic connectivity under random failures

In this section we present simulation results on the distribution of the algebraic connectivity. We first discuss the considered network models. Then we analyze simulation results where we apply topological changes in the form of random node removal.

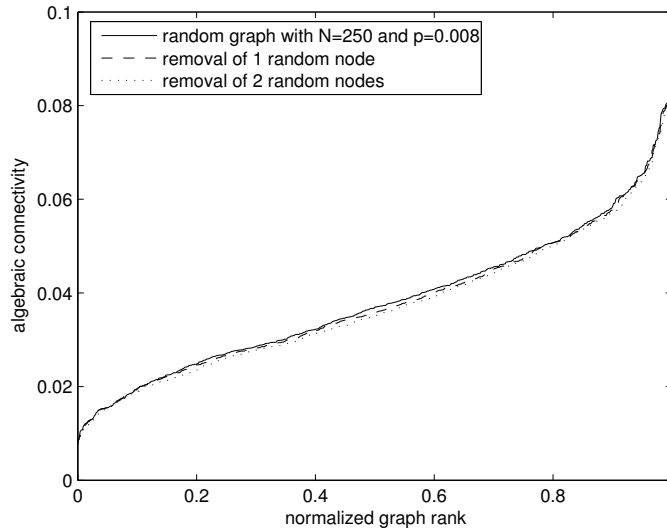


Figure 9.3: Distribution of the algebraic connectivity for the random graph of Erdős-Rényi with 250 nodes and a link probability $p = 0.008$. The distribution of the algebraic connectivity is also measured when the model is subject to topological changes in the form of random node removal.

9.3.1 Random graph of Erdős-Rényi

In the previous section we have seen that in the Erdős-Rényi random graph, the parameters of interest are N and p . We simulate for each combination of N and p , 10^3 independent $G_p(N)$ graphs. The number of nodes is 250 and the link probability takes one of the following two values, $p = 0.008$ and $p = 0.016$. From each combination of N and p , we compute the algebraic connectivity λ_{N-1} and sort the 10^3 values in increasing order. We plot them so that the i^{th} smallest algebraic connectivity value λ_{N-1}^i , $1 \leq i \leq 10^3$, is drawn at (x, y) with $x = \frac{(i-1)}{(10^3-1)}$ and $y = \lambda_{N-1}^i$. In this way all values in the x-axis are in the $[0, 1]$ range. In Figure 9.3, the distribution of the algebraic connectivity is given for the Erdős-Rényi random graph with $N = 250$ and $p = 0.008$. Note that the considered values of the link probability p are smaller than the value for which a random graph with $N = 250$ almost surely becomes connected (equation 2). For that reason, we compute the algebraic connectivity of the largest connected component. We also compute the distribution when the random graph is affected by random node failures: in two consecutive steps, we randomly remove a node and all its connections and compute the algebraic connectivity from the remaining largest connected component.

Figure 9.3 shows that the random graph subject to random node removal, exhibits a slight decrease in the distribution of the algebraic connectivity. A similar behavior has been observed for the random graph with twice the number of links (not shown).

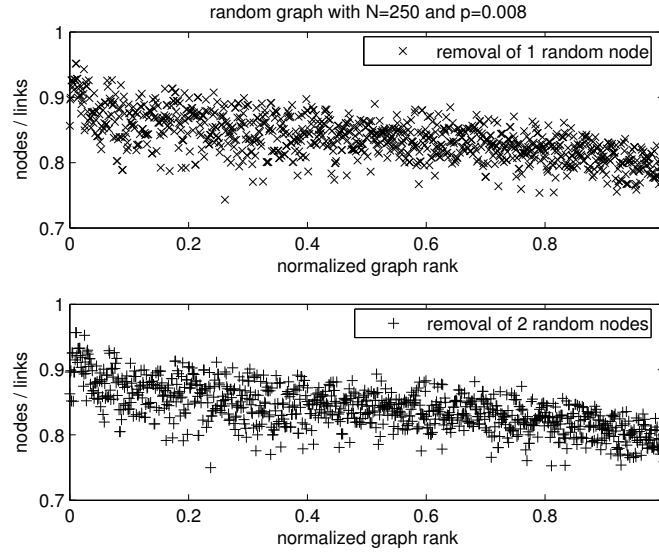


Figure 9.4: The ratio of the number of nodes to the number of links after the random graph of Erdős-Rényi with 250 nodes and $p = 0.008$ is subject to the topological changes in the form of random node removal.

The analytical results presented in Section 9.1 explain the decrease in the algebraic connectivity under the condition that removal of links would negatively affect the link density of the resulting graph (third result on the algebraic connectivity). Taking this result into account, it seems that the random removal of nodes leaves a subgraph that has approximately the same (or a slightly smaller) proportion of nodes to links as the original graph. Figure 9.4 supports this observation by showing that nodes to links ratios after one or two node removals are the same. Consequently, the removal of random nodes results in a subgraph that has approximately the same structure as the random graph upon which the changes were applied.

9.3.2 Small-World graph of Watts-Strogatz

The parameters of interest in the small-world graph of Watts-Strogatz are N , s and p_r . N stands for the number of nodes, s for the number of clockwise neighbors a node is connected to, and p_r for the probability at which each link (connected to a clockwise neighbor) is rewired to a new node chosen uniformly in the ring lattice. The rewiring process allows the small-world model to interpolate between a regular lattice ($p_r = 0$) and something which is similar, though not identical, to a random graph ($p_r = 1$). For already small p_r the small-world becomes a locally clustered network in which two arbitrary nodes are connected by a small number of intermediate links [81].

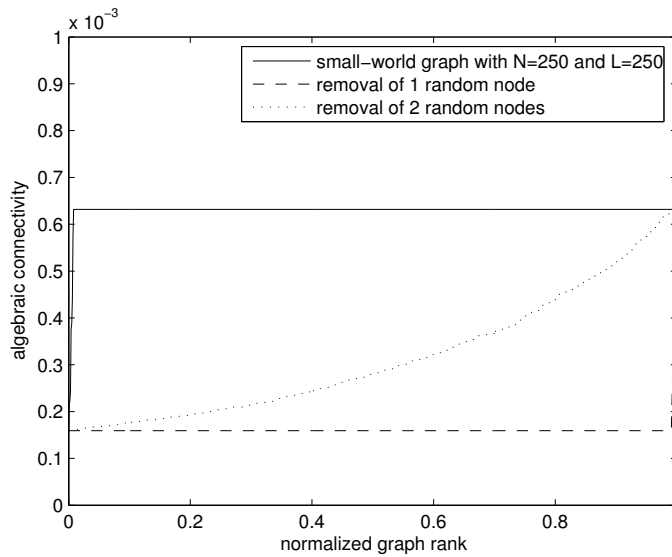


Figure 9.5: Distribution of the algebraic connectivity for the small-world of Watts-Strogatz with 250 nodes and 250 links. The distribution of the algebraic connectivity is also measured when the model is subject to topological changes in the form of random node removal.

We simulate, for N and s , 10^3 independent Watts-Strogatz small-world graphs. The number of nodes N is 250 and s takes one of the following two values, $s = 1$ and $s = 2$. The rewiring probability is set to be $p_r = 0.008$. Similarly to results for the Erdős-Rényi random graph, we plot in Figures 9.5 and 9.6 the distribution of the algebraic connectivity as a function of the normalized graph rank. We also compute and plot the distribution when the small-world network is affected by random node failures: in two consecutive steps, we randomly remove a node and all its connections and compute the algebraic connectivity from the remaining largest connected component.

From Figures 9.5 and 9.6 we observe that the value of the algebraic connectivity is constant for nearly all simulated small-world graphs. We also observe that the higher value of the algebraic connectivity is a consequence of a graph with a higher value of the link density: for $N = 250$ and $s = 2$, the small-world graph has twice as many links, i.e. $L = 500$, as the small-world with $N = 250$ and $s = 1$. The constant behavior of the algebraic connectivity comes from the fact that the small-world process introduces $p_r N s$ non-lattice links, which for $p_r = 0.008$ results in only few rewired links. Therefore, the resulting small-world graphs are almost regular ring lattices so that their values of the algebraic connectivity are constant for almost all simulated graphs.

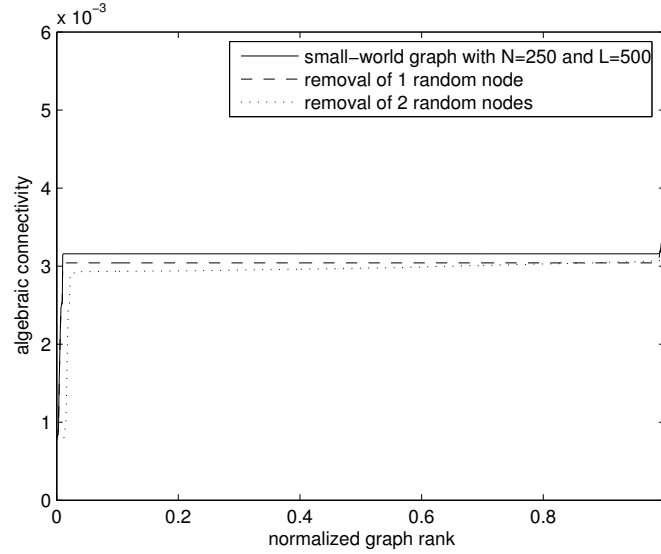


Figure 9.6: Distribution of the algebraic connectivity for the small-world of Watts-Strogatz with 250 nodes and 500 links. The distribution of the algebraic connectivity is also measured when the model is subject to topological changes in the form of random node removal.

Furthermore, the removal of random nodes in a non sparse small-world graph (see dashed lines in Figure 9.6) results in a slight decrease in the distribution of the algebraic connectivity. Then, the remaining largest connected component have the same proportion of nodes to links, i.e. $\frac{N}{L} = 0.5$, as the initial small-world graph (not shown).

On the other hand, for the sparse small-world graph, i.e. $L = N$, the distribution of the algebraic connectivity is surprisingly different. After random removal of one node, there seems to be a substantial decrease in the robustness of the remaining largest connected component. However, after random removal of two nodes, the distribution of the algebraic connectivity is likely to be increasing. Take notice of an unchanged ratio between the number of nodes and the number of links, i.e. the ratio after random removal of one or two nodes is equal to the ratio in a given small-world graph with $N = L$. As evident from the figure, this does not mean that remaining largest connected components after random removal of nodes will have the same robustness as the original graph. While this may appear inconsistent with prior results, it can be easily understood: random removal of nodes in a practically regular graph, i.e. here the ring graph with small number of rewired links, most likely fragments the small-world so that in the remaining largest connected component a highly regular structure is observed again. This process is repeated until the network is fragmented in such a way that the regularity is hardly perceptible.

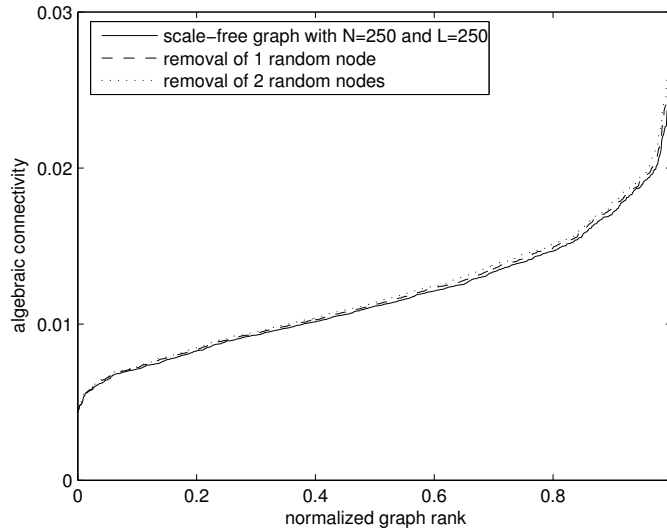


Figure 9.7: Distribution of the algebraic connectivity for the scale-free graph of Barabási-Albert with 250 nodes and 250 links, with and without random nodes removal.

9.3.3 Scale-Free graph of Barabási-Albert

The parameters of interest in the scale-free graph of Barabási-Albert are N , m_0 and t . The Barabási-Albert model starts with a small number m_0 of fully-meshed nodes, followed at every step by a new node attached to $m \leq m_0 = 2m + 1$ nodes already present in the system. After t steps this procedure results in a graph with $N = t + m_0$ nodes and $L = \frac{m_0(m_0-1)}{2} + mt$ links.

We simulate, for N and m , 10^3 independent Barabási-Albert scale-free graphs. The number of nodes is 250 and the parameter m takes one of the following two values, $m = 1$ and $m = 2$. For each combination of N and m , we compute the algebraic connectivity λ_{N-1} , and sort the 10^3 independent values in increasing order. We plot them in the same way as in the simulations for the Erdős-Rényi and Watts-Strogatz graph.

In Figures 9.7 and 9.9, the distribution of the algebraic connectivity is given, respectively for the Barabási-Albert scale-free graph with $m = 1$ and with $m = 2$. Note that for the Barabási-Albert scale-free with $m = 2$, the number of links, i.e. $L = 500$, is twice as large as for the Barabási-Albert scale-free with $m = 1$, i.e. $L = 250$. Along with the standard behavior of the algebraic connectivity, we also show the distribution when the model is exposed to the random node removal: in two consecutive steps, we randomly remove a node and all its connections and calculate the algebraic connectivity from the remaining largest connected component.

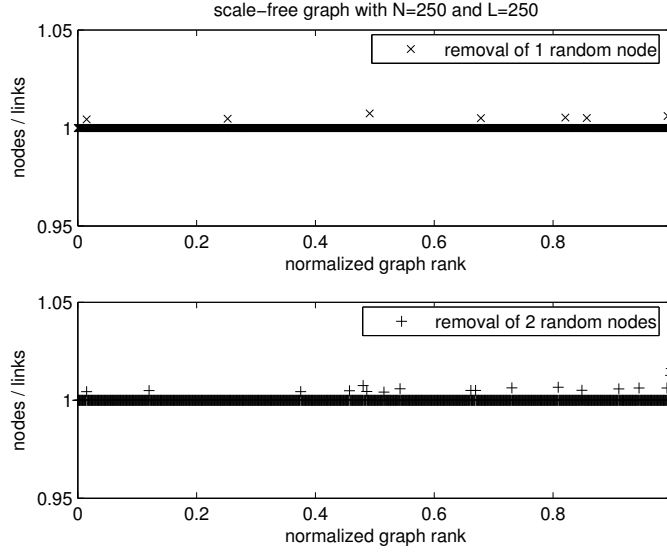


Figure 9.8: The ratio of the number of nodes to the number of links after the scale-free graph of Barabási-Albert with 250 nodes and 250 links is exposed to the topological changes in the form of random node removal.

The first observation from Figure 9.7 is that there seems to be a slight increase in the value of the algebraic connectivity due to the random removal of nodes. On the contrary, from Figure 9.9 we observe that the value of the algebraic connectivity decreases due to the random removal of nodes. To illustrate how different link densities influence the robustness, we show in Figures 9.8 and 9.10 the ratio of the number of nodes to the number of links after the scale-free graph of Barabási-Albert undergoes random node removals. We also see on Figures 9.9 and 9.10 that the gap between the algebraic connectivity before and after the removals reflects the decreasing degrees of the removed nodes. This illustrates the effect of the centrality of a node on network robustness.

From the figures showing the ratio between the number of nodes and the number of links, we observe that for the scale-free graph of Barabási-Albert with $N = L$, the random removal of nodes results in a similar topological structure, i.e. it results in a graph where the number of nodes approximately equals the number of links $N \approx L$. On the other hand, for the scale-free graph of Barabási-Albert with $N \ll L$, the random removal of nodes results in a graph where the ratio of nodes to links increases compared to the original graph (where the ratio is 0.5). Note that for the scale-free graph of Barabási-Albert with $N = 250$ and $L = 500$, the random removal of i nodes results in a connected component with a number of nodes always equals to $N - i$.

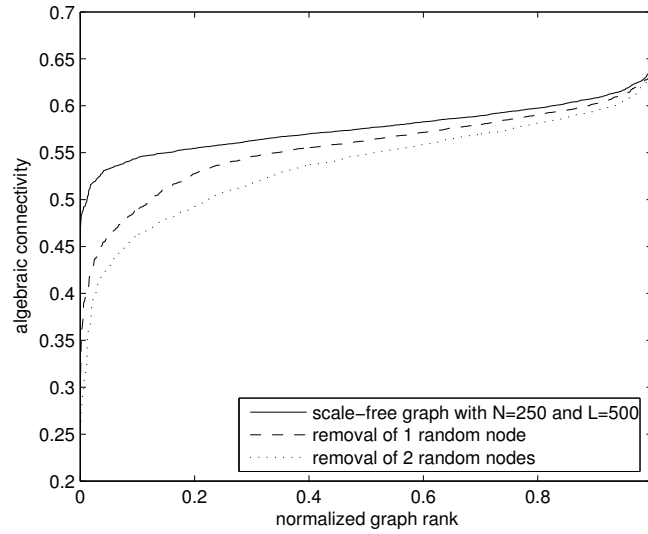


Figure 9.9: Distribution of the algebraic connectivity for the scale-free graph of Barabási-Albert with 250 nodes and 500 links, with and without random nodes removal.

The analysis of the random node failures allow us to conjecture the existence of networks belonging to two different robustness classes:

- The value of the algebraic connectivity slightly increases or remains the same: a graph has approximately the same number of nodes and links. After random node or link removal, a graph has on average the same number of nodes and links. This is either due to random nodes or random links whose removal results in a subgraph with the proportional number of nodes and links as in the original graph.
- The value of the algebraic connectivity decreases: a graph has more links than nodes. After random node or link removal, a graph has still on average more nodes than links. This decreasing algebraic connectivity is either due to
 1. random nodes whose removal results in a subgraph with the node set that spans almost all nodes and consequently much less links than in the given graph.
 2. random links whose removal results in a subgraph with the same node set as in the given graph.

Sparse and dense graphs hence will exhibit different behaviors of the algebraic connectivity under topological changes in the form of random node and link removal.

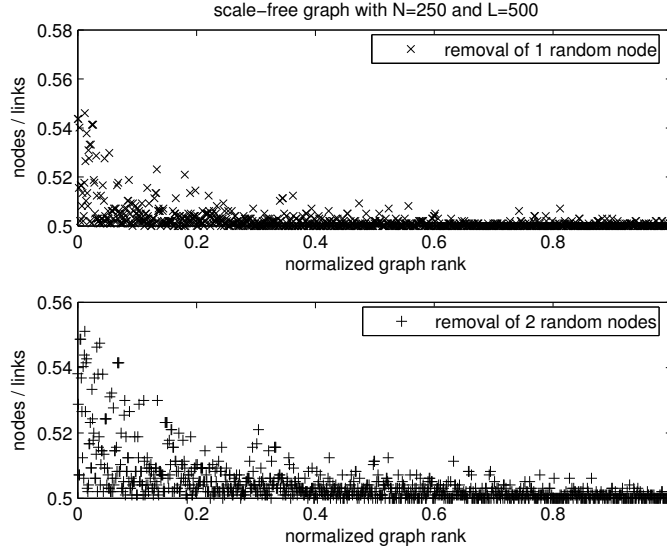


Figure 9.10: The ratio of the number of nodes to the number of links after the scale-free graph of Barabási-Albert with 250 nodes and 500 links is exposed to the topological changes in the form of random node removal.

9.4 Conclusion

In this chapter we have studied how the algebraic connectivity is affected by topological changes in the form of random node removal. For the Erdős-Rényi random graph, our findings have confirmed its well-known property: random removal of nodes results in a distribution similar to the distribution of the original random graph. Consequently, the remaining largest connected component has a similar structure and hence is equally robust to random removal of nodes. With regard to this observation, we deduce that the random link removal strategy most probably results in a decreasing distribution of the algebraic connectivity: the remaining largest connected component most probably has an unchanged set of nodes but a subset of the original links. Similar behaviour has been found for the small-world graph of Watts-Strogatz, though only for graphs with higher link densities than their sparse counterparts (where a non-trivial robustness to random node failures is observed).

For the scale-free graph of Barabási-Albert, we have observed that the distribution of the algebraic connectivity provides information on the type of failure the considered network has undergone: random node or link removal will increase the value of the algebraic connectivity only if the resulting subgraphs have approximately the same number of nodes and links. On the other hand, we have observed that the random node or link removal results in a decreased value of the algebraic connectivity only if the resulting subgraphs have a larger number of nodes than links.

Chapter 10

Conclusions

In this concluding chapter we summarize the main findings of the research performed on the subject of characterization of complex networks and its application to robustness analysis. As our findings may have implications for future work, we strive to include remarks regarding possible extensions on the subject matter.

In the thesis we exemplified that the study of complex networks has a highly interdisciplinary character where methods from a wide range of disciplines are employed to contribute to better understanding of complex networks. Furthermore, we described that recent work on the subject matter has been mainly concerned with the development of graph theoretical models and graph theoretical measures. Both seek to capture some of the relevant topological properties observed in large-scale network data. Founded on these guiding principles, in this thesis we gained new insights in the form of quantitative measures that may be applied to capture, more tightly, the qualitative topological properties that characterize technological, or any other type of complex networks. In particular, we gained new insights in the form of quantitative measures that may be applied to the problem of analyzing different topological aspects of the robustness of complex networks. In the end, we provided theoretical and empirical basis for novel results and illustrated that in this thesis followed line of research will continue to deepen into a rich and informative complex networks theory. With reference to the in this thesis addressed research questions, the main findings may be summarized as follows:

1. Concise characterization of complex networks

Over the past several years, a number of graph measures have been introduced to characterize the topology of complex networks. We performed a statistical analysis of large-scale network data sets, representing the topology of different empirical networks. We showed that some graph measures are either fully related to other measures or that they are significantly limited in the range of their possible values. In addition, we observed that subsets of graph measures are highly correlated, indicating redundancy

among them. This study implies that the set of commonly used graph measures is too extensive to concisely characterize the topology of complex networks. It also provides an important basis for classification and unification of a definite set of graph measures that would serve in future topological studies of complex networks.

2. Spectral classification of topological properties

In this thesis we studied the structural properties of both generic complex network models and empirical networks throughout the spectrum of the Laplacian matrix and the normalized Laplacian matrix. We found that among considered generic complex network models, random graph and small-world graph show a similar Laplacian spectral behavior, which differs considerably from that of the scale-free graphs. This discrepancy in the Laplacian spectrum points to explicit difference between homogenous and heterogenous network structures. We also found that both Laplacian and normalized Laplacian spectra of empirical networks provide evidence of unique spectral properties that are shared by networks within similar domains of application. Those uniquely shared spectral properties emerge as a result of the characteristic structure of specific classes of networks. This study thus implies that the (normalized) Laplacian is a qualitative spectral method that may serve towards a classification of network properties that uniquely characterize specific classes of networks.

3. Spectral measures as quantifiers of topological aspects of robustness

In this thesis we introduced the algebraic connectivity as a quantifier of the robustness to disconnection or component separation in complex networks. We studied the behavior of the algebraic connectivity in a random graph model and estimated analytically the mean and the variance of the algebraic connectivity. We used simulations to emphasize the accuracy of the analytical estimation, also for networks having small sizes. Hereby we improved a known expression for the asymptotic behavior of the algebraic connectivity.

Furthermore, we studied the algebraic connectivity in relation to the node and the link connectivity, two classical graph measures that quantify the number of nodes and links that have to be removed in order to disconnect a graph. We showed the algebraic connectivity at all times increases with the increasing node and link connectivity, implying that the larger the algebraic connectivity the more it is difficult to cut the network into independent components. In other words, we showed that the larger the algebraic connectivity, the larger number of node- or link-disjoint paths or the more nodes and links will have to be removed in order to disconnect a graph. This study implies thus that the algebraic connectivity quantifies the extent to which a network is strong and unlikely to break or have failures. Despite what has just been referred to, we proved that finding the exact relationship between the algebraic connectivity and

the node and the link connectivity is not a trivial problem. In fact, we proved that this relationship is very dependent on network structures related to different classes of complex networks.

In this thesis we also introduced the spectral radius as a quantifier of robustness to virus propagation in complex networks. We studied how well-known upper bounds for the spectral radius of graphs match the spectral radii of a number of empirical networks. This study was motivated by the fact that an sharper upper bound for the spectral radius gives a tighter lower bound for the epidemic threshold for virus propagation in networks. In addition, we explored the spectral radius of both generic complex network models as well as empirical networks. We found that the spectral radius of empirical networks with homogeneous structure in value is closest to the spectral radius of random graph models. For empirical networks with heterogeneous structure, we found that the value of the spectral radius is closest to that of scale-free graphs. This study implies that a smaller value for the spectral radius is more likely to occur for networks with homogeneous structure.

4. Spectral measures to quantify influence of network structure on robustness to different types of failures

Finally, in this thesis we relied on the algebraic connectivity to study the influence of the network structure on the robustness of complex networks. To provide the answer to this research question we considered the behavior of the algebraic connectivity for two structurally opposite network classes that are subject to topological changes in the form of different types of failures. The results from this study demonstrated that the robustness to a given number component failures significantly differs between generic complex network models. This points to an explicit influence of the network structure on the robustness to different types of failures. For example, we found that the homogeneous structure of random graphs implies an invariant robustness under random node failures. On the other hand, we found that the heterogeneous structure of small-world and scale-free graphs implies a non-trivial robustness to random node and link failures.

Appendix A

Summary Statistics of Graph Measures for Number of Empirical Networks

What follows in this appendix are summary statistics of a set of graph measures for a number empirical complex networks. We have mostly used publicly available data sets, representing the topology of complex networks from a wide range of systems in nature and society, i.e. technological, social, biological and linguistic. A detailed description of graph measures is presented in Chapter 3. Technological system we considered include the following empirical networks:

- the Dutch road infrastructure [52];
- a European national railway infrastructure;
- a European Internet Service Provider (ISP);
- a European city area power grid;
- the western states power grid of the United States [81];
- the air transportation network representing the world wide airport connections, documented at the Bureau of Transportation Statistics database [13], and the connection between United States airports [75];
- the Internet network at the autonomous system [14] and the router level [15].

Social systems include the following empirical networks:

- the network representing soccer players association from the Dutch soccer team [45];
- the network representing actor appearance in movies [6];
- the network representing collaboration among scientists [66].

Biological systems include the following empirical networks:

- the network representing frequent associations between dolphins [56];
- the network representing protein interaction of the yeast *Saccharomyces cerevisiae* [19, 39].

Linguistic systems include the following real-world networks:

- the network representing common adjacencies among words in English, French and Spanish [61].

Topological measures	Road	Rail1	Rail2	Air1	Air2
Number of nodes	14098	8710	689	500	2179
Number of links	18687	11332	778	2980	31326
Link density	0,0002	0,0003	0,0033	0,0239	0,0132
Average degree	2,7	2,6	2,3	11,9	28,8
Average neighbor degree	2,9	2,8	2,5	53,8	140,5
Assortativity coefficient	0,093	-0,022	0,098	-0,268	-0,046
Rich-club coefficient	0,0005	0,0008	0,0172	0,0621	0,3395
Clustering coefficient	0,0912	0,0212	0,0731	0,6175	0,4849
Average node distance	80,6	79,0	34,1	2,9	3,0
Average node eccentricity	177,4	158,6	65,0	5,2	5,9
Average node coreness	2,4	2,4	2,0	8,2	19,1
Average node betweenness	0,0560	0,0090	0,0481	0,0040	0,0090
Average link betweenness	0,00220	0,00350	0,02190	0,00050	0,00005
Algebraic connectivity	0,0001	0,0695	0,0008	0,1186	0,2082

Topological measures	Power1	Power2	Power3	Power4	ISP
Number of nodes	4940	3419	1713	1205	29902
Number of links	6594	3953	2043	1385	32707
Link density	0,0005	0,0007	0,0014	0,0019	0,0001
Average degree	2,7	2,3	2,4	2,3	2,2
Average neighbor degree	3,9	3,8	2,9	3,1	45,7
Assortativity coefficient	0,004	-0,128	0,022	0,108	-0,036
Rich-club coefficient	0,0026	0,0042	0,0056	0,0204	0,0085
Clustering coefficient	0,0801	0,0120	0,0145	0,0171	0,0306
Average node distance	18,5	21,1	38,0	12,3	7109,9
Average node eccentricity	34,1	38,9	71,8	22,6	14250,0
Average node coreness	2,2	2,0	2,1	2,1	2,1
Average node betweenness	0,0036	0,0059	0,0216	0,0094	0,2377
Average link betweenness	0,00140	0,00270	0,00930	0,00450	0,10870
Algebraic connectivity	0,0009	0,0003	0,0001	0,0022	0,0440

Topological measures	AS-level	Router	Protein	Soccer	Dolphins
Number of nodes	20906	29064	4626	685	62
Number of links	42994	62260	14801	10310	159
Link density	0,0002	0,0001	0,0014	0,0440	0,0841
Average degree	4,1	4,3	6,4	30,1	5,1
Average neighbor degree	230,9	21,0	24,2	45,0	6,8
Assortativity coefficient	-0,201	-0,039	-0,137	-0,063	-0,044
Rich-club coefficient	0,0101	0,0037	0,0196	0,2605	0,4127
Clustering coefficient	0,2114	0,0232	0,0912	0,7507	0,2589
Average node distance	3,9	7,1	4,2	4,5	3,4
Average node eccentricity	8,0	14,7	8,1	8,6	6,5
Average node coreness	2,9	3,0	4,4	20,2	4,5
Average node betweenness	0,0001	0,0002	0,0007	0,0050	0,0380
Average link betweenness	0,00005	0,00006	0,00014	0,00022	0,01060
Algebraic connectivity	0,0152	0,0059	0,1173	0,1612	0,1730

Topological measures	Actor	Scientific	English	French	Spanish
Number of nodes	10143	13861	7377	8308	11558
Number of links	147907	44619	44205	23832	43050
Link density	0,0029	0,0005	0,0016	0,0007	0,0006
Average degree	29,2	6,4	11,9	5,7	7,4
Average neighbor degree	83,6	13,5	320,7	218,0	457,6
Assortativity coefficient	0,026	0,157	-0,237	-0,233	-0,282
Rich-club coefficient	0,0399	0,0042	0,0588	0,0240	0,0340
Clustering coefficient	0,7551	0,6514	0,4085	0,2138	0,3764
Average node distance	3,7	6,6	2,8	3,2	2,9
Average node eccentricity	9,6	12,4	5,6	6,7	7,6
Average node coreness	21,4	4,9	7,5	3,9	4,9
Average node betweenness	0,0003	0,0004	0,0002	0,0003	0,0002
Average link betweenness	0,00040	0,00007	0,00003	0,00007	0,00003
Algebraic connectivity	0,0004	0,0292	0,1875	0,1197	0,0782

Appendix B

Dependence of Two Arbitrary Degrees in Random Graph of Erdős-Rényi

We first compute the joint probability $\Pr[D_i = k, D_j = l]$, where node i and node j are random nodes in $G_p(N)$. The dependence lies in the possible direct link between node i and j . By the law of total probability [60], we have

$$\begin{aligned} \Pr[D_i(N) = k, D_j(N) = l] &= \Pr[D_i(N) = k, D_j(N) = l | a_{ij} = 1] \Pr[a_{ij} = 1] \\ &\quad + \Pr[D_i(N) = k, D_j(N) = l | a_{ij} = 0] \Pr[a_{ij} = 0] \end{aligned}$$

where a_{ij} is the matrix element of the adjacency matrix. Since $\Pr[a_{ij} = 1] = p$, and, in absence of the direct link, D_i and D_j are independent, we obtain

$$\begin{aligned} \Pr[D_i(N) = k, D_j(N) = l] &= p \Pr[D_i(N) = k, D_j(N) = l | a_{ij} = 1] \\ &\quad + (1 - p) \Pr[D_i(N - 1) = k] \Pr[D_j(N - 1) = l] . \end{aligned}$$

Further, given the existence of the direct link, the direct link is counted both in D_i and in D_j such that

$$\Pr[D_i(N) = k, D_j(N) = l | a_{ij} = 1] = \Pr[D_i(N - 1) = k - 1] \Pr[D_j(N - 1) = l - 1] .$$

Combining the contributions and introducing the binomial density $\Pr[D_i(N) = k] = \binom{N-1}{k} p^k (1-p)^{N-1-k}$, we obtain

$$\begin{aligned} \Pr[D_i(N) = k, D_j(N) = l] &= p \binom{N-2}{k-1} p^{k-1} (1-p)^{N-1-k} \binom{N-2}{l-1} p^{l-1} (1-p)^{N-1-k} \\ &\quad + (1-p) \binom{N-2}{k} p^k (1-p)^{N-2-k} \binom{N-2}{l} p^l (1-p)^{N-2-l} . \end{aligned}$$

The joint expectation is

$$\begin{aligned} E[D_i(N), D_j(N)] &= \sum_{k=0}^{N-1} \sum_{l=0}^{N-1} lk \Pr[D_i(N) = k, D_j(N) = l] \\ &= p - 2Np^2 + N^2p^2. \end{aligned}$$

The covariance is

$$\begin{aligned} Cov[D_i(N), D_j(N)] &= E[D_i(N), D_j(N)] - E[D_i(N)] E[D_j(N)] \\ &= p(1 - p) \end{aligned}$$

Since the variance $Var[D_i(N)] = Var[D_j(N)] = (N - 1)p(1 - p)$, we find that the correlation coefficient [60] is

$$c(D_i(N), D_j(N)) = \frac{Cov[D_i(N), D_j(N)]}{\sqrt{Var[D_i(N)]}\sqrt{Var[D_j(N)]}} = \frac{1}{N - 1}$$

which shows that, although arbitrary degrees $D_i(N)$ and $D_j(N)$ are not independent, their degree of correlation decreases with N . Hence, for large N , both $D_i(N)$ and $D_j(N)$ can be regarded as almost independent.

Appendix C

Minimum of Normalized i.i.d. Binomially Distributed Sequence of Degrees in Random Graph of Erdős-Rényi

For large N and constant p , independent of N , the normalized i.i.d. binomially distributed sequence $\{D_i^*\}_{1 \leq i \leq N}$ of all degrees in $G_p(N)$ tends to be Gaussian distributed, as specified in (6.2). The minimum of the sequence $\{D_i^*\}_{1 \leq i \leq N}$ possesses the distribution

$$\Pr[\min_{1 \leq i \leq N} D_i^* \leq x] = 1 - \prod_{i=1}^N \Pr[D_i^* > x] = 1 - (\Pr[D_i^* > x])^N$$

or simplified

$$\Pr[\min_{1 \leq i \leq N} D_i^* > x] = (1 - F_{D_i^*}(x))^N$$

where $F_{D_i^*}(x) = \Pr[D_i^* \leq x]$.

Consider now the limiting process of the minimum of a set $\{D_i^*\}_{1 \leq i \leq N}$ when $N \rightarrow \infty$. By choosing an appropriate sequence $\{x_N\}$ such that ζ is finite, a scaling law for the minimum of a sequence can be obtained as

$$\lim_{N \rightarrow \infty} \Pr[\min_{1 \leq i \leq N} D_i^* > x_N] = e^{-\zeta}$$

where $\lim_{N \rightarrow \infty} NF_{D_i^*}(x_N) = \zeta$. Hence, we have to find the appropriate x_N in such a way that the limit is independent of N but equal to ζ . For a Gaussian random variable D_i^* the condition $\lim_{N \rightarrow \infty} NF_{D_i^*}(x_N) = \zeta$ becomes

$$\lim_{N \rightarrow \infty} \frac{N}{\sqrt{2\pi}} \int_{-\infty}^{x_N} e^{-\frac{u^2}{2}} du = \zeta. \quad (\text{C.1})$$

After substitution of $u = -\sqrt{2}t$, the finite integral is rewritten in term of the $\operatorname{erfc}(x)$ function as

$$\int_{-\infty}^{x_N} e^{-\frac{u^2}{2}} du = \sqrt{2} \int_{\frac{-x_N}{\sqrt{2}}}^{\infty} e^{-t^2} dt = \sqrt{2} \frac{\sqrt{\pi}}{2} \operatorname{erfc}\left(\frac{-x_N}{\sqrt{2}}\right)$$

and using the asymptotic expansion of the $\operatorname{erfc}(x)$ function [1, Section 7.1.23],

$$\int_{-\infty}^{x_N} e^{-\frac{u^2}{2}} du = \frac{e^{-\frac{x_N^2}{2}}}{-x_N} \left(1 + O\left(\frac{1}{x_N^2}\right)\right).$$

In order to have a finite limit in (C.1), x_N must be negative and tending to minus infinity for large N . Substituting the asymptotic expansion into (C.1), we have for large N ,

$$\frac{N}{\sqrt{2\pi}} \frac{e^{-\frac{x_N^2}{2}}}{-x_N} \sim \zeta$$

or

$$-x_N \sim \frac{N}{\zeta \sqrt{2\pi}} e^{-\frac{x_N^2}{2}}.$$

After squaring and taking the logarithm of both sides, we obtain with $y = -\log \zeta$

$$\log(-x_N)^2 = \log \frac{N^2}{2\pi} + 2y - (-x_N)^2$$

or

$$-x_N = \sqrt{2 \log N - \log 2\pi + 2y - 2 \log(-x_N)}. \quad (\text{C.2})$$

This is a non-linear equation in $-x_N > 0$. An order estimate for large N of the solution $(-x_N)$ is obtained as follows. Substitution of the right-hand side of (C.2) into (C.2) gives

$$\begin{aligned} -x_N &= \sqrt{2 \log N - \log 2\pi + 2y - \log(2 \log N - \log 2\pi + 2y - 2 \log(-x_N))} \\ &= \sqrt{2 \log N} \sqrt{1 + \frac{2y - \log 2\pi - \log(\log \frac{N^2}{2\pi} + 2y - 2 \log(-x_N))}{2 \log N}}. \end{aligned}$$

Furthermore,

$$\begin{aligned} \log\left(\log \frac{N^2}{2\pi} + 2y - 2 \log(-x_N)\right) &= \log\left(\log \frac{N^2}{2\pi} \left(1 + \frac{2y - 2 \log(-x_N)}{\log \frac{N^2}{2\pi}}\right)\right) \\ &= \log\left(\log \frac{N^2}{2\pi}\right) + \log\left(1 + \frac{2y - 2 \log(-x_N)}{\log \frac{N^2}{2\pi}}\right). \end{aligned}$$

With (C.2) we obtain $\frac{2y-2\log(-x_N)}{\log \frac{N^2}{2\pi}} = O\left(\frac{\log \log N}{\log N}\right)$ and $\log\left(1 + \frac{2y-2\log(-x_N)}{\log \frac{N^2}{2\pi}}\right) = O\left(\frac{\log \log N}{\log N}\right)$, where the Taylor expansion of $\log(1+x) = x + O(x^2)$ for small x has been used. Hence, for large N ,

$$\log\left(\log \frac{N^2}{2\pi} + 2y - 2\log(-x_N)\right) = \log\left(\log \frac{N^2}{2\pi}\right) + O\left(\frac{\log \log N}{\log N}\right)$$

and

$$-x_N = \sqrt{2\log N} \sqrt{1 + \frac{2y - \log 2\pi - \log\left(\log \frac{N^2}{2\pi}\right) + O\left(\frac{\log \log N}{\log N}\right)}{2\log N}}.$$

Using the Taylor expansion $(1-x)^\alpha = 1 + \alpha x + O(x^2)$ for small x gives

$$-x_N = \sqrt{2\log N} \left(1 + \frac{2y - \log 2\pi - \log\left(\log \frac{N^2}{2\pi}\right) + O\left(\frac{1}{\log N}\right)}{4\log N} + O\left(\frac{\log \log N}{\log^2 N}\right)\right)$$

from which (C.3) follows. The appropriate solution thus yields

$$x_N = -\sqrt{2\log N} - \frac{y}{\sqrt{2\log N}} + \frac{\log\left(\sqrt{2\pi \log \frac{N^2}{2\pi}}\right)}{\sqrt{2\log N}} + O\left(\frac{\log \log N}{\log^{\frac{3}{2}} N}\right) \quad (\text{C.3})$$

from which we finally arrive at

$$\lim_{N \rightarrow \infty} \Pr \left[\min_{1 \leq k \leq N} D_i^* > -\sqrt{2\log N} - \frac{y}{\sqrt{2\log N}} + \frac{\log\left(\sqrt{2\pi \log \frac{N^2}{2\pi}}\right)}{\sqrt{2\log N}} \right] = e^{-e^{-y}}$$

or

$$\lim_{N \rightarrow \infty} \Pr \left[-\sqrt{2\log N} \min_{1 \leq i \leq N} D_i^* - 2\log N + \log\left(\sqrt{2\pi \log \frac{N^2}{2\pi}}\right) < y \right] = e^{-e^{-y}}. \quad (\text{C.4})$$

Thus,

$$Y = -\sqrt{2\log N} \min_{1 \leq i \leq N} D_i^* - 2\log N + \log\left(\sqrt{2\pi \log \frac{N^2}{2\pi}}\right)$$

is a Gumbel random variable such that

$$D_{\min}^* = \frac{-Y - 2\log N + \log\left(\sqrt{2\pi \log \frac{N^2}{2\pi}}\right)}{\sqrt{2\log N}}.$$

Appendix D

Symbols and Acronyms

Ω	adjacency matrix
Δ	degree matrix
Λ	Laplacian matrix
Π	normalized Laplacian matrix
α	value of variable
κ_N	node connectivity
κ_L	link connectivity
λ_{N-1}	algebraic connectivity
ρ	spectral radius
ρ_{ER}	spectral radius of Erdős and Rény random graph
ρ_{WS}	spectral radius of Watts and Strogatz small-world graph
ρ_{BA}	spectral radius of Barabási and Albert scale-free graph
$\sigma[X]$	standard deviation of random variable X
ϕ_x	rich-club coefficient induced by nodes with degrees larger than given value x
\mathcal{L}	set of links
\mathcal{N}	set of nodes

a_{ij}	(i, j) entry of adjacency matrix
$Cov[X, X']$	covariance of two random variables X and X'
c	correlation coefficient
c_i	clustering coefficient of node i
c_G	average of clustering coefficients of graph
$diam$	diameter
D_i	degree of node i
D_{\max}	maximum degree
D_{\min}	minimum degree
$E[X]$	mean of random variable X
$E[X, X']$	joint mean of two random variables X and X'
G	graph
$G_p(N)$	random graph with parameters p and N
H_i	hopcount or shortest path length of a node i
H_{\max}	maximum hopcount
H_{\min}	minimum hopcount
(i, j)	entry of a matrix that lies in the i -th row and the j -th column
K_N	complete graph with N nodes
$K_{1,N-1}$	star graph with N nodes
$K_{N,N}$	complete bipartite graph with $N + N$ nodes
L	number of links
m	number of links a newly attached node has in scale-free graph
m_0	number of fully-meshed nodes at beginning of construction of scale-free graph
N	number of nodes
O	big O notation used either to characterize residual term of asymptotic series or to characterize complexity of algorithms
P_N	path graph with N nodes
$\Pr[X]$	probability of random variable X
$\Pr[X, X']$	joint probability of two random variables X and X'
p	probability of having link between any two nodes in random graph
p_c	probability at which in random graph a giant component is formed
p_r	rewiring probability in small-world graph
q	link density
R_N	ring with N nodes
$R_{N,s}$	ring lattice with N nodes and Ns links
r	assortativity coefficient
$Sdev[X]$	standard deviation of random variable X
s	number of clockwise neighbor node has in small-world graph
t	number of new attached nodes in the scale-free graph
$Var[X]$	variance of random variable X

AS	Autonomous System
BA	scale-free graph of Barabási and Albert
ER	random graph of Erdős and Rényi
IP	Internet Protocol
ISP	Internet service provider
MST	minimum spanning tree
PCA	principal component analysis
ST	spanning tree
WS	small-world graph of Watts and Strogatz

Bibliography

- [1] M. Abramowitz and I.A. Stegun. *Handbook of mathematical functions*. Dover Publications, 1999.
- [2] R. Albert and A.-L. Barabási. Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74:47–94, 2002.
- [3] Z. Bai, J. Demmel, J. Dongarra, A. Ruhe, and H. van der Vorst. *Templates for the solution of algebraic eigenvalue problems: A practical guide*. SIAM, Philadelphia, 2000.
- [4] A. Banerjee and J. Jost. Spectral plot properties: Towards a qualitative classification of networks. In *European Conference on Complex Systems*, 2007.
- [5] A.-L. Barabási. *Linked: The new science of networks*. Perseus Books Group, 2002.
- [6] A.-L. Barabási and R. Albert. Emergence of scaling in random networks. *Science*, 286:509–512, 1999.
- [7] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D.-U. Hwang. Complex networks: Structure and dynamics. *Physics Reports*, 424:175–308, 2006.
- [8] B. Bollobás. The evolution of sparse graphs. In *Graph Theory and Combinatorics. A volume in honour of Paul Erdős*, pages 35–57. Academic Press, 1984.
- [9] B. Bollobás. *Modern graph theory*. Springer-Verlag, 1998.
- [10] B. Bollobás. *Random Graphs*. Cambridge University Press, 2001.
- [11] B. Bollobás and A.G. Thomason. Random graphs of small order. *Annals of Discrete Mathematics*, 28:47–97, 1985.
- [12] S.P. Borgatti and M.G. Everett. Models of core/periphery structures. *Social Networks*, 21:375–395, 1999.
- [13] Bureau of Transportation Statistics. <http://www.bts.gov>.

- [14] CAIDA AS-Level Topology Measurements. http://www.caida.org/tools/measurement/skitter/as_adjacencies.xml.
- [15] CAIDA Router-Level Topology Measurements. http://www.caida.org/tools/measurement/skitter/router_topology.
- [16] D. Cao. Bounds on eigenvalues and chromatic numbers. *Linear Algebra and its Applications*, 270:1–13, 1998.
- [17] F.R.K. Chung. *Spectral graph Theory (CBMS regional conference series in mathematics)*. American Mathematical Society, 1997.
- [18] S.M. Cioaba, D.A. Gregory, and V. Nikiforov. Extreme eigenvalues for nonregular graphs. *Journal of Combinatorial Theory B*, 97:483–486, 2006.
- [19] V. Colizza, A. Flammini, M.A. Serrano, and A. Vespignani. Detecting rich-club ordening in complex networks. *Nature Physics*, 2:110–115, 2006.
- [20] D.M. Cvetković, M. Doob, I. Gutman, and A. Torgašev. *Recent Results in the theory of Graph Spectra*. North-Holland, 1988.
- [21] D.M. Cvetković, M. Doob, and H. Sachs. *Spectra of graphs, theory and applications*. Johann Ambrosius Barth Verlag, 1995.
- [22] L. da F. Costa, F.A. Rodrigues, G. Travieso, and P.R. Villas Boas. Characterization of complex networks: A survey of measurements. *Advances of Physics*, 56:167–242, 2007.
- [23] D.J. Daley and J. Gani. *Epidemic modelling: An Introduction*. Cambridge University Press, 1999.
- [24] K.C. Das and P. Kumar. Some new bounds on the spectral radius of graphs. *Discrete Mathematics*, 281:149–161, 2004.
- [25] Y. Dinic. Algorithm for solution of a problem of maximum-flow in a network. *Soviet Mathematics - Doklady*, 11:1277–1280, 1970.
- [26] S.N. Dorogovtsev, A.V. Goltsev, J.F.F. Mendes, and A.N. Samukhin. Spectra of complex networks. *Physical Review E*, 68:046109, 2003.
- [27] Dutch soccer team network. The data was obtained from <http://www.voetbalstats.nl>, which gives the line-ups for all Dutch international soccer matches. We have considered all matches up till Ireland-Holland (16 august 2006), which was match number 644.

- [28] J. Edmonds and R.M. Karp. Theoretical improvements in algorithmic efficiency for network flow problems. *Journal of the ACM*, 19:248–264, 1972.
- [29] P. Erdős and A. Rényi. On random graphs. *Publicationes Mathematicae*, 6:290–297, 1959.
- [30] P. Erdős and A. Rényi. On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, 5:17–61, 1960.
- [31] P. Erdős and A. Rényi. On the strength of connectedness of a random graph. *Acta Mathematica Scientia Hungary*, 12:261–267, 1961.
- [32] L. Euler. Solutio problematis ad geometriam situs pertinentis (the solution of a problem relating to the geometry of position). *Commentarii academiae scientiarum Petropolitanae*, 8:128–140, 1741.
- [33] M. Faloutsos, P. Faloutsos, and C. Faloutsos. On power-law relationships of the Internet topology. In *Proceedings of ACM SIGCOMM 1999*, 1999.
- [34] M. Fiedler. Algebraic connectivity of graphs. *Czechoslovak Mathematical Journal*, 23:298–305, 1973.
- [35] A. Gibbons. *Algorithmic graph theory*. Cambridge University Press, 1985.
- [36] C. Gkantsidis, M. Mihail, and E. Zegura. Spectral analysis of Internet topologies. In *Proceedings of IEEE Infocom 2003*, 2003.
- [37] A.V. Goldberg. A new max-flow algorithm. In *Technical Memo MIT/LCS/TM-291*, 1985.
- [38] R. Grone, R. Merris, and V.S. Sunder. The Laplacian spectrum of a graph. *SIAM Journal on Matrix Analysis*, 11:218–238, 1990.
- [39] A.-L. Barabási H. Jeong, S. Mason and Z.N. Oltvai. Centrality and lethality of protein networks. *Nature*, 411:41–42, 2001.
- [40] S. Hakimi. On the realizability of a set of integers as degree of the vertices of a graph. *SIAM Journal on Applied Mathematics*, 10:496–506, 1962.
- [41] V. Havel. A remark on the existence of finite graphs. *Casopis Pest. Mat.*, 80:477–480, 1955.
- [42] P. Holme, B.J. Kim, C.N. Yoon, and S.K. Han. Attack vulnerability of complex networks. *Physical Review E*, 65:056109, 2002.

- [43] Y. Hong. A bound on the spectral radius of graphs. *Linear Algebra and its Applications*, 108:135–140, 1988.
- [44] Y. Hong, J.-L. Shu, and K. Fang. A sharp upper bound on the spectral radius of graphs. *Journal of Combinatorial Theory B*, 81:177–183, 2001.
- [45] A. Jamakovic, R.E. Kooij, P. Van Mieghem, and E.R. van Dam. Robustness of networks against viruses: The role of the spectral radius. In *Proceedings of the 13th Annual Symposium of the IEEE/CVT Benelux*, 2006.
- [46] A. Jamakovic and P. Van Mieghem. The Laplacian spectrum of complex networks. In *European Conference on Complex Systems*, 2006.
- [47] A. Jamakovic and P. Van Mieghem. On the robustness of complex networks by using algebraic connectivity. In *IFIP Networking Conference*, 2008.
- [48] A. Jamakovic and S. Uhlig. Influence of network structure on robustness. In *15th IEEE International Conference on Networks*, 2007.
- [49] A. Jamakovic and S. Uhlig. On the relationship between the algebraic connectivity and graph’s robustness to node and link failures. In *3rd EuroNGI Conference on Next Generation Internet Networks*, 2007.
- [50] A. Jamakovic and S. Uhlig. On the relationships between topological measures in real-world networks. *Networks and Heterogeneous Media Journal*, 3:345–359, 2008.
- [51] A. Jamakovic, S. Uhlig, and I. Theisler. On the relationships between topological metrics in real-world networks. In *European Conference on Complex Systems*, 2007.
- [52] A. Jamakovic, H. Wang, and P. Van Mieghem. Topological characteristics of the Dutch road infrastructure. In *Seminar Infrastructure Reliability TU Delft*, 2006.
- [53] S. Janson, D.E. Knuth, T. Luczak, and B. Pittel. The birth of the giant component. *Random Structures and Algorithms*, 4:233–358, 1993.
- [54] I.T. Jolliffe. *Principal component analysis*. Springer-Verlag, 2002.
- [55] F. Juhaz. The asymptotic behaviour of Fiedler’s algebraic connectivity for random graphs. *Discrete Mathematics*, 96:59–63, 1991.
- [56] D. Lusseau, K. Schneider, O.J. Boisseau, P. Haase, E. Slooten, and S.M. Dawsen. The bottlenose dolphin community of Doubtful Sound features a large proportion of long-lasting associations. Can geographic isolation explain this unique trait? *Behavioral Ecology and Sociobiology*, 54:396–405, 2003.

- [57] P. Mahadevan, D. Krioukov, M. Fomenkov, X. Dimitropoulos, kc claffy, and A. Vahdat. The Internet AS-level topology: Three data sources and one definitive metric. *ACM SIGCOMM Computer Communication Review*, 36:17–26, 2006.
- [58] R. Merris. Laplacian matrices of graphs: A survey. *Linear Algebra and its Applications*, 197:143–176, 1994.
- [59] R. Merris. A survey of graph Laplacians. *Linear and Multilinear Algebra*, 39:19–31, 1995.
- [60] P. Van Mieghem. *Performance Analysis of Communication Systems and Networks*. Cambridge University Press, 2006.
- [61] R. Milo, S. Itzkovic, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon. Superfamilies of evolved and designed networks. *Science*, 303:1538–1542, 2004.
- [62] B. Mohar. Laplace eigenvalues of graphs: A survey. *Discrete Mathematics*, 109:171–183, 1992.
- [63] B. Mohar. Some applications of Laplace eigenvalues of graphs. *Graph Symmetry: Algebraic Methods and Applications, NATO ASI Ser. C*, 497:225–275, 1997.
- [64] B. Mohar, Y. Alavi, G. Chartrand, O.R. Oellermann, and A.J. Schwenk. The Laplacian spectrum of graphs. *Graph Theory, Combinatorics and Applications*, 2:53–64, 1991.
- [65] W. Mühlbauer, A. Feldmann, O. Maennel, M. Roughan, and S. Uhlig. Building an AS-topology model that captures route diversity. In *Proceedings of ACM SIGCOMM 2006*, 2006.
- [66] M.E.J. Newman. The structure of scientific collaboration networks. *The Proceedings of the National Academy Sciences Online (US)*, 98:404–409, 2001.
- [67] M.E.J. Newman. Assortative mixing in networks. *Physical Review Letters*, 89:208701, 2002.
- [68] M.E.J. Newman. The structure and function of complex networks. *SIAM Review*, 45(2):167–256, 2002.
- [69] R. Olfati-Saber. Ultra-fast consensus in small-world networks. In *Proceedings of American Control Conference*, 2005.
- [70] R. Oliveira, B. Zhang, and L. Zhang. Observing the evolution of Internet AS topology. In *Proceedings of ACM SIGCOMM 2007*, 2007.

- [71] PlanetLab Consortium. <http://www.planet-lab.org/>.
- [72] RIPE NCC Test Traffic Measurements Service. <http://www.ripe.net/ttm>.
- [73] RouteViews, University of Oregon Route Views Project. <http://www.routeviews.org/>.
- [74] S.H. Strogatz. Exploring complex networks. *Nature*, 410:268–276, 2001.
- [75] R. Pastor-Satorras V. Colizza and A. Vespignani. Reaction-diffusion processes and metapopulation models in heterogeneous networks. *Nature Physics*, 3:276–282, 2007.
- [76] A.X.C.N. Valente, A. Sarkar, and H.A. Stone. Two-peak and three-peak optimal complex networks. *Physical Review Letters*, 92:118702, 2004.
- [77] E. R. van Dam and W.H. Haemers. Which graphs are determined by their spectrum? *Linear Algebra and its Applications*, 373:241–272, 2003.
- [78] D. Vukadinovic, P. Huang, and T. Erlebach. On the spectrum and structure of Internet topology graphs. In *Proceedings of the second international workshop on innovative Internet computing systems*, 2002.
- [79] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos. Epidemic spreading in real networks: An eigenvalue viewpoint. In *Proceedings of the 22nd Symposium in Reliable Distributed Computing*, 2003.
- [80] D.J. Watts. *Small Worlds: The dynamics of networks between order and randomness*. Princeton University Press, 1999.
- [81] D.J. Watts and S.H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393:440–442, 1998.
- [82] Wolfram Mathworld Web Mathematics Resource. <http://mathworld.wolfram.com/RandomGraph.html>.
- [83] S. Zhou, G.-Q. Zhang, and G.-Q. Zhang. Chinese Internet AS-level topology. *IET Communications*, 1:209–214, 2007.

Samenvatting

Karakterisering van complexe netwerken: Toepassing op robuustheids analyse

De recente vooruitgangen in complexiteitswetenschap hebben onthuld dat *complexe netwerken* op vele schalen en in vele verschillende onderzoeksdomeinen, evenals natuur voorkomen. Dergelijke netwerken, het is nu reeds wel gevestigd, bezitten gemeenschappelijke eigenschappen in termen van hun niet alledaagse netwerk structuur, de *netwerk topologie* genoemd.

In deze thesis wordt een bepaalde onderzoekslijn van complexe netwerken gevolgd welke hoofdzakelijk betrekking heeft op de karakterisering van niet alledaagse topologische eigenschappen van complexe netwerken. Dit onderzoekslijn van complexe netwerken wordt uitgebreid door die elementaire *graafmetrieken* te analyseren, nader geclassificeerd in *structurele- en spectrale- metrieken*, die van belang zijn tijdens het kwantificeren van verschillende topologie-gerelateerde aspecten van de *robuustheid* van complexe netwerken. Deze thesis maakt de volgende contributies bij het onderzoeksdomain van complexe netwerken. In het eerste inleidende deel van de thesis presenteren wij 1) een algemene beschouwing van het onderzoek naar complexe netwerken door de toepassing van grafentheorie, 2) een korte beschrijving van de generieke modellen die gebruikt worden voor de modellering van complexe netwerken, en 3) een kort overzicht van praktisch belangrijke graafmetrieken. In het hoofddeel van deze thesis beantwoorden wij de vier onderzoeksvragen, welke als volgt zijn samengevat:

1. Analyse van relaties tussen een verscheidenheid van bestaande structurele en spectrale metrieken om een definitieve set te introduceren, geschikt voor het weergeven van de meest relevante topologische eigenschappen van complexe netwerken.
2. Studie naar de toepasbaarheid van spectrale metrieken voor het classificeren van de kwalitatieve topologische eigenschappen die specifieke klassen van complexe netwerken kenmerken.
3. Studie naar de toepasbaarheid van spectrale metrieken voor het kwantificeren

van de verschillende topologische aspecten van de robuustheid van complexe netwerken.

4. Praktische toepassing van spectrale metrieken om te kwantificeren hoe de robuustheid van verschillende soorten storingen in het onderliggende complexe netwerken structuur tot uiting komt.

Als eerste analyseren wij de relaties onder een verscheidenheid van de bestaande graafmetrieken om een eerste fundamentele stap te maken in het introduceren van een definitieve set, geschikt voor het weergeven van de meest relevante topologische eigenschappen van complexe netwerken. Wij bouwen deze studie uit door verder te focussen op de topologische eigenschappen die betrekking hebben op de eigenwaarden van een karakteristieke matrix van het netwerk. In het bijzonder passen wij de spectrale grafentheorie toe om de spectrale eigenschappen van generieke en empirische netwerken te analyseren en laten wij zien hoe deze methode beoogt de kwalitatieve karakterisering van verschillende klassen van netwerken weer te geven. Wij illustreren verder het gebruik van spectrale metrieken bij de onderzoeksvraag van kwantitatieve karakterisering van verschillende topologische aspecten van de robuustheid van complexe netwerken: wij introduceren een bepaalde eigenwaarde van een karakteristieke matrix van het netwerk, nader aangeduid als de algebraïsche connectiviteit, als een maatstaf van de robuustheid tegen disconnectiviteit of opsplitsing in complexe netwerken. Wij introduceren hierbij ook een andere eigenwaarde van een karakteristieke matrix van het netwerk, nader aangeduidt als de spectrale radius, als een maatstaf van de robuustheid tegen viruspropagatie in complexe netwerken. Na de introductie van mogelijke maatstaaven van de topologie-gerelateerde aspecten van de robuustheid, analyseren wij het verband tussen de algebraïsche connectiviteit en de klassieke metrieken die de mate weergeven waarin een netwerk zich kan aanpassen aan het falen van zijn componenten. In het laatste deel van de thesis bestuderen wij een praktische toepassing van de algebraïsche connectiviteit om te kwantificeren hoe de robuustheid tegen verschillende soorten storingen in de onderliggende complexe netwerken structuur tot uiting komt. Deze studie is een directe consequentie van de analyse van de relaties tussen de algebraïsche connectiviteit en de klassieke metrieken die de mate weergeven waarin een netwerk zich kan aanpassen aan zijn componentenstoringen. Uiteindelijk recapituleren wij de belangrijkste resultaten en verstrekken een uitgangspunt voor het verder leren in deze bepaalde onderzoekslijn van complexe netwerken.

Acknowledgements

This thesis is a result of research I have performed during four-year period within the Network Architectures and Services Group at the Faculty of Electrical Engineering, Mathematics, and Computer Science of Delft University of Technology, in the Netherlands. Several people contributed to its successful accomplishment and I would like to thank them.

First, I would like to thank my promotor Prof. Piet Van Mieghem for his rigorous guidance towards my doctorate. I would also like to express my gratitude to all dissertation defense committee members. It is indeed a great honour that they took an active part in the life of this thesis.

I am also thankful for having had numerous pleasurable discussions with all colleagues from both the Network Architectures and Services Group, and the Wireless and Mobile Communications Group. I am especially grateful for having had collaborations with peers without whose knowledge this thesis would be shortened. Therefore, I am indebted to Dr. Steve Uhlig, Dr. Rob Kooij and Dr. Fernando Kuipers. I am also very pleased to have worked with students in their efforts towards a M.Sc. degree. Teamwork is the best way to achieve results immortality...

Finally, I am really an most grateful to the those who are the true ones: my parents for endlessly being proud of me, my sister for deeply loving me and my friends for at all times encouraging me.

Almerima

Rotterdam, April 2008

Curriculum Vitae

First Name	Almerima
Last Name	Jamaković
Title	MSc
Date of Birth	14 August 1977
Place of Birth	Sarajevo, Bosnia and Herzegovina
Gender	Female
Nationalities:	Dutch, Bosnian

Almerima Jamakovic was born in Sarajevo, Bosnia and Herzegovina. She obtained a Bachelor of Science (BSc) degree in Electrical Engineering at Breda Institute of Technology, the Netherlands, in the year 2001. Subsequently, she obtained a Master of Science (MSc) degree in Electrical Engineering at Delft University of Technology, the Netherlands, in the year 2004. After finishing her study she joined the Network Architectures and Services (NAS) Group to work towards a Ph.D. degree. Under the supervision of the NAS group chair Prof. Van Mieghem she performed research in the field of complex networks, focusing on the robustness of such "real-world" structures.

During her PhD studies at Delft University of Technology she assisted in lecturing the Telecommunication Networks and the Performance Analysis, as well as guided two students in their efforts towards a MSc degree in Electrical Engineering. She also served as a reviewer for various international conferences and journals in the field of telecommunications. In the period January - May 2007, she visited CAIDA, the Co-operative Association for Internet Data Analysis, San Diego Supercomputer Center at University of California in San Diego. For this working visit, she has been awarded a full scholarship from the Netherlands Organization for Scientific Research (NWO).

Scientific publications:

1. D. Fay, H. Haddadi, A. Thomason, A. Moore, R. Mortier, A. Jamakovic, S. Uhlig and M. Rio, 2008, *Weighted spectral distribution for Internet topology analysis: Theory and applications*, IEEE/ACM Transactions on Networking Journal.

2. H. Haddadi, D. Fay, A. Thomason, S. Uhlig, A. Jamakovic, A.W. Moore, R. Mortier and M. Rio, 2008, *A brief history of the Internet topology: Inconsistent views of a consistent evolution*, Proceedings of the ACM CoNEXT Conference.
3. H. Haddadi, D. Fay, S. Uhlig, A.W. Moore, R. Mortier, A. Jamakovic and M. Rio, 2008, *Tunning topology generators using spectral distribution*, Proceedings of the SPEC International Performance Evaluation Workshop, Darmstadt, Germany.
4. R.E. Kooij, A. Jamakovic, F. van Kesteren, T.C.M. de Koning, I. Theisler and P. Veldhoven, 2008, *The Dutch soccer team as a social network*, Connections Journal.
5. R.E. Kooij, A. Jamakovic, F. van Kesteren, T.C.M. de Koning, I. Theisler and P. Veldhoven, 2008, *Het Nederlands elftal als complex netwerk*, Nieuw Archief voor Wiskunde Journal.
6. A. Jamakovic and S. Uhlig, 2008, *On the relationships between topological measures in real-world networks*, Networks and Heterogeneous Media Journal.
7. A. Jamakovic and P. Van Mieghem, 2008, *On the robustness of complex networks by using the algebraic connectivity*, Proceedings of the IFIP Networking Conference, Singapore City, Singapore.
8. A. Jamakovic and S. Uhlig, 2007, *Influence of network structure on robustness*, Proceedings of the 15th IEEE International Conference on Networks, Adelaide, Australia.
9. A. Jamakovic, S. Uhlig and I. Theisler, 2007, *On the relationships between topological metrics in real-world networks*, Proceedings of the European Conference on Complex Systems, Dresden, Germany.
10. A. Jamakovic and S. Uhlig, 2007, *On the relationship between the algebraic connectivity and graph's robustness to node and link failures*, Proceedings of the 3rd EURO-NGI Conference on Next Generation Internet Network, Trondheim, Norway.
11. A. Jamakovic, R.E. Kooij, P. Van Mieghem and E.R. van Dam, 2006, *Robustness of networks against viruses: the role of the spectral radius*, Proceedings of the 13th Annual Symposium of the IEEE/CVT Benelux, Liège, Belgium.

12. A. Jamakovic, H. Wang and P. Van Mieghem, 2006, *Topological characteristics of the Dutch road infrastructure*, Seminar Infrastructure Reliability, Delft, The Netherlands.
13. A. Jamakovic and P. Van Mieghem, 2006, *The Laplacian spectrum of complex networks*, Proceedings of the European Conference on Complex Systems, Oxford, UK.

NGInfra PhD Thesis Series on Infrastructures

1. Strategic behavior and regulatory styles in the Netherlands energy industry
Martijn Kuit, 2002, Delft University of Technology, The Netherlands.
2. Securing the public interest in electricity generation markets, The myths of the invisible hand and the copper plate
Laurens de Vries, 2004, Delft University of Technology, The Netherlands.
3. Quality of Service Routing in the Internet: Theory, Complexity and Algorithms
Fernando Kuipers, 2004, Delft University of Technology, The Netherlands.
4. The role of power exchanges for the creation of a single European electricity market: market design and market regulation
François Boisseleau, 2004, Delft University of Technology, The Netherlands, and University of Paris IX Dauphine, France.
5. The ecology of metals
Ewoud Verhoef, 2004, Delft University of Technology, The Netherlands.
6. MEDUSA, Survivable information security in critical infrastructures
Semir Daskapan, 2005, Delft University of Technology, The Netherlands.
7. Transport infrastructure slot allocation
Kaspar Koolstra, 2005, Delft University of Technology, The Netherlands.
8. Understanding open source communities: an organizational perspective
Ruben van Wendel de Joode, 2005, Delft University of Technology, The Netherlands.

9. Regulating beyond price, integrated price-quality regulation for electricity distribution networks
Viren Ajodhia, 2006, Delft University of Technology, The Netherlands.
10. Networked Reliability, Institutional fragmentation and the reliability of service provision in critical infrastructures
Mark de Bruijne, 2006, Delft University of Technology, The Netherlands.
11. Regional regulation as a new form of telecom sector governance: the interactions with technological socio-economic systems and market performance
Andrew Barendse, 2006, Delft University of Technology, The Netherlands.
12. The Internet bubble - the impact on the development path of the telecommunications sector
Wolter Lemstra, 2006, Delft University of Technology, The Netherlands.
13. Multi-Agent Model Predictive Control with Applications to Power Networks
Rudy Negenborn, 2007, Delft University of Technology, The Netherlands.
14. Dynamic bi-level optimal toll design approach for dynamic traffic networks
Dusica Joksimovic, 2007, Delft University of Technology, The Netherlands.
15. Intertwining uncertainty analysis and decision-making about drinking water infrastructure
Machtelt Meijer, 2007, Delft University of Technology, The Netherlands.
16. The New EU Approach to Sector Regulation in the Network Infrastructure Industries
Richard Cawley, 2007, Delft University of Technology, The Netherlands.
17. A functional legal design for reliable electricity supply, How technology affects law
Hamilcar Knops, 2008, Delft University of Technology, The Netherlands and Leiden University, The Netherlands.
18. Improving Real-Time Train Dispatching: Models, Algorithms and Applications
Andrea D'Ariano, 2008, Delft University of Technology, The Netherlands.
19. Exploratory modelling and analysis: A promising method to deal with deep uncertainty
Datu Buyung Agusdinata, 2008, Delft University of Technology, The Netherlands.

20. Characterization of Complex Networks: Application to Robustness Analysis
Almerima Jamakovic, 2008, Delft University of Technology, The Netherlands.