MEASURE OF ROBUSTNESS FOR COMPLEX NETWORKS

by

MINA NABIL YOUSSEF

B.S., Alexandria University, Egypt, 2004
M.S., Kansas State University, USA, 2008

AN ABSTRACT OF A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2012

Abstract

Critical infrastructures are repeatedly attacked by external triggers causing tremendous amount of damages. Any infrastructure can be studied using the powerful theory of complex networks. A complex network is composed of extremely large number of different elements that exchange commodities providing significant services. The main functions of complex networks can be damaged by different types of attacks and failures that degrade the network performance. These attacks and failures are considered as disturbing dynamics, such as the spread of viruses in computer networks, the spread of epidemics in social networks, and the cascading failures in power grids. Depending on the network structure and the attack strength, every network differently suffers damages and performance degradation. Hence, quantifying the robustness of complex networks becomes an essential task.

In this dissertation, new metrics are introduced to measure the robustness of technological and social networks with respect to the spread of epidemics, and the robustness of power grids with respect to cascading failures. First, we introduce a new metric called the Viral Conductance (VC_{SIS}) to assess the robustness of networks with respect to the spread of epidemics that are modeled through the susceptible/infected/susceptible (SIS) epidemic approach. In contrast to assessing the robustness of networks based on a classical metric, the epidemic threshold, the new metric integrates the fraction of infected nodes at steady state for all possible effective infection strengths. Through examples, VC_{SIS} provides more insights about the robustness of networks than the epidemic threshold. In addition, both the paradoxical robustness of Barabási-Albert preferential attachment networks and the effect of the topology on the steady state infection are studied, to show the importance of quantifying the robustness of networks. Second, a new metric VC_{SIR} is introduced to assess the

robustness of networks with respect to the spread of susceptible/infected/recovered (SIR) epidemics. To compute VC_{SIR}, we propose a novel individual-based approach to model the spread of SIR epidemics in networks, which captures the infection size for a given effective infection rate. Thus, VC_{SIR} quantitatively integrates the infection strength with the corresponding infection size. To optimize the VC_{SIR} metric, a new mitigation strategy is proposed, based on a temporary reduction of contacts in social networks. The social contact network is modeled as a weighted graph that describes the frequency of contacts among the individuals. Thus, we consider the spread of an epidemic as a dynamical system, and the total number of infection cases as the state of the system, while the weight reduction in the social network is the controller variable leading to slow/reduce the spread of epidemics. Using optimal control theory, the obtained solution represents an optimal adaptive weighted network defined over a finite time interval. Moreover, given the high complexity of the optimization problem, we propose two heuristics to find the near optimal solutions by reducing the contacts among the individuals in a decentralized way. Finally, the cascading failures that can take place in power grids and have recently caused several blackouts are studied. We propose a new metric to assess the robustness of the power grid with respect to the cascading failures. The power grid topology is modeled as a network, which consists of nodes and links representing power substations and transmission lines, respectively. We also propose an optimal islanding strategy to protect the power grid when a cascading failure event takes place in the grid.

The robustness metrics are numerically evaluated using real and synthetic networks to quantify their robustness with respect to disturbing dynamics. We show that the proposed metrics outperform the classical metrics in quantifying the robustness of networks and the efficiency of the mitigation strategies.

In summary, our work advances the network science field in assessing the robustness of complex networks with respect to various disturbing dynamics.

MEASURE OF ROBUSTNESS FOR COMPLEX NETWORKS

by

MINA NABIL YOUSSEF

B.S., Alexandria University, Egypt, 2004
M.S., Kansas State University, USA, 2008

A DISSERTATION

submitted in partial fulfillment of the
requirements for the degree

DOCTOR OF PHILOSOPHY

Department of Electrical and Computer Engineering
College of Engineering

KANSAS STATE UNIVERSITY
Manhattan, Kansas
2012

Approved by:

Major Professor
Caterina Scoglio

Copyright

Abstract

Critical infrastructures are repeatedly attacked by external triggers causing tremendous amount of damages. Any infrastructure can be studied using the powerful theory of complex networks. A complex network is composed of extremely large number of different elements that exchange commodities providing significant services. The main functions of complex networks can be damaged by different types of attacks and failures that degrade the network performance. These attacks and failures are considered as disturbing dynamics, such as the spread of viruses in computer networks, the spread of epidemics in social networks, and the cascading failures in power grids. Depending on the network structure and the attack strength, every network differently suffers damages and performance degradation. Hence, quantifying the robustness of complex networks becomes an essential task.

In this dissertation, new metrics are introduced to measure the robustness of technological and social networks with respect to the spread of epidemics, and the robustness of power grids with respect to cascading failures. First, we introduce a new metric called the Viral Conductance (VC_{SIS}) to assess the robustness of networks with respect to the spread of epidemics that are modeled through the susceptible/infected/susceptible (SIS) epidemic approach. In contrast to assessing the robustness of networks based on a classical metric, the epidemic threshold, the new metric integrates the fraction of infected nodes at steady state for all possible effective infection strengths. Through examples, VC_{SIS} provides more insights about the robustness of networks than the epidemic threshold. In addition, both the paradoxical robustness of Barabási-Albert preferential attachment networks and the effect of the topology on the steady state infection are studied, to show the importance of quantifying the robustness of networks. Second, a new metric VC_{SIR} is introduced to assess the

robustness of networks with respect to the spread of susceptible/infected/recovered (SIR) epidemics. To compute VC_{SIR}, we propose a novel individual-based approach to model the spread of SIR epidemics in networks, which captures the infection size for a given effective infection rate. Thus, VC_{SIR} quantitatively integrates the infection strength with the corresponding infection size. To optimize the VC_{SIR} metric, a new mitigation strategy is proposed, based on a temporary reduction of contacts in social networks. The social contact network is modeled as a weighted graph that describes the frequency of contacts among the individuals. Thus, we consider the spread of an epidemic as a dynamical system, and the total number of infection cases as the state of the system, while the weight reduction in the social network is the controller variable leading to slow/reduce the spread of epidemics. Using optimal control theory, the obtained solution represents an optimal adaptive weighted network defined over a finite time interval. Moreover, given the high complexity of the optimization problem, we propose two heuristics to find the near optimal solutions by reducing the contacts among the individuals in a decentralized way. Finally, the cascading failures that can take place in power grids and have recently caused several blackouts are studied. We propose a new metric to assess the robustness of the power grid with respect to the cascading failures. The power grid topology is modeled as a network, which consists of nodes and links representing power substations and transmission lines, respectively. We also propose an optimal islanding strategy to protect the power grid when a cascading failure event takes place in the grid.

The robustness metrics are numerically evaluated using real and synthetic networks to quantify their robustness with respect to disturbing dynamics. We show that the proposed metrics outperform the classical metrics in quantifying the robustness of networks and the efficiency of the mitigation strategies.

In summary, our work advances the network science field in assessing the robustness of complex networks with respect to various disturbing dynamics.

Table of Contents

Table of Contents viii

List of Symbols xi

List of Figures xii

List of Tables xiv

Acknowledgements xv

Dedication xvi

1 Introduction 1
- 1.1 Background . . . 1
  - 1.1.1 Spread of epidemics . . . 1
  - 1.1.2 Cascading failures . . . 2
- 1.2 Motivation . . . 3
  - 1.2.1 Robustness of networks with respect to spread of epidemics . . . 4
  - 1.2.2 Robustness of power grids with respect to cascading failures . . . 5
- 1.3 Contribution . . . 6
- 1.4 Organization . . . 9

2 Robustness of networks with respect to the spread of SIS epidemics 10
- 2.1 Related work . . . 12
- 2.2 Review of epidemic models . . . 13
  - 2.2.1 Epidemic spread on regular, complete bi-partite, and random networks 14
  - 2.2.2 Average infection fraction versus the epidemic threshold . . . 17
- 2.3 Viral Conductance . . . 19
  - 2.3.1 Definition of viral conductance VC_{SIS} . . . 19
  - 2.3.2 Viral Conductance versus the epidemic threshold . . . 20
  - 2.3.3 Paradoxical robustness of Barabási-Albert preferential attachment networks . . . 21
- 2.4 Properties of steady state infected population fraction . . . 22
- 2.5 Computation of VC_{SIS} and bounds . . . 26
  - 2.5.1 A heuristic for VC_{SIS} . . . 26
  - 2.5.2 Upper and lower bounds for VC_{SIS} . . . 29
- 2.6 Numerical evaluations . . . 30
  - 2.6.1 Assortative and disassortative preferential attachment networks . . . 31

2.6.2 Watts-Strogatz small world model . . . 33
2.6.3 Barabási-Albert Preferential attachment network model . . . 35
2.6.4 VC_{SIS} versus λ_{max} . . . 36
2.6.5 Internet AS-level networks . . . 37
2.6.6 Survey-based network . . . 38
2.7 Summary . . . 40

3 Robustness of networks with respect to the spread of SIR epidemics 42
3.1 Related work . . . 44
3.1.1 SIR homogeneous mean field approach . . . 45
3.1.2 SIR heterogeneous mean field approach . . . 46
3.1.3 Mitigation strategies for SIR epidemics . . . 47
3.2 Continuous time Markov chain SIR epidemic model . . . 50
3.3 Individual-based SIR approach . . . 51
3.3.1 Numerical evaluations . . . 53
3.4 Properties of the individual-based SIR approach . . . 56
3.4.1 Epidemic threshold . . . 56
3.4.2 The existence of a maximum number of infected individuals . . . 57
3.4.3 The effect of the network spectrum . . . 58
3.5 Markov chain model and mean field approximation . . . 60
3.5.1 Numerical evaluations . . . 63
3.6 Viral conductance VC_{SIR} . . . 63
3.6.1 Numerical evaluations . . . 65
3.7 Optimal mitigation of epidemics in weighted networks . . . 68
3.7.1 Weighted SIR epidemic approach . . . 69
3.7.2 Optimal dynamical weights . . . 69
3.7.3 Optimal control formulation . . . 70
3.8 Computational heuristics for the optimal mitigation strategies . . . 77
3.8.1 Bang-Bang controller heuristic . . . 78
3.8.2 Piecewise nonlinear controller heuristic . . . 78
3.8.3 Numerical evaluations . . . 79
3.9 Summary . . . 83

4 Robustness of power grids with respect to cascading failures 86
4.1 Related work . . . 89
4.2 Problem definition . . . 91
4.3 Computation algorithm of the robustness metric . . . 94
4.3.1 Power grid networks . . . 95
4.3.2 Numerical evaluations of the robustness metric . . . 96
4.4 Optimal islanding mitigation strategy for the power grids . . . 98
4.4.1 Power flow model constraints . . . 102
4.4.2 Network flow model constraints . . . 103
4.4.3 Numerical evaluation for the optimal islanding . . . 104

4.5 Summary 104

5 Conclusions and future research work **107**
5.1 Conclusions 107
5.2 Future work 109

Bibliography **123**

List of symbols

Symbol	Definition
R_0	Reproductive number
τ_c	Epidemic threshold
β	Infection rate
δ	Cure rate
τ	Effective infection rate
λ_{max}	Spectral radius
N	Number of nodes/individuals
k	Node degree
$y(t)$	Infection size at time t
s	Effective cure rate
$\overline{y_\infty}$	Average steady state infected population
$v_\infty^i(s)$	Steady state probability of infection for node i
m_o	Size of initial topology in preferential attachment networks
w	Contact rate, contact weight
$s(t)$	Fraction of susceptible nodes at time t
$i(t)$	Fraction of infected nodes at time t
$r(t)$	Fraction of recovered nodes at time t
Q	Infinitesimal matrix
$W(t)$	State vector at time t
$E[.]$	Expectation, average
V_i	Voltage at bus i
$z_{i,j}$	Impedance of transmission line (i, j)
$b_{i,j}$	Admittance of transmission line (i, j)

List of Figures

1.1 State diagram for epidemic models 5

2.1 Infection size vs the effective infection rate 16
2.2 Relationship between the reciprocal of the epidemic threshold and the average infection 18
2.3 Infection curve for different networks with the same spectral radius 20
2.4 Paradoxical robustness of preferential attachment networks 22
2.5 Infection curves of different networks with the same average degree 24
2.6 The steady state infected population for the effective cure rate is half the average connectivity 25
2.7 VC_{SIS} heuristic curve 27
2.8 Upper and lower bounds for the robustness measure 29
2.9 Robustness measure of assortative and disassortative preferential attachment networks 32
2.10 Robustness measure of Watts-Strogatz networks 34
2.11 Robustness measure of Barabási-Albert networks 35
2.12 Robustness measure versus the reciprocal of the epidemic threshold 36
2.13 Map of the weighted contact network of Clay Center, Kansas 39

3.1 Markov chain state diagram for spread of SIR epidemics 50
3.2 Evaluation of individual-based SIR approach in time domain 54
3.3 Numerical evaluation of the individual-based approach, Monte Carlo simulation and heterogeneous mean field approach 54
3.4 SIR state transition of individuals 60

3.5 Evaluation of the individual-based approach and heterogeneous mean field approach in time domain 64
3.6 Node degree correlations 66
3.7 Evaluation of individual-based approach on correlated preferential attachment networks 67
3.8 Optimal solution for the controller variable and the infection size 76
3.9 Finite time of the mitigation strategy 77
3.10 Numerical solution of the infection cases for the Bang-Bang and piecewise nonlinear controller heuristics 80
3.11 Numerical solution of the weight reduction cost for the Bang-Bang and piecewise nonlinear controller heuristics 81
3.12 Mitigation strategies for a survey-based weighted network 82

4.1 Power degradation during the cascading failure 96
4.2 Example of islanding mitigation strategy 99
4.3 Island structures 105

List of Tables

2.1 Networks with the same epidemic thresholds 21

2.2 The steady state infected population for the effective cure rate is half the average connectivity 25

2.3 Robustness measure of random and regularized networks 38

2.4 Robustness measure of mitigated social networks 40

3.1 Evaluation of the robustness measure on correlated preferential attachment networks 68

3.2 Data inputs and decision variables 70

4.1 Robustness measure with respect to the cascading failures 97

4.2 Definitions of data inputs and decision variables 100

Acknowledgments

I can do all things through Christ who strengthens me (Philippians 4:13). I am very grateful to the Lord for all His works.

I would like to thank Dr. Caterina Scoglio for all the efforts and the time that she dedicated for me. I really appreciate her academic and social support that I found all the time. I would like to thank Dr. Don Gruenbacher for all the support that he provided during my studies. I also would like to thank Dr. Robert Kooij for his research guidance during my visit to Delft University of Technology in summer 2009 and his follow up after the visit. I also appreciate all research advise from Dr. Noel Schulz, Dr. Pietro Poggi-Corradini, Dr. Piet Van Mieghem, and Dr. Huijuan Wang.

I can not forget to thank my best friends Phillip Schumm, Sakshi Pahwa, Ali Sydney, Xinyi Yu, Ranjni Chand, Karim Morcos, and the Schumm family for their constant support during the good and difficult times. I would like to thank all the members of Sunflower Networking Group and K-State Epicenter for all their efforts. I also would like to thank Prof. Medhat Morcos and Prof. Hani Melhem for encouraging me. I wish to thank Martha Schumm, and Maryam Bottros for reviewing the dissertation.

I would like to express my gratefulness to my parents Nabil and Angel and my sisters Miral and Mariam for their great support and their prayers. I also would like to thank my uncle Father Andrew Khalil and my aunt Nadia Khalil who are my mentors. I cannot forget to thank my talented cousin John Farag for all his support. I would like to thank my lovely fiancée Maryam Bottros for all her significant support.

Pope Shenouda III really played a major role in my life through his constant prayers, his love and support to his people and to the Coptic church. He passed away while I was writing this acknowledgment on Saturday March 17, 2012 - Paremhat 8, 1728 Coptic calendar.

Dedication

To the Coptic Orthodox Church

To Pope Shenouda III

To all the victims during the Egyptian Youth revolution

To Khaled Saeed, Mariam Fekry, Sally Zahran, and Mina Daniel

To Father Wesa the Syrian

To my family

Chapter 1

Introduction

1.1 Background

Large networks can suffer from dynamics that degrade their performances leading to extreme damages. The disturbing dynamics exist due to either random failures of one or more network elements, or targeted attacks aiming to cause the largest amount of damages. In both cases, the amount of damage mainly depends on the attack strength and the network structure. Large networks can represent infrastructures, which are modeled using complex network theory. A complex network is composed of many elements that are interconnected together showing heterogeneous properties. For instance, computer networks are composed of routers and communication links, social networks are composed of individuals who have contacts among them, and power grids are composed of substations that are interconnected through transmission lines. Among the various scenarios of attacks that cause damages for complex networks, the spread of viruses and cascading failures are considered among the most harmful disturbing dynamics that can take place on a complex network.

1.1.1 Spread of epidemics

The spread of epidemics caused by large scale attacks has become a crucial issue affecting modern life. For example, in 2001 and 2004, respectively, the Code Red and Sasser computer viruses infected numerous computers, resulting in costly global damages. Also, the Trojan

computer viruses have caused damages to numerous computing machines that are used for different purposes. Therefore, it is essential to study the spread of viruses and epidemics in networks through mathematical models to predict the damages.

Epidemics are usually studied via compartmental models, where individuals or nodes can be in different states, such as susceptible, infected, or recovered. On one hand, for infections like the common cold among humans in social networks and malware infections in computer networks, the susceptible/infected/susceptible (SIS) epidemic model [43, 77, 106] is considered appropriate, since individuals are again susceptible after contracting the infection and recovering. On the other hand, Influenza-like and other human contagious diseases are better modeled by the susceptible/infected/recovered (SIR) process, where immunity can be obtained after contracting the disease. These compartmental models are also integrated with prediction methods to estimate the infection parameters given the number of confirmed infected cases during pandemics and epidemics. They can also be utilized to predict the peak of the infected cases and to define and test efficient mitigation strategies [25, 102].

1.1.2 Cascading failures

Another type of disturbing dynamics is the cascading failure in power grids. During the last decade, many electric power grids have suffered from enormous cascading failures leading to major disasters in different parts of the world such as the cascading failure that took place in the states of Arizona and California in 2011, the black out in Brazil in 2009, and the famous blackout in North America in 2003. Such disasters resulted in blackouts that left millions of people without electrical energy.

Many types of triggers can disturb the normal functionality of the electric grid including but not limited to the voltage dips (voltage sags), brief voltage increases (swells), and transient events. In addition to the voltage faults that can harm the control devices and motor speeds, the instability of the frequency of generated voltage with large deviation may lead

to asynchronization of the generators, and in turn, to a dramatic reduction of the electric power causing blackouts. Moreover, the weather storms and lightening may lead to shutting down some substations and damaging power transmission lines.

1.2 Motivation

Due to the existence of different types of attacks that disturb the performance of complex networks, it is vital to propose new metrics reflecting the robustness of networks with respect to different disturbing dynamics. In particular, the new metrics proposed in this dissertation give insights about the ability of a network to persistently function while its performance is degraded due to the presence of a disturbing dynamic with a given attack strength. Thus, we consider the disturbing dynamics as dynamical systems having inputs and corresponding outputs. The inputs represent the attack strengths, while the outputs represent the amount of damages in the network. Previous efforts focused on assessing the robustness of networks based on a single input, which usually represent a critical attack strength. Such critical attack strength is described as the maximum strength that a network can sustain without any damage. However, assessing the robustness of networks based on a single attack strength point is not efficient, since it does not consider the whole attack range that a network can experience. In addition to the critical attack strength, previous efforts focused on quantifying the robustness of networks based on the amount of damages. However, this method neglects both the effect of the network structure on the damages and the attack strength. Therefore, quantifying the robustness of networks is not a trivial task, and proposing new metrics becomes vital in complex network theory. Below, we highlight the challenges in quantifying the robustness of networks with respect to both the spread of SIS and SIR epidemics in technological and social networks, and the cascading failures in power grids.

1.2.1 Robustness of networks with respect to spread of epidemics

The epidemic spreading mechanism is mathematically represented via compartmental models such as SIS and SIR models [7, 57]. The SIS model describes how an epidemic spreads in the network, and its dynamic is described by any change of a node's state with time. Thus, an infected node infects any susceptible neighbor with infection rate β, which is the number of infection trials per unit time. Then, the infected node cures itself with cure rate δ, and it becomes susceptible to a new infection. The ratio between β and δ is called the effective spreading rate τ. The epidemic threshold τ_c is a function of network characteristics, and it is a specific value of the effective spreading rate above which an epidemic outbreak takes place. When an epidemic outbreak takes place in the network, a persistent fraction of infected population exists at the steady state, and this fraction does not depend on the initial condition of the infection in the network, but it only depends on the effective infection rate and the epidemic threshold. Similar to SIS model, in the SIR model, each individual can be in one of the three states, namely, S susceptible, I infected, and R recovered (cured). To clarify, an infected individual infects its susceptible neighbors with rate β. Also, an infected individual can cure itself with a cure rate δ. The curing process represents either the death (removal) or the complete recovery of the individual after the infection. Again, the ratio between β and δ is called the effective infection rate, and its reciprocal is the effective cure rate ($\frac{\delta}{\beta}$). The epidemic threshold τ_c in the SIR model is a specific effective infection rate value, and it is a function of the network characteristics. Figure 1.1 shows the state diagram for the SIS and the SIR epidemic models.

Based on the epidemic models, assessing the robustness of complex networks not only allows us to compare the robustness among different network structures, but also gives insights about mitigation strategies and the design of future network infrastructures. However, even given the network topology, assessing the robustness of networks is difficult. On one hand, one way to measure the impact of the epidemic is by the number of infected nodes, which is not explicitly related to the topological characteristics of the networks. On the other hand,

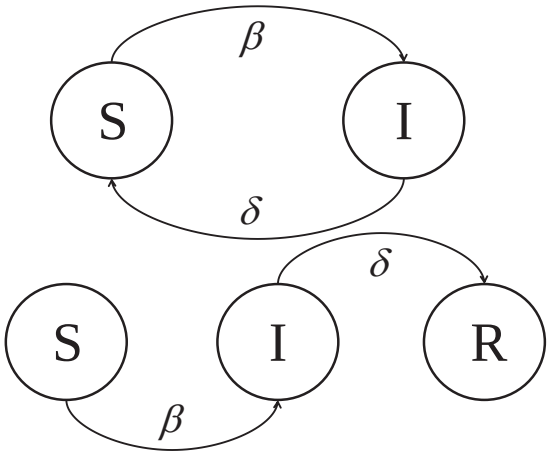


Figure 1.1: *The state diagram for both the Susceptible/Infected/Susceptible SIS (top diagram) and the Susceptible/Infected/Recovered SIR (bottom diagram) epidemic models given the transition rates β and δ.*

the epidemic threshold τ_c, used as a measure for the network robustness in the literature [50, 91], does not consider the number of infected nodes. Ultimately, the epidemic threshold is not adequate for assessing the network robustness. Therefore, it becomes necessary to propose a new robustness measure that integrates both various infection strengths and the corresponding infection sizes.

1.2.2 Robustness of power grids with respect to cascading failures

A power grid can be characterized by its topology and the power flows on it. The topology represents the connectivity of substations (generation, transmission and distribution), while the power flows represent the dynamics that are controlled by the electrical characteristics of the grid, and are delivered from the generators to the distribution substations. The electrical characteristics of the grid are the capacity and the inductance of transmission lines, the voltage values, the differences between the voltage phase angles at the terminals of each transmission line, and the loads at the distribution substations. A cascading failure takes place when a series of single or multiple faults happen in a grid. Due to the line fault(s), the power flow is redistributed on the existing transmission lines leading to overloading other lines, and the faulty lines are tripped from the grid. The redistribution of the power flow is

called transmission loading relief procedures (TLR) [80]. Therefore, during a cascading failure, the stress on the transmission lines increases causing blackout. Since any transmission line can be tripped due to a fault, it becomes very critical to predict the amount of failures and power loss that the grid can experience. Therefore, the robustness of the grid can not be measured only through the topology of the grid, but the power flows in the grid have to be involved as well.

1.3 Contribution

In this dissertation, we propose new metrics to assess the robustness of complex networks with respect to the spread of epidemics, and the cascading failures.

Firstly, we propose the Viral Conductance VC_{SIS} and VC_{SIR} to assess the robustness of networks with respect to the spread of SIS and SIR epidemics, respectively. The new metrics are applied to different types of networks to study the effect of network structure on the spread of epidemics. In addition, the new metrics are used to quantify the performance of efficiency of different mitigation strategies in reducing the infection size. Differently from the classical methods that are used in the literature, viral conductance distinguishes the robustness of networks, which have similar topological characteristics such as average connectivity, spectral radius, and node degree distribution.

Secondly, we propose a new individual-based SIR approach to evaluate the spread of epidemics in networks. Basically, the new approach is based on the continuous time Markov chain model in which the state of every node is independently studied in the network. The individual-based SIR approach is utilized to evaluate the total infection size for a given effective cure rate, which are used to evaluate the robustness metric VC_{SIR}. In addition, the new approach is used to confirm that the epidemic threshold is the reciprocal of the network spectral radius.

Thirdly, we present an optimal mitigation strategy to the spread of SIR epidemics in social contact networks. In a social network, every individual has contact strength with other

individuals in the network. The contact strength between any two individuals is quantified through the frequency of contact, which is presented as a weight ranging between 0 and 1 in the contact network. The mitigation strategy is based on reducing the contacts among the individuals during the prevalence of an epidemic, while maintaining a minimum social contact rate. Therefore, the objective of the mitigation strategy is to minimize both the total infection size, and the contact reduction among the individuals. In addition to the obtained optimal solution, we propose two heuristics to find near optimal solutions in a decentralized way. In the first heuristic, the contact rate between any two individuals switches between its normal value and the minimum value, while in the second heuristic, the contact rate is continuously reduced over a period of time. For both heuristics, the normal contact rates are retrieved to avoid large contact reduction cost. Extensive numerical simulations are performed on synthetic networks as well as a survey based network to evaluate the performance of the mitigation strategies.

Finally, a new metric is proposed to quantify the robustness of power grids. The new measure mainly depends on number of cascading stages and the frequency of link failures due to the initial removal of a single transmission line from the grid. The DC power flow model is utilized to evaluate the power flow in the grid. The new measure is applied to real and synthetic power grids for evaluations. Moreover, a new strategy is proposed to mitigate the power grid from cascading failure events. The new strategy is based on the grid islanding in which, the optimal island topologies are found such that the maximum electric loads are satisfied in the island and in the remaining part of the grid. The mitigation strategy is applied to real networks showing the structure of the islands, the amount of load shedding, and the power generation reduction.

We summarize the contributions as follows:

- Robustness of networks with respect to the spread of SIS epidemics
  - Introducing a new robustness measure VC_{SIS} with respect to the spread of SIS

epidemics

  - Comparing VC_{SIS} with the epidemic threshold and showing cases where the epidemic threshold fails to assess the robustness of networks
  - Introducing a computational heuristic for VC_{SIS} and upper and lower bounds
  - Presenting a framework to compare the robustness of Watts-Strogatz (WS) networks and Barabási-Albert (BA) preferential attachment (PA) networks
  - Applying VC_{SIS} to a weighted social network to evaluate the efficiency of mitigation strategies

- Robustness of networks with respect to the spread of SIR epidemics
  - Proposing an individual-based SIR approach
  - Outlining the role of the eigenvalues and the eigenvector centrality in the spread of epidemics
  - Validating the new individual-based SIR approach and providing guidelines on its accuracy
  - Extending the individual-based approach to propose a new robustness metric VC_{SIR} with respect to the spread of SIR epidemics
  - Proposing an optimal control framework for mitigation strategies to SIR epidemics in contact networks
  - Considering realistic scenarios in social networks
  - Proposing two heuristics to find near optimal solutions

- Robustness of power grids with respect to cascading failures
  - Proposing a new metric to quantify the robustness of the electric power grids
  - Utilizing the power flow model and the electric parameters in assessing the robustness of the grid

- Outlining the role of the link survival probability and the depth of the cascading failure
- Proposing an optimal islanding mitigation strategy in power grids

1.4 Organization

The robustness of complex networks is discussed in three main chapters. In chapter 2, we introduce the metric viral conductance VC_{SIS} to measure the robustness of networks with respect to the spread of SIS epidemics. The robustness of well-known network models in the literature is discussed, and we highlight the importance of the new metric with respect to the classical metric, epidemic threshold. In chapter 3, we propose a new individual-based model to SIR epidemics. In addition, the viral conductance is extended to propose a new measure VC_{SIR} to quantify the robustness of networks with respect to the spread of SIR epidemics. Moreover, a new optimal mitigation strategy to the spread of SIR epidemics is proposed using optimal control theory. Chapter 4 addresses the robustness of power grids with respect to cascading failures. Finally, chapter 5 concludes the dissertation, and guidelines for future research work are also discussed in detail.

Chapter 2

Robustness of networks with respect to the spread of SIS epidemics

Complex networks own conductivity properties for many types of dynamics, which are classified as good dynamics and disturbing dynamics. Good dynamics carry the actual provided services, while disturbing dynamics degrade the network performance. Among the disturbing dynamics, the spread of viruses can take place in networks. Thus, a connected network has a viral conductivity property, which depends on the spreading strength and the self curing process of the network elements. In this chapter, we propose a novel metric Viral Conductance (VC) [60, 110] to assess the robustness of networks with respect to the spread of SIS epidemics. The new metric considers the value of epidemic threshold and the number of infected nodes at steady state above the threshold. Thus the higher the value of VC, the lower the robustness of the network. One very interesting result obtained using VC, concerns the robustness of Barabási-Albert (BA) preferential attachment (PA) networks. While previous work stressed the fact that the epidemic threshold is close to zero in those networks, VC can quantify the fraction of infected nodes for increasing value of the effective spreading rates above the threshold. As a matter of fact, infinite Barabási-Albert (BA) preferential attachment (PA) networks with vanishing epidemic thresholds can still require strong epidemics to have major outbreaks, while in homogeneous random networks, an epidemic does not need to be very much beyond the threshold to cause major outbreaks. Due

to the computational complexity of the new metric, we propose a heuristic to actually compute the VC for large networks. Moreover, we derive upper and lower bounds for VC. We performed extensive simulations to validate the new metric, the bounds, and the heuristic. The numerical evaluation focuses on assessing the robustness of different types of networks, e.g. correlated preferential attachment (PA) networks, Watts-Strogatz (WS) small world networks [107] and Barabási-Albert (BA) preferential attachment (PA) networks [10]. Our results show that assortative PA networks have lower VC than disassortative PA networks when only the average node degree is preserved. Qualitatively, we compare between the robustness of Watts-Strogatz (WS) networks and Barabási-Albert (BA) preferential attachment (PA) networks, and we show cases where Watts-Strogatz (WS) networks are less robust than Barabási-Albert (BA) preferential attachment (PA) networks and vice versa. In addition, for a given irregular network, we rewired the network to make it almost regular by decreasing the variance of the node degrees, and we computed the VC before and after the regularization. Finally, the new robustness metric is used to evaluate the performance of different mitigation strategies for social networks.

The chapter is organized as follows: In Section 2.1, we review the literature and related work. In Section 2.2, we review the mathematical models of epidemic spread and their applications to different types of networks, and we compare the epidemic threshold with average fraction of infection. In addition, we propose the new robustness metric with respect to epidemic spread, the viral conductance VC, we compare between the classical measure and VC, and we discuss the robustness of Barabási-Albert (BA) preferential attachment networks in Section 2.3. In Section 2.4, a summary of some useful properties of infected nodes at steady state are presented. In Section 2.5, we propose upper and lower bounds,and a heuristic to reduce the computational complexity of the viral conductance. The numerical results are discussed in Section 2.6, and finally the chapter is summarized in Section 2.7.

2.1 Related work

The history of epidemic modeling in biology dates back to the eighteenth century when Bernoulli proposed the first deterministic epidemic model for smallpox, which was followed by other deterministic and stochastic epidemic models during the last two centuries [6, 7, 67, 73]. Those models introduced the basic reproductive number R_0, which is the average number of secondary infections due to a single infected case in the population. If $R_0 > 1$, the epidemic spreads in the population, while if $R_0 < 1$, the epidemic dies out without causing an outbreak. In complex network theory, the work in [90, 91] found the epidemic threshold, which is a function of R_0, for SIS epidemics using the heterogeneous mean field approach. The authors of [90, 91] conclude that the threshold is a function of the heterogeneity of the network represented by the second moment of the node degree distribution. Later, an exact expression of the epidemic threshold emerged in the framework of spectral graph theory in [31, 77]. The work in [91] also concludes that epidemics can spread on any scale-free network regardless of its effective infection rate when the number of nodes in the network approaches infinity. In other words, all scale-free networks are vulnerable to epidemic spread. However, finite scale-free networks have a non-zero epidemic threshold. Therefore, they are not always vulnerable to epidemic attacks.

Recently, biological epidemic models have been adapted to study the spread of viruses in technological networks as shown in [21, 55, 77, 106]. For example, in [55], the authors propose deterministic and probabilistic models to predict the size of the infected population in homogeneous networks. Unfortunately, the models do not consider the heterogeneous structure of networks, and hence, they can not be used to measure the robustness of generic networks with respect to the spread of epidemics. In [106], the authors propose a discrete-time epidemic model to predict the infection size, finding that the epidemic threshold is the reciprocal of the largest eigenvalue of the network adjacency matrix. Meanwhile, the N-Intertwined model in [77] reproduced the epidemic threshold in [106] based on a continuous time Markov chain process. Also the N-Intertwined model explicitly reveals the role of

the network in spreading epidemics. However, it does not introduce a metric to assess the robustness of networks.

In summary, using only the epidemic threshold as a measure for robustness is common practice: the larger the epidemic threshold, the more robust a network is against the spread of an epidemic, as in [50]. However, in Section 2.3, we will show some examples of networks where the epidemic threshold fails to assess their robustness, properly. Moreover, none of the works in the literature combine both the effective spreading rate and the infected population to describe the robustness of the network with respect to epidemic spreading. Specifically, then, we show that the robustness of networks depends on the value of the epidemic threshold τ_c (or its reciprocal) as well as the fraction of the infected population above the threshold.

2.2 Review of epidemic models

In this section, we review basic results about the spread of Susceptible-Infected-Susceptible (SIS) epidemics on networks. The SIS infection model, which arose in mathematical biology, is often used to model the spread of epidemics [7, 38, 55, 82, 90], epidemic algorithms for information dissemination in unreliable distributed systems like P2P and ad-hoc networks [20], [34], and propagation of faults and failures in BGP networks [23]. The SIS model analytically reveals how a node's state changes between the two S and I states in complex networks. To clarify, during the spread of an epidemic, a node is in one of the two states. First, an infected node can infect susceptible neighbors with infection rate β. Also, an infected node can cure itself with a cure rate δ and become susceptible to re-infection. Additionally, the ratio between β and δ is called the effective infection rate, $\tau = \frac{\beta}{\delta}$. Moreover, the epidemic threshold τ_c can be defined as follows: for effective spreading rates below τ_c, the epidemic contamination in the network dies out; the mean epidemic lifetime is of order $log\ n$, while for effective spreading rates above τ_c, the epidemic is prevalent, i.e. a persistent fraction of nodes remains infected with the mean epidemic lifetime [38] of the

order e^{n^α}, for τ sufficiently large. Thus, the epidemic threshold was found to be $\tau_c = \frac{1}{\lambda_{max}}$, i.e. the reciprocal of the spectral radius λ_{max} of the network adjacency matrix [77, 106]. Another basic epidemic model in the literature is the SIR [57], which differs from the SIS model in many aspects. In the SIR model, the susceptible node becomes infected and later recovers without being susceptible to further infection. Therefore, this model has no steady state infected population fraction since all infected nodes recover, while in the SIS model, a steady state infected population fraction exists, and it depends on the effective infection rate and the network structure. Below, we review the important results of the spread of SIS epidemics on regular, bi-partite, and random networks, and we also discuss how the epidemic threshold is related to the average fraction of infection at steady state.

2.2.1 Epidemic spread on regular, complete bi-partite, and random networks

We compare the fraction of infected nodes in the SIS model for different example networks and show that the value of the effective spreading rate τ determines for which network this fraction is higher. The first and second example networks belong to the class of regular and complete bi-partite networks, respectively, while the last example addresses two different types of random network.

Epidemic spread on regular networks

This model is based on a classical result for SIS models by Kephart and White [55, 58]. We consider a connected network on N nodes where every node has degree k. We denote the number of infected nodes in the population at time t by $Y(t)$. If the population N is sufficiently large, we can convert $Y(t)$ to $y(t) = Y(t)/N$, yielding a continuous quantity representing the fraction of infected nodes. Now the rate at which the fraction of infected nodes changes is due to two processes: susceptible nodes become infected and infected nodes become susceptible. Obviously, the rate at which the fraction $y(t)$ grows is proportional to the fraction of susceptible nodes, i.e. $1 - y(t)$. Therefore, for every susceptible node, the

rate of infection is the product of the infection rate per node (β), the degree of node (k), and the probability that on a given link the susceptible node connects to an infected node ($y(t)$). Therefore, we obtain the following differential equation describing the time evolution of $y(t)$:

$$\frac{dy(t)}{dt} = \beta k y(t)(1 - y(t)) - \delta y(t) \tag{2.1}$$

The steady state solution $y_\infty(\tau)$ of Eq. 2.1 satisfies

$$y_\infty(\tau) = \frac{\beta k - \delta}{\beta k} = 1 - \frac{1}{\tau k} \tag{2.2}$$

Because an epidemic state only exists if $y_\infty(\tau) > 0$, we conclude that the epidemic threshold satisfies

$$\tau_c = \frac{\beta}{\delta} = \frac{1}{k} \tag{2.3}$$

Because for k-regular networks, the spectral radius of the adjacency matrix is equal to k, see [28], Eq. 2.3 is in line with the result in [106].

Epidemic spread on complete bi-partite networks

A complete bi-partite network $K_{M,N}$ consists of two disjointed sets containing respectively M and N nodes, such that all nodes in one set are connected to all nodes in the other set, while within each set no connections occur. Notice that (core) telecommunication networks often can be modeled as a complete bi-partite topology. For instance, the so-called double-star topology (i.e. $K_{M,N}$ with $M = 2$) is quite commonly used because it offers a high level of robustness against link failures. For example, the Amsterdam Internet Exchange (see www.ams-ix.net), one of the largest public Internet exchanges in the world, uses this topology to connect its four locations in Amsterdam to two high-density Ethernet switches. Sensor networks are also often designed as complete bi-partite networks.

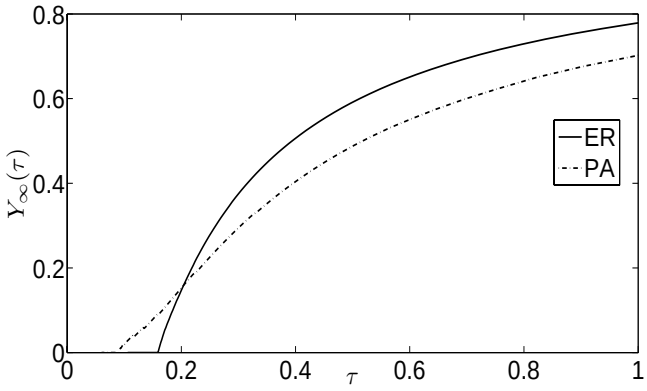


Figure 2.1: *Fraction of infected nodes for an ER network and a Barabási-Albert (BA) preferential attachment (PA) network with* $N = 10^4$.

In [86] and [77], a model for epidemic spreading on the complete bi-partite network $K_{M,N}$ is presented. Using differential equations and two-state Markov processes, the authors show that above the epidemic threshold $\tau_c = \frac{1}{\sqrt{MN}}$, the fraction of infected nodes at the steady state for $K_{M,N}$ satisfies

$$y_\infty(\tau) = \frac{(MN\tau^2 - 1)((M+N)\tau + 2)}{\tau(M+N)(M\tau+1)(N\tau+1)} \tag{2.4}$$

It is easy to verify that for the case $M = N$, Eq. 2.4 reduces to Eq. 2.2, with $k = N$.

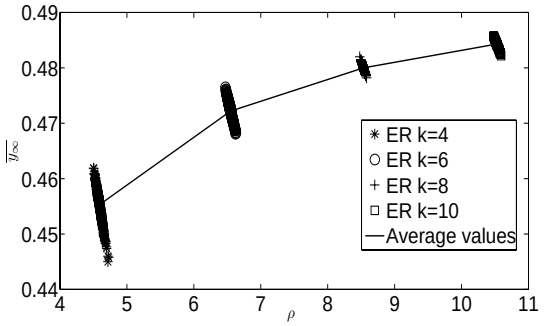
Epidemic spread on random networks

Many classes of random networks are described by their statistical properties, for example Erdös Rényi (ER), Watts-Strogatz (WS) and Barabási-Albert (BA) preferential attachment (PA) networks. Erdös-Rényi networks (ER) are homogeneous networks with binomial node degree distribution and average node degree k. In Erdös Rényi networks (ER), every node selects its neighbors with probability p independently. For large number of nodes N and for $pN = k$, the node degree distribution becomes Poisson distribution with a tail that decays exponentially for large node degrees. Meanwhile, Barabási-Albert networks BA are built using the preferential attachment (PA) mechanism in which each node prefers to connect with high node degree neighbors, and therefore the node degree distribution follows the

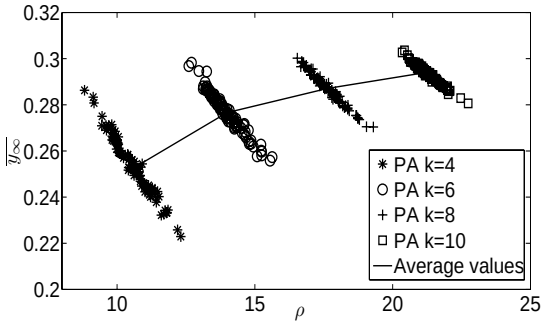
power law distribution. We will use PA and BA equivalently to refer to Barabási-Albert preferential attachment networks. The literature shows that large BA networks are scale-free (SF) networks, and consequently are the most vulnerable networks to spread of epidemics. However, we study a counter example in which an Erdös Rényi (ER) network can be more vulnerable to spread of epidemics than Barabási-Albert (BA) preferential attachment (PA) network given the same number of nodes. Figure 2.1 shows the fraction of infected nodes at steady state $y_\infty(\tau)$ due to the spread of SIS epidemics for different effective infection rates $\tau = \frac{\beta}{\delta}$ on an ER network and a PA network with 10^4 nodes. We observed that the epidemic threshold of PA is smaller than the epidemic threshold of an ER network showing that PA network is more vulnerable than ER network. However, there is an inversion in $y_\infty(\tau)$ curves and after the inversion, the ER network has higher steady state infection fraction than the PA network. In this region, the ER network is more vulnerable than the PA network. Thus, the two networks have two opposing features and it is not trivial to measure their robustness. For a given network, because the range of τ values for which the epidemic prevails is infinitely large, from now on, instead of considering the effective spreading rate τ, we look at the reciprocal of τ, that is the effective curing rate $s = \frac{\delta}{\beta}$. We are interested in $y_\infty(s)$, the fraction of infected nodes in steady state, as a function of the effective curing rate. Note that the behavior of $y_\infty(s)$ around $s = 0$ reflects the behavior of the original system for $\tau \to \infty$. Moreover, the value of $y_\infty(s)$ for $s = \lambda_{max}$ (the reciprocal of the epidemic threshold) is 0. Such a conversion leads to a closed area under the $y_\infty(s)$ curve. We denote $\overline{y_\infty}$ to be the average fraction of infection at steady state defined over $0 \leq s \leq \lambda_{max}$ and it is given by $\overline{y_\infty} = \frac{1}{\lambda_{max}} \int_0^{\lambda_{max}} y_\infty(s) ds$.

2.2.2 Average infection fraction versus the epidemic threshold

Next, we hypothesize a case study, which shows that the epidemic threshold is not capable to assess the robustness of networks having the same average node degree k within the same network class. Assume that the epidemic can have any effective cure rate $0 \leq s = \frac{\delta}{\beta} \leq \lambda_{max}$,



(a) Erdös Rényi (ER) networks



(b) Barabási-Albert (BA) preferential attachment (PA) networks

Figure 2.2: *The relationship between λ_{max} and $\overline{y_\infty}$ for Erdös-Rényi (ER) and Barabási-Albert (BA) preferential attachment (PA) networks with $N = 10^4$ and $k = 4$, 6, 8 and 10. For network class and every average node degree k, there are 100 samples of networks. The solid line represents the average values of λ_{max} and the average values of $\overline{y_\infty}$.*

and every effective cure rate has a probability of infection at steady state $y_\infty(s)$. The average value of $y_\infty(s)$ over the defined range of s is $\overline{y_\infty}$ infection fraction. Figure 2.2 shows how the average infection fraction $\overline{y_\infty}$ is inversely proportional to the maximum eigenvalue λ_{max} (reciprocal of epidemic threshold) for networks that not only belong to the same class, but also have the same average node degree. Moreover, in Figure 2.2(b), many networks with low average node degree and low maximum eigenvalue have a higher average infection fraction than other networks with high maximum eigenvalues and high average node degree. However, the average infection fraction shows a general tendency to increase with the maximum eigenvalue over different average node degree values. This tendency reflects that as the number of links increases, the average node degree increases too, so the chance for an outbreak to take place becomes more relevant and causes a higher number of incidences at steady state.

We conclude that neither the fraction of infection at steady steady nor the epidemic threshold can solely describe the robustness of networks with respect to spread of SIS epidemics in networks. Therefore, we need a new metric to quantify the robustness of networks considering both the fraction of infection at steady state and the epidemic threshold.

2.3 Viral Conductance

Based on the above conclusions and the mathematical background, in this section, we propose a new metric to quantify the robustness of networks with respect to spread of epidemics.

Definition 2.3.1. *Viral Conductance VC is a robustness measure of a given network G with respect to the spread of epidemics. It represents the average fraction of infected nodes for all types of epidemic attacks that are capable of producing outbreaks in the network.*

Consequently, the definition of viral conductance is applied to any type of epidemic. In this Chapter, we extend the concept of viral conductance to be the robustness measure of networks with respect to the spread of SIS epidemics, and we denote the new measure by VC_{SIS}.

2.3.1 Definition of viral conductance VC_{SIS}

Because we are considering the effective curing rate s as an independent variable, the area under the curve $y_\infty(s)$ is bounded. A logical way to consider the range of s values as well as the fraction of infected nodes is to evaluate the area under the $y_\infty(s)$ curve. We can now introduce a new robustness measure with respect to epidemic spread, viral conductance VC_{SIS}, that takes into account all values of s, and hence τ.

Mathematically, for non-negative continuous variable, $y_\infty(s) = \frac{1}{N}\sum_{i\in N} v_\infty^i(s)$ where $v_\infty^i(s) = \frac{\sum_{j\in neighbors(i)} v_\infty^j(s)}{s+\sum_{j\in neighbors(i)} v_\infty^j(s)}$ is the probability that node i is infected at the steady state [77]. Therefore, VC_{SIS} is defined as follows:

$$VC_{SIS} = \int_0^{\lambda_{max}} y_\infty(s)ds = \lambda_{max}\overline{y_\infty} \tag{2.5}$$

where λ_{max} is the spectral radius (i.e. maximum eigenvalue) of the adjacency matrix of network G and $\overline{y_\infty}$ is the average value of the fraction of infected nodes for all $0 \le s \le \lambda_{max}$.

We will now state two theorems for the viral conductance VC_{SIS}, of a network G.

Theorem 2.3.2. *For regular networks G_k, where every node has k neighbors, $VC_{SIS}(G_k) = \frac{k}{2}$.*

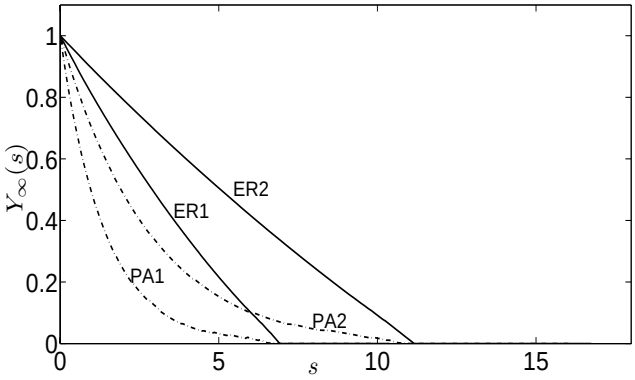


Figure 2.3: *Two examples of ER and PA networks with $N = 10^4$ and with the same λ_{max} (i.e. the same epidemic threshold).*

Proof. This follows directly from Eq. 2.2. □

Theorem 2.3.3. *For complete bi-partite networks $K_{M,N}$, $VC_{SIS}(K_{M,N})$ is as follows:*

$$\begin{aligned} VC_{SIS}(K_{M,N}) = & \frac{(M+N)\sqrt{MN} - MN}{M+N} \\ & + \frac{(M-N)(Nln(M+\sqrt{MN}))}{M+N} \\ & - \frac{(M-N)(Mln(N+\sqrt{MN}))}{M+N} \\ & + \frac{(M-N)(MlnM - NlnN)}{M+N} \end{aligned} \tag{2.6}$$

Proof. This follows from applying Eq. 2.5 to Eq. 2.4. □

The viral conductance VC_{SIS} can also be applied to random networks. For example in Figure 2.1, the values of VC_{SIS} for Erdös-Rényi (ER) and Barabási-Albert (BA) preferential attachment (PA) networks are 5.103 and 2.887, respectively. In this case, PA is more robust than ER.

2.3.2 Viral Conductance versus the epidemic threshold

Traditionally, the epidemic threshold has been used to evaluate the robustness of networks with respect to spread of epidemics [50, 91]. However, we present a case-study in which the

Metric	ER1	PA1	ER2	PA2
λ_{max}	7.16	7.16	10.98	10.98
VC_{SIS}	3.093	1.652	5.097	2.672

Table 2.1: *The values of VC_{SIS} of the networks in Figure 2.3 with the same λ_{max} values.*

epidemic threshold fails to assess the robustness of networks. Figure 2.3 show examples of networks with $N = 10^4$ that almost have the same maximum eigenvalues (i.e. same epidemic thresholds), in which the epidemic threshold thus fails to accurately assess their robustness. Note that the difference between the maximum eigenvalue values in every pair of networks, $\{ER1, PA1\}$ and $\{ER2, PA2\}$, is very small and in the order of 10^{-4}. Figure 2.3 represents two ER networks and two PA networks. ER1 and ER2 networks have the same epidemic threshold as PA1 and PA2 networks, respectively.

Table 2.1 shows the numerical values of λ_{max} and VC_{SIS} of the four networks discussed in Figure 2.3. We notice that the value of λ_{max} for both ER1 and PA1 networks is the same, while the VC_{SIS} value of ER1 network is higher than the VC_{SIS} value of PA1 network. The same observation is applied on ER2 and PA2 networks. The difference between VC_{SIS} and λ_{max} is that VC_{SIS} represents the area under the $y_\infty(s)$ curve.

2.3.3 Paradoxical robustness of Barabási-Albert preferential attachment networks

We addressed the robustness of Erdös-Rényi (ER) networks and Barabási-Albert (BA) preferential attachment (PA) networks in Figure 2.1 and we showed that there are two opposing features relate to the epidemic threshold and the fraction of infection at steady state. We summarize the opposing features as follows: *On one hand, the PA network has a lower epidemic threshold than the ER network showing that the PA network is more vulnerable than ER network. On the other hand, the ER network has a higher fraction of infection at steady state than does the BA network.* Therefore, looking independently look at the epidemic threshold or the steady state infection fraction to measure the robustness of networks is not sufficient. Additionally, to address the paradoxical robustness of PA networks, let us study

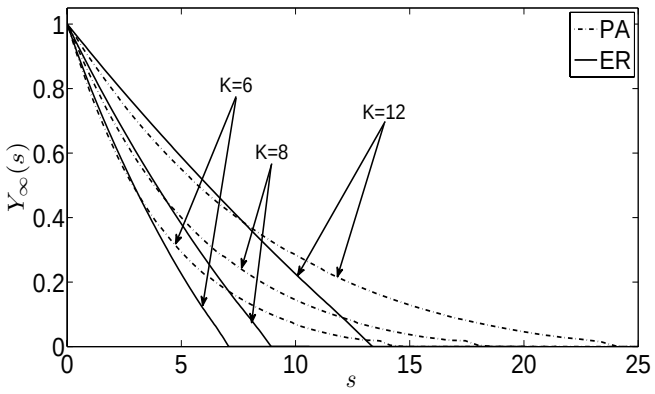


Figure 2.4: *Examples of the fraction of infection at steady state $y_\infty(s)$ given different values of $s = \frac{\delta}{\beta}$ for both ER and PA networks with $N = 10^4$ and $k = 6,\ 8,$ and 12.*

the steady state infection curves for finite ER and PA networks as shown in Figure 2.4. Because of the shrinking tail behavior of the fraction of infection at steady state, PA networks with very high λ_{max} (very low epidemic threshold) would still have tiny fractions of infected population within the region beyond and away from the maximum eigenvalue (reciprocal of the threshold). Infinite PA networks with large maximum eigenvalues (vanishing epidemic threshold) can still require strong epidemics to have major outbreaks, while in ER networks, an epidemic does not need to be very much beyond the reciprocal of the epidemic threshold to cause a major outbreak.

2.4 Properties of steady state infected population fraction

Here, we summarize basic properties of the steady state infected population $y_\infty(s)$ presented in the literature, which are very useful for the computational heuristic for VC_{SIS} in Section 2.5. We also show new results and ideas based on the basic properties of an infected population at steady state. First, the role of the epidemic threshold was studied in the literature and found that for any connected network G, denote λ_{max} is the maximum eigenvalue of the adjacency matrix, such that $y_\infty(\lambda_{max})$=0. Secondly, the fraction of infected

population at the steady state $y_\infty(s)$ can be written as $y_\infty(s) = 1 - \sigma s + O(s^2)$, where $\sigma = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{d_i}$ and d_i is the degree of node i. Lastly, given a regular network with node degree $d = k$, the fraction of infected nodes $y_\infty(s)$ for $s = \frac{k}{2}$ equals 1/2. This result directly follows from Eq. 2.2. For further reading and proofs, we refer the reader to [77]. Below, we show new results about $y_\infty(s)$ for $s = \frac{k}{2}$ (i.e. $\tau = \frac{2}{k}$) for any network with average node degree k.

Lemma 2.4.1. *For any complete bi-partite network and for $s = \frac{k}{2}$, the fraction of infected nodes $y_\infty(\frac{k}{2})$ is bounded by 0.5147.*

Proof. Let the total number of nodes in any complete bi-partite network be $T = M + N$. Then, substituting $M = T - N$ in Eq. 2.4 for $s = \frac{k}{2} = \frac{MN}{M+N}$, we obtain the following equation:

$$y_\infty(\frac{k}{2}) = \frac{T^4 + T^3N - 3T^2N^2 + 4TN^3 - 2N^4}{2T^4 + NT^3 - T^2N^2} \tag{2.7}$$

Note that Eq. (2.7) is symmetric at $N = \frac{T}{2}$. By differentiating $y_\infty(\frac{k}{2})$ with respect to N, we find that the N values at which the function obtains a maximum are $0.5T$, $2.5908T$, $-1.5908T$, $0.8587T$ and $0.1413T$. Note that we only consider solutions that satisfy the condition $0 \leq N \leq T$. Due to the symmetry of Eq. 2.7, the maximum value of $y_\infty(\frac{k}{2})$ is 0.5147 for $N = 0.8587T$ and $N = 0.1413T$. □

To address the effect of the network structure on the steady state probability of infection when $s = \frac{k}{2}$, we show a simple example on a ring network structure ($\lambda_{max} = k = 2$), where we rewired every link towards a common node. Figure 2.5 represents the $y_\infty(s)$ curve of every rewiring step. The figure shows that all the steady state infected population curves pass close to the point $(\frac{k}{2}, 0.5)$. In addition, we performed extensive simulations to validate our assumption on different network classes with network sizes ranging from 100 nodes up to 3×10^5 nodes with different connectivities. All the simulations are averaged over 10^3 runs. We randomly rewired a 100-node regular network with $k = 10$ and a 100-node Barabási-Albert (BA) preferential attachment (PA) network with $k = 7.2$ each with 10^3 rewiring steps, and

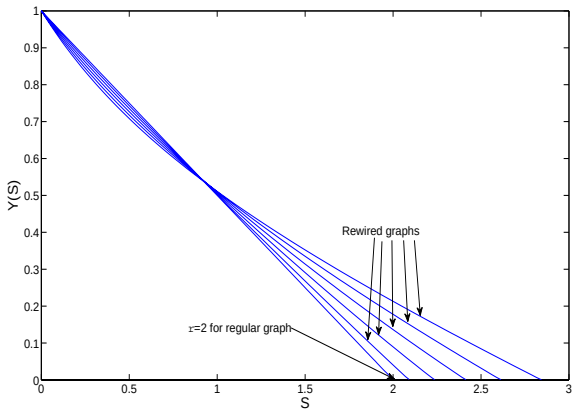
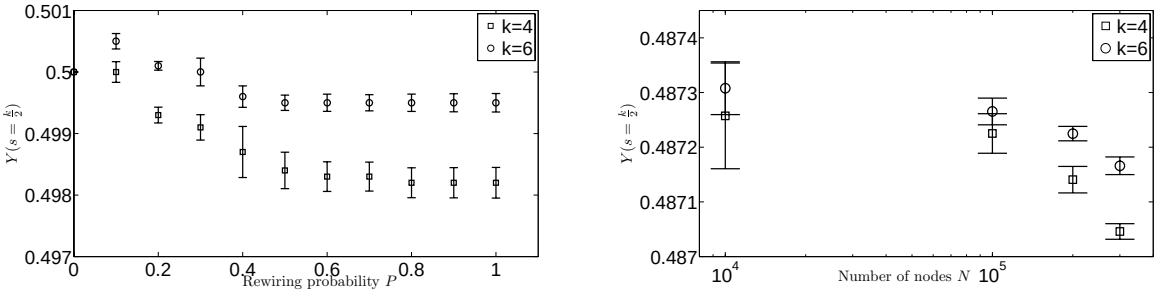


Figure 2.5: *The steady infected populations of a ring network and its rewired versions.*

we show that $y_\infty(\frac{k}{2})$ is very close to 0.5 as Table 2.2 shows. Moreover, we created 10^4-node Barabási-Albert (BA) preferential attachment (PA) networks with $k = 8,\ 12,\ 16$ and 20. For every type of network, we computed the average and the variance values of $y_\infty(\frac{k}{2})$ as in Table 2.2. The results show that the average of samples of $y_\infty(\frac{k}{2})$ is very close to 0.5. Furthermore, we evaluated $y_\infty(s = \frac{k}{2})$ for Watts-Strogatz (WS) networks with network size $N = 10^4$, average node degree $k = 4$, and 6, and rewiring probability $0 \leq p \leq 1$ as shown in Figure 2.6(a). We noticed that Watts-Strogatz (WS) networks with $k = 4$ have a larger deviation from the value 0.5 than the networks with $k = 6$. Additionally, the deviation converges with the rewiring probability p reaches 1. Thus, the results show that $y_\infty(\frac{k}{2})$ can be approximated to 0.5 with small deviation. To study the effect of the finite network size on the value of $y_\infty(\frac{k}{2})$, we created Barabási-Albert (BA) preferential attachment (PA) networks with different sizes and different average node degrees, and we evaluated $y_\infty(s = \frac{k}{2})$ as shown in Figure 2.6(b). Again, the results validate our assumption. We believe that for any connected network $G(N, L)$ with average node degree k, for $s = \frac{k}{2}$, the fraction of infected nodes at steady state is $0.5 + O(\epsilon)$ where $|\epsilon| \ll 1$. Therefore, in Section 2.5, where we propose a heuristic for the new robustness metric with respect to spread of epidemics, we neglect the parameter ϵ and assume that $y_\infty(\frac{k}{2}) = 0.5$.

Network	N	k	Average	Variance
Randomly rewired regular	10^2	10	0.499	2.7843×10^{-7}
preferential attachment	10^2	7.2	0.4849	2.59×10^{-5}
preferential attachment	10^4	8	0.4885	2.34×10^{-5}
preferential attachment	10^4	12	0.4883	2.01×10^{-5}
preferential attachment	10^4	16	0.4881	1.806×10^{-5}
preferential attachment	10^4	20	0.4881	1.8013×10^{-5}

Table 2.2: *The average and variance values of $y_\infty(s)$ are evaluated at $s = \frac{k}{2}$ for a randomly rewired regular network and different Barabási-Albert (BA) preferential attachment (PA) networks.*



(a) The value of $y_\infty(s = \frac{k}{2})$ for Watts-Strogatz (WS) networks given $N = 10^4$, $k = 4$ and 6, and rewiring probability $0 \leq p \leq 1$.

(b) The value of $y_\infty(s = \frac{k}{2})$ for Barabási-Albert (BA) preferential attachment (PA) given $k = 4$ and 6, and different network size N.

Figure 2.6: *Evaluation of the steady state infected population $y_\infty(s)$ for $s = \frac{k}{2}$ on Watts-Strogatz (WS) and Barabási-Albert (BA) preferential attachment (PA) networks given different network sizes N and different average node degrees k.*

2.5 Computation of VC_{SIS} and bounds

For general networks with heterogeneous structure, we cannot analytically compute the fraction of infected nodes $y_\infty(s)$, and hence the viral conductance, explicitly. Moreover, computing the solution of $y_\infty(s)$ for $0 \leq s \leq \lambda_{max}$ numerically is not feasible for large scale networks. Therefore, in this section, we propose a heuristic for computing the viral conductance for general networks. We will use the lemmas and theorems of the previous section to construct a heuristic to compute the new robustness metric, VC_{SIS}. Our objective is to make the heuristic as simple as possible to avoid computation complexity.

2.5.1 A heuristic for VC_{SIS}

The heuristic mainly depends on fitting linear and non-linear functions passing through three main points on the steady state infected population curve $y_\infty(s)$. The three points (s,$y_\infty(s)$) are as follows:

- Point $(0, 1)$ where s equals 0 (i.e. $\delta = 0$) and hence the whole network is infected at steady state (i.e. $y_\infty(0) = 1$).
- Point $(\frac{k}{2}, 0.5)$ as discussed in Section 2.4.
- Point $(\lambda_{max}, 0)$, which is the reciprocal of the epidemic threshold where the network is cured.

The basic heuristic equation is as follows:

$$y_{heuristic\infty}(s) = \begin{cases} \frac{y_{1\infty}(s)+y_{2\infty}(s)}{2} & 0 \leq s \leq \frac{k}{2} \\ \frac{y_{1\infty}(s)+y_{3\infty}(s)}{2} & \frac{k}{2} \leq s \leq \lambda_{max} \end{cases} \tag{2.8}$$

The function $y_{heuristic\infty}(s)$ is defined over two intervals of the range of s. In each interval, $y_{heuristic\infty}(s)$ is the average of two fitting curves. For example, in the first interval where $0 \leq s \leq \frac{k}{2}$, $y_{1\infty}(s)$ is a decreasing exponential function, it continues to the second interval to the point $(\lambda_{max}, 0)$ and therefore is constrained to pass by the three main points. Figure 2.7

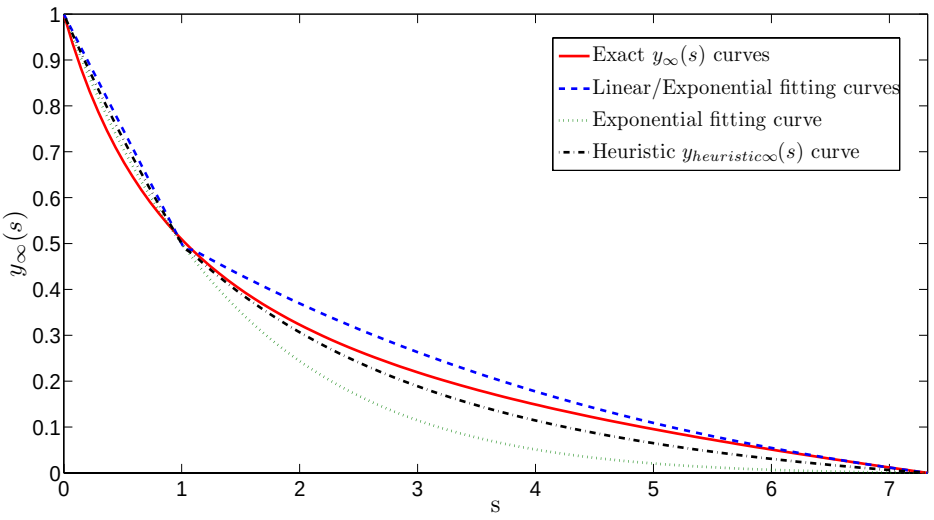


Figure 2.7: *Demonstration of the heuristic fitting curves and the three points on the $y_\infty(s)$ curve for a preferential attachment network with average node degree of 2.*

shows the heuristic fitting curves and the three points on the $y_\infty(s)$ curve. The function $y_{1\infty}(s)$ is defined as follows:

$$y_{1\infty}(s) = (\frac{\lambda_{max} - s}{\lambda_{max}})e^{-as} \tag{2.9}$$

To obtain the value of the exponent coefficient a, we solve $y_{1\infty}(s)$ at the point $(\frac{k}{2}, 0.5)$. Therefore, the value of a becomes as follows:

$$a = -\frac{2}{k} ln(\frac{\lambda_{max}}{2\lambda_{max} - k}) \tag{2.10}$$

Next, the second function in the first interval $y_{2\infty}(s)$ is a linear decreasing function, proposed to equalize any underestimation due to the exponential function $y_{1\infty}(s)$. It is defined as follows:

$$y_{2\infty}(s) = 1 - \frac{s}{k} \tag{2.11}$$

In the second interval, the function $y_{3\infty}(s)$ is a decreasing exponential function that passes through the points $(\frac{k}{2}, 0.5)$ and $(\lambda_{max}, 0)$, proposed to follow the tail of the $y_\infty(s)$

curve that depends on the irregularity of the network. Thus, the function $y_{3\infty}(s)$ is defined as follows:

$$y_{3\infty}(s) = b(\lambda_{max} - s)e^{-cs} \tag{2.12}$$

The exponent coefficient c depends on the irregularity of the network since $\lambda_{max} \geq k$ for all irregular networks [28]. Also, the total exponent of the exponential function of $y_{3\infty}(s)$ should be unitless; therefore, the exponent coefficient c has to hold a unit of (node degree)$^{-1}$. To compute the value of c, we propose the following equation:

$$c = \frac{1}{2}\sqrt{\frac{-1 + \sum_{i=1}^{N-1} 1_{i \in ND}}{\lambda_{max} k}} \tag{2.13}$$

$$1_{i \in ND} = \begin{cases} 1 & if\ i \in ND \\ 0 & Otherwise \end{cases} \tag{2.14}$$

where $1_{i \in ND}$ is a set function of the node degree, and ND is the set of node degrees that exists in the network, and therefore $\sum_{i=1}^{N-1} 1_{i \in ND}$ represents the irregularity of the network. Then, the constant b is computed as follows:

$$b = \frac{e^{c\frac{k}{2}}}{2\lambda_{max} - k} \tag{2.15}$$

Note that for regular networks with degree k ($k = \lambda_{max}$), the functions $y_{1\infty}(s)$ and $y_{3\infty}(s)$ become linear and they are similar to $y_{2\infty}(s)$, which is the exact $y_{\infty}(s)$ curve for regular networks.

By integrating $y_{heuristic\infty}(s)$ in Eq. 2.8, we obtain the final mathematical formula for the $VC_{SIS-heuristic}$ as follows:

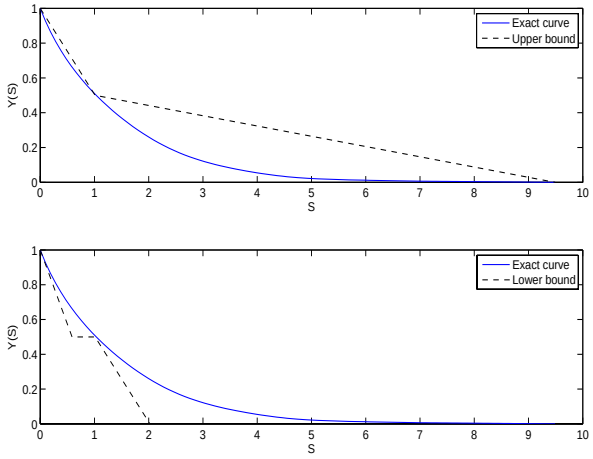


Figure 2.8: *An example of the upper and lower bounds of the viral conductance for the Abilene network with N=886, k=2.022 and λ_{max}=9.489. The upper figure represents the upper bound VC_{SIS-UB}, while the lower figure shows the lower bound VC_{SIS-LB}.*

$$VC_{SIS-heuristic} = \frac{1}{2}[\frac{3k}{8} + \frac{1}{a} - \frac{1}{\lambda_{max}a^2} - \frac{e^{-a\lambda_{max}}}{a} - \frac{b\lambda_{max}}{c}e^{-c\lambda_{max}} + \frac{b\lambda_{max}}{c}e^{-c\frac{k}{2}} + \frac{e^{-a\lambda_{max}}}{\lambda_{max}}(\frac{\lambda_{max}}{a} + \frac{1}{a^2}) + be^{-c\lambda_{max}}(\frac{\lambda_{max}}{c} + \frac{1}{c^2}) - be^{-c\frac{k}{2}}(\frac{k}{2c} + \frac{1}{c^2})] \quad (2.16)$$

2.5.2 Upper and lower bounds for VC_{SIS}

We formulated upper and lower bounds for VC_{SIS}, depending on the topological characteristics of the network, to avoid underestimating and overestimating the actual value of VC_{SIS}.

Upper bound

We know that the steady state infection population $y_\infty(s)$ is always a convex function [76] since connecting two points on the curve with a linear decreasing function renders the area under the linear function greater than the area under the actual curve (see Figure 2.8). Therefore, we computed that area under the following linear functions: 1) a linear function

that connects the points $(0, 1)$ and $(\frac{k}{2}, 0.5)$, and 2) a linear function that relates the points $(\frac{k}{2}, 0.5)$ and $(\lambda_{max}, 0)$. Thus, we formulated the upper bound as follows:

$$VC_{SIS-UB} = \int_0^{\frac{k}{2}} (1 - \frac{s}{k}) + ds + \int_{\frac{k}{2}}^{\lambda_{max}} \frac{1}{k - 2\lambda_{max}} (s - \lambda_{max}) ds \tag{2.17}$$

The upper bound VC_{SIS-UB} is as follows:

$$VC_{SIS-UB} = \frac{1}{4}(k + \lambda_{max}). \tag{2.18}$$

Lower bound

Section 2.4 shows that the slope of the steady state $y_\infty(s)$ at $s = 0$ satisfies $-\sigma$, where $\sigma = \frac{1}{N}\sum_{i=1}^{N} \frac{1}{d_i}$. Then using the line connecting the points $(\frac{k}{2}, 0.5)$ and $(k, 0)$ and the intersection point between the tangent line $y_\infty(s) = 1 - \sigma s$ and the line $y_\infty(s) = 0.5$, we constructed a lower bound VC_{SIS-LB} as shown in Figure 2.8. By applying Eq. 2.5, we found the following value for the lower bound:

$$VC_{SIS-LB} = \frac{1}{8}(\frac{1}{\sigma} + 3k) \tag{2.19}$$

For any regular network with node degree k, both the upper bound VC_{SIS-UB} and lower bound VC_{SIS-LB} lead to the same value of $VC_{SIS} = \frac{k}{2}$.

2.6 Numerical evaluations

In this section, we numerically evaluate the robustness metric VC_{SIS}, the accuracy of the heuristic, and the bounds on different types of networks like synthetic networks with 10^4 nodes, real-world networks, and survey-based networks. The results are averaged over 10^3 runs.

2.6.1 Assortative and disassortative preferential attachment networks

In this subsection, we show how the new robustness measure VC_{SIS} can differentiate between assortative and disassortive networks in which node degree correlation is observed. Such correlation is important since correlated networks exist in the real world. For example, social networks are classified as assortative networks, while technological and biological networks are classified as disassortative networks [84]. In assortative networks, nodes with similar node degrees are connected together, while in disassortative networks, nodes with different nodes degrees are connected together. Accordingly, we generated assortative and disassortative preferential attachment (PA) networks using the algorithm in [47]. The algorithm starts with a connected network having $m_0 \ll N$ nodes. Every new node is connected to already existing nodes through two stages: In the first stage, a new node is connected to an existing node u with probability $\pi_u = \frac{d_u}{\sum_j d_j}$ where d_u is the degree of node u; in the second stage, a new link between the new node and one of the neighbors s of the chosen node u in the first stage is added with probability $p_s = \frac{d_s^\alpha}{\sum_{v \in \Gamma_u} d_v^\alpha}$, where α is an assortative tuning coefficient, and Γ_u is the set of neighbors of individual u chosen in the first stage. The generated assortative and disassortative preferential attachment (PA) networks have different node degree distributions. We will address the analytical and numerical studies of the robustness of correlated PA networks having the same node degree distribution in our future work.

To show how VC_{SIS} can distinguish among correlated networks and also that the heuristic is accurate and close to the exact VC_{SIS}, we evaluated VC_{SIS} and $VC_{SIS-heuristic}$ on correlated PA networks with 10^4 nodes given different average node degrees k as in Figure 2.9. We noticed that $VC_{SIS-heuristic}$ increases with k, and it is close to the exact VC_{SIS} for both types of networks. In addition, both exact VC_{SIS} and $VC_{SIS-heuristic}$ of assortative networks are lower than their corresponding values for disassortative networks, showing that the new heuristic is capable of evaluating the robustness of networks. Moreover, we com-

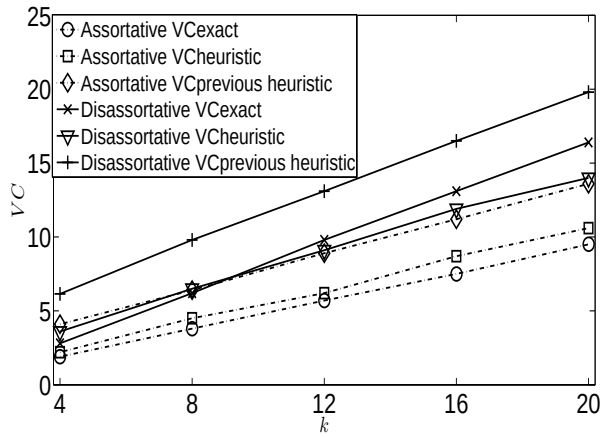


Figure 2.9: *The exact value of VC_{SIS} is compared with the heuristic approach $VC_{SIS-heuristic}$ in Eq. 2.16, and with the previous heuristic approach $VC_{SIS-previous\ heuristic}$ presented in [60] on assortative and disassortative preferential attachment (PA) networks with 10^4 nodes given different average node degrees $k = 4,\ 8,\ 12,\ 16$ and 20.*

pared the new heuristic presented in this chapter with a previous heuristic presented in [60]. In Figure 2.9, we observed that the new heuristic $VC_{SIS-heuristic}$ is closer to $VC_{SIS-exact}$ than the previous heuristic $VC_{SIS-previous\ heuristic}$. Additionally, the values of $VC_{SIS-heuristic}$ of disassortative networks do not overlap with $VC_{SIS-exact}$ and $VC_{SIS-heuristic}$ for assortative networks. Consequently, the new heuristic outperforms the previous heuristic, and it is capable of evaluating the robustness of large networks for which the computation of $VC_{SIS-exact}$ may not be feasible.

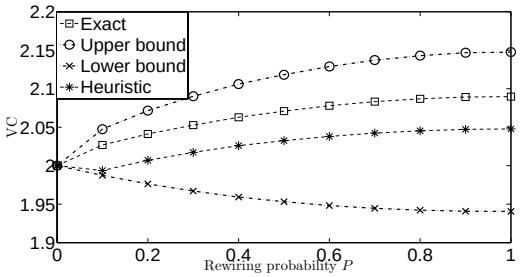
The robustness of assortative networks with respect to the spread of epidemics is also discussed in [84] showing that the giant component in assortative networks is smaller than in dissasortative networks. On one hand, nodes with high degrees are connected together causing any epidemic to persist in the network. On the other hand, the epidemic survives in only a small portion of the network. Here, our results show that disassortative preferential attachment (PA) networks are more vulnerable than assortative preferential attachment (PA) networks. Notably, the algorithm used to create the correlated PA networks does not constrain the degree of the nodes in the networks. Therefore, few nodes with very large node

degrees appear in the disassortative networks, and they are connected with low node degree nodes. Consequently they increase the network heterogeneity properties; once an epidemic reaches any node with large node degree, a major outbreak takes place in the network.

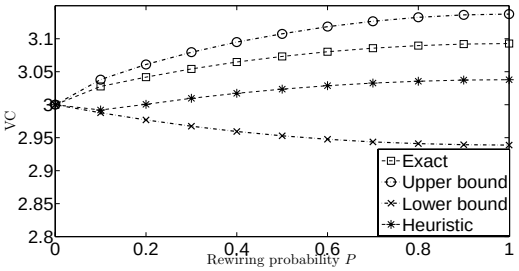
2.6.2 Watts-Strogatz small world model

We generated Watts-Strogatz (WS) small world networks [107] with a given number of nodes N and average node degree k. To create a WS small world network, we started with a k-regular network; each node has $\frac{k}{2}$ links that connect it to its nearest counterclockwise neighbors, while the other $\frac{k}{2}$ links connect the same node to its nearest clockwise neighbors. Given a rewiring probability p, we started rewiring the clockwise links for every node. We created WS networks given $N = 10^4$ nodes with average node degrees $k = 4, 6, 8$, and 10 and different probability of rewiring p ranging from 0 to 1. Then, we used those networks to evaluate the heuristic value $VC_{SIS-heuristic}$ compared with the exact VC_{SIS} values, and to study the effect of rewiring the network links on the values of the VC_{SIS}.

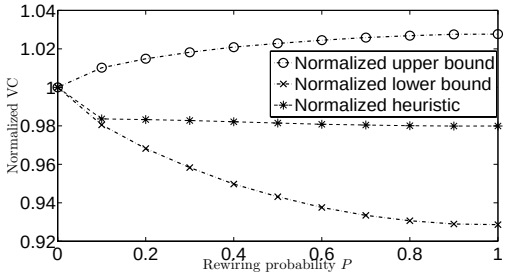
Figures 2.10(a) and 2.10(b) show how the exact value of VC_{SIS} changes with a probability of rewiring p given $k = 4$ and 6, respectively. For networks with given regular node degrees ($p = 0$), VC_{SIS} equals $\frac{k}{2}$. In addition, the exact VC_{SIS} value and $VC_{SIS-heuristic}$ non-linearly increase with the probability of rewiring p because the irregularity of the network increases with p. We also verified the validity of the upper and lower bounds VC_{SIS-UB} and VC_{SIS-LB}, noticing that for all networks with different average node degrees and different probability of rewiring, exact values of VC_{SIS} as well as heuristic values $VC_{SIS-heuristic}$ are bounded. To evaluate the deviation of $VC_{SIS-heuristic}$, the upper bound and the lower bound, we normalized the results with respect to the exact value of VC_{SIS} as shown in Figures 2.10(c) and 2.10(d) for $k = 4$ and 6, respectively. We observed that the deviation of the heuristic value of VC_{SIS} is bounded with the increase of rewiring probability p. All the above analysis of VC_{SIS} was also applied to the WS small world networks with $k = 8$ and 10 with the same observations as discussed in this subsection.



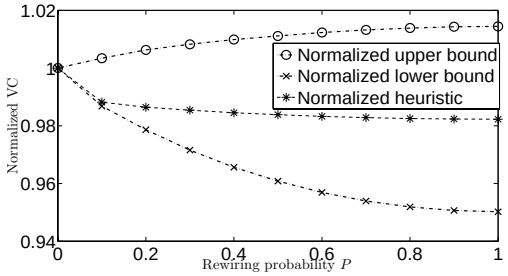
(a) Absolute values of VC_{SIS} given $k = 4$



(b) Absolute values of VC_{SIS} given $k = 6$

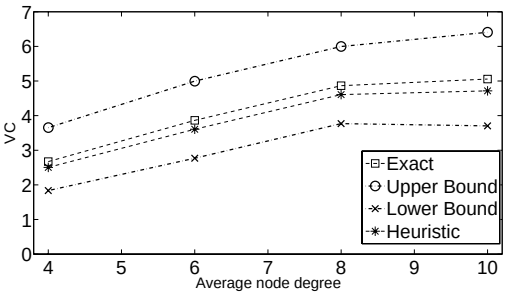


(c) Normalized values of VC_{SIS} given $k = 4$

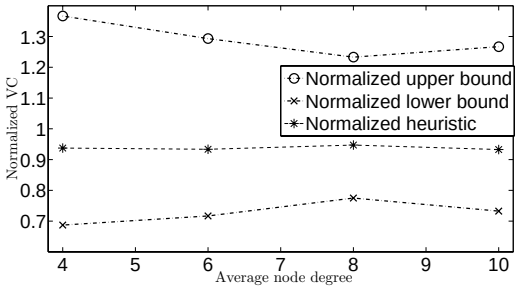


(d) Normalized values of VC_{SIS} given $k = 6$

Figure 2.10: *Absolute values of exact VC_{SIS}, heuristic VC_{SIS}, and upper and lower bounds VC_{SIS} in Figures 2.10(a) and 2.10(b); normalized values of heuristic VC_{SIS}, and upper and lower bounds VC_{SIS} with respect to exact value of VC_{SIS} in Figures 2.10(c) and 2.10(d) for Watts-Strogatz (WS) networks given $N = 10^4$ and probability of rewiring $0 \leq p \leq 1$.*



(a) Absolute values of VC_{SIS} given $k = 4, 6, 8$ and 10



(b) Normalized values of VC_{SIS} given $k = 4$ $k = 4, 6, 8$ and 10

Figure 2.11: *Absolute values of exact VC_{SIS}, heuristic VC_{SIS}, and upper and lower bounds VC_{SIS} in Figure 2.11(a); normalized values of heuristic VC_{SIS}, and upper and lower bounds VC_{SIS} with respect to exact value of VC_{SIS} in Figure 2.11(b) for Barabási-Albert BA networks given $N = 10^4$ and different average node degrees.*

2.6.3 Barabási-Albert Preferential attachment network model

We generated Barabási-Albert (BA) preferential attachment (PA) networks as follows: we started with a small number m_0 of disconnected nodes; next we connected a new node to an existing node u with probability $\pi_u = \frac{d_u}{\sum_j d_j}$, where d_u is the node degree of the existing node u. The generated BA network is characterized as an uncorrelated Barabási-Albert (BA) preferential attachment (PA) network. We evaluated VC_{SIS} on the generated BA networks with $N = 10^4$ and $k = 4, 6, 8$ and 10 as shown in Figure 2.11(a). The value of VC_{SIS} increases as k increases, and also the heuristic value of VC_{SIS} is close to its corresponding VC_{SIS} value. Moreover, both exact and heuristic VC_{SIS} values are bounded. Also, we evaluated the deviation of the heuristic VC_{SIS} values, and the upper and lower bounds by computing their normalized values with respect to the exact values of VC_{SIS} as shown in Figure 2.11(b). We found that the heuristic slightly deviates from its exact value, and therefore it perfectly estimates the exact value of VC_{SIS}.

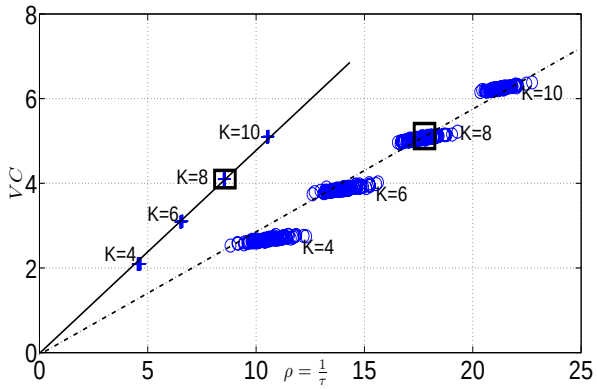


Figure 2.12: *Numerical relationship between VC_{SIS} and $\lambda_{max} = \rho = \frac{1}{\tau}$ (the reciprocal of the epidemic threshold) collected from Watts-Strogatz (WS) networks '+' with rewiring probability $p = 1$ and Barabási-Albert (BA) preferential attachment (PA) networks 'o' with $N = 10^4$ given different average node degrees.*

2.6.4 VC_{SIS} versus λ_{max}

Here, we numerically present the differences between VC_{SIS} and λ_{max} as shown in Figure 2.12, which comprises models with 100 samples of results for a given network class with a certain average node degree k. In general, the values of VC_{SIS} tends to increase with λ_{max}, and simultaneously both measures increase with the average node degrees. Note that within a module of samples, VC_{SIS} does not always need to increase with λ_{max}. We also observed that the slope for Watts-Strogatz (WS) networks results is higher than the slope in Barabási-Albert (BA) preferential attachment (PA) networks results. To relate the slope with VC_{SIS} and λ_{max}, recall that $VC_{SIS} = \overline{y_\infty}\lambda_{max}$, and therefore, the slope of the trend line is $\overline{y_\infty}$. To address the differences between VC_{SIS} and λ_{max}, let us compare the robustness of a WS network with $k = 8$ and a PA network with the same value of k as shown in the black box in Figure 2.12. The value of λ_{max} of the PA network is at least twice the value of λ_{max} of WS network, and therefore the value of $\overline{y_\infty}$ of PA network is at most half its corresponding value of WS network. Such a trade-off was discussed earlier in this dissertation, so on one hand, we can not measure the robustness of networks by considering only λ_{max} or the average infection $\overline{y_\infty}$, yet on the other hand, VC_{SIS} combines both λ_{max} and

$\overline{y_\infty}$ to account for them to measure the robustness of networks. Moreover, for any average node degree in PA networks, the samples widely scatter in the direction of λ_{max} than VC_{SIS} in a given module, while samples from WS networks do not scatter largely. As a result, for a given value of λ_{max}, it is not difficult to obtain distinct values of VC_{SIS} for different networks. For example, for a given value of λ_{max}, WS networks are less robust than PA networks. Therefore, it is not necessary that PA networks are always more vulnerable than WS networks, and also λ_{max} is not a unique robustness measure with respect to the spread of epidemics.

2.6.5 Internet AS-level networks

Next, we apply VC_{SIS} to measure networks like the Internet AS-level networks. Specifically, we employed three different networks, SKITTER [22] with 9204 nodes and $k = 6.29$, BGP [113] with 17446 nodes and $k = 4.68$, and WHOIS [114] with 7485 and $k = 15.22$, to measure their robustness and to study the effect of rewiring the networks with the value of VC_{SIS}. The rewiring process aims to decrease the variance of node degree by regularizing the network using the algorithm in [87]. Next, we compared the robustness of each network in two cases: the original network and its regularized network. Table 2.3 shows that regularized networks have a lower $VC_{SIS-heuristic}$ than original networks. In addition, we used the metric κ to assess the heterogeneity of the irregular and regularized networks. The metric κ was introduced in the literature [24] as the ration between the second moment and the first moment of the node degree distribution. After regularizing a network, the heterogeneity of the network decreases leading to decrease the degrees of nodes with large degrees and to increase the degrees of nodes with low degrees. Therefore, all the node degrees are centered around the average degree value k with low variance in node degrees, while the maximum degree and the minimum degree are $\lceil k \rceil$ and $\lfloor k \rfloor$, respectively. Consequently, $y_\infty(s)$ curve of the regularized networks shrinks dramatically and approaches a linear decreasing function. Assessing the $VC_{SIS-heuristic}$ and κ, both metrics confirm that regular networks are more

	VC_{SIS-LB}	$VC_{SIS-heuristic}$	VC_{SIS-UB}	λ_{max}	Heterogeneity $\frac{\kappa}{k}$
SKITTER	2.6041	4.3661	21.4559	79.53	6.683
Regularized SKITTER	3.1415	3.1428	3.1559	6.33	1.006
WHOIS	6.0708	9.7964	41.5197	150.86	8.9104
Regularized WHOIS	7.6082	7.61	7.6130	15.23	1.0002
BGP	1.9750	3.2877	19.4348	73.06	12.2295
Regularized BGP	2.3334	2.3385	2.3549	4.73	1.0089

Table 2.3: *$VC_{SIS-heuristic}$, VC_{SIS-UB}, VC_{SIS-LB}, λ_{max} and κ of Internet AS-level networks and their regularized networks.*

robust than irregular networks against the spread of epidemics. Moreover, the values of λ_{max} for irregular and regularized networks in Table 2.3 show that both values of VC_{SIS} and λ_{max} decrease when the networks are regularized. For every network, this behavior occurs due to the decrease in the maximum node degree, which upper-bounds the spectral radius of the network, resulting in a largely decreased area under the curve, and therefore VC_{SIS} decreases too.

2.6.6 Survey-based network

In this subsection, we applied VC_{SIS} to a social network created through a survey to study the spread of epidemics in rural regions. The survey was conducted in Clay Center, the county seat of Clay County in the State of Kansas, and the network was created based on the responses of Clay Center residents. The survey included questions about three main places, {R,W,G}, that the residents visit, and questions about three levels of contact each respondent makes with other individuals. The three levels of contact were defined as follows: 1) *Proximity contact*, which happens when another person is passing within five feet, 2) *Direct-Low contact*, which happens when a person is directly touching another person for short time period, and 3) *Direct-High contact*, which happens when a person is directly touching another person for a long time period. Every respondent i provided the number of contacts $n_{x,i}$ for every contact level x. We used the survey responses to create a weighted contact network with 138 nodes (respondents) and 9222 links (contacts) as shown

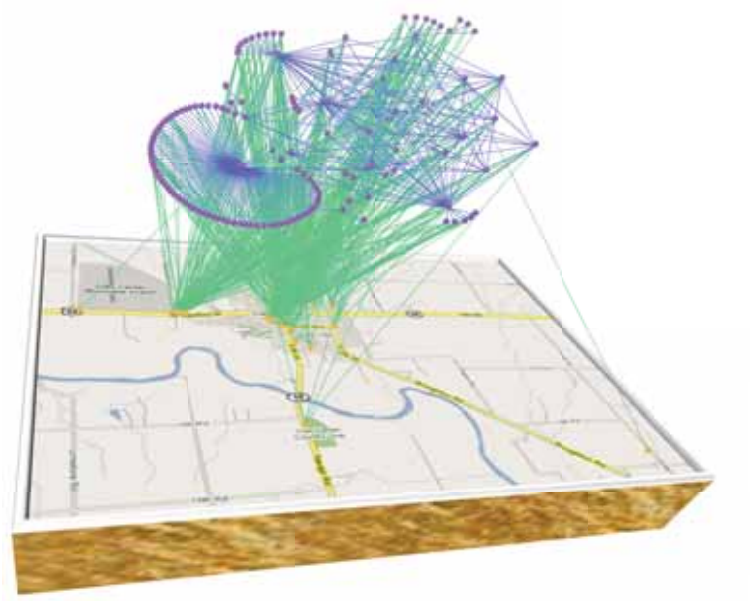


Figure 2.13: *The survey-based network of individuals and popular locations in Clay Center, Kansas, where the nodes (survey respondents) in the cloud network are connected via green edges to the locations in Clay Center according to the survey responses. The map is courtesy of Google.*

in Figure 2.13. The weight between respondent i and respondent j is the average of the three contact levels. For contact level x, we proposed the following equation:

$$w_{x,i,j} = (1 - (1 - \mu_{i,j}\pi_x)^{n_{x,i}})(1 - (1 - \mu_{i,j}\pi_x)^{n_{x,j}}) \tag{2.20}$$

where π_x is a constant that depends on the level of contact x and $\mu_{i,j}$ quantifies the location responses for both respondents i and j as follows:

$$\mu_{i,j} = \frac{1 + l_{i,j}}{1 + d} \tag{2.21}$$

where d is the total number of locations, and $l_{i,j}$ represents number of common locations that respondents i and j used to visit. For more details about the survey questions and the link weights, we refer the reader to [99].

We applied the measure VC_{SIS} to quantify the robustness of the social network with respect to the spread of SIS epidemics. We also applied some mitigation strategies to the

Mitigation strategy	VC_{SIS}	λ_{max}	Immunized nodes
None (weighted network)	71.4985	31.91	0
High strength node immunization	49.1791	23.89	14
High strength node immunization in location R	49.4439	24.08	14

Table 2.4: *The value of VC_{SIS} for the survey-based social network in case of the absence of mitigation strategies and the most two effective mitigation strategies.*

network and used VC_{SIS} to rank them. Also, we studied mitigation strategies where 10% of the nodes were immunized. From a network point of view, the immunized nodes were removed from the network. Immunized nodes were selected based on 1) node strength (the sum of link weights of a node), 2) random selection, 3) random selection from a specific location, and 4) highest strength nodes from a specific location. Table 2.4 shows the values of VC_{SIS} and λ_{max} in case of the absence of mitigation strategies and the most two effective mitigation strategies. Clearly, we notice that VC_{SIS} values, when mitigation strategies are applied, are lower than the VC_{SIS} value in absence of mitigation. Also, the highest strength mitigation strategy outperforms other mitigation strategies since it has the lowest VC_{SIS} value because highest strength nodes play a major role in spreading any epidemic. Also, the highest strength mitigation strategy outperforms the highest strength mitigation strategies that are applied at different locations because the former considers all nodes with the highest strength in the network regardless of their locations. This result agrees with the effect of mitigation strategies presented in the literature (for example see [92]). Therefore, the highest strength mitigation strategy has the best effectiveness for reducing the spread of epidemics. Observing the values of robustness metrics in our example, we see that both VC_{SIS} and λ_{max} rank the mitigation strategies similarly.

2.7 Summary

We summarize the above results and analysis of robustness with respect to spread of epidemics in the following conclusions.

- *Viral Conductance is a better measure than the epidemic threshold for robustness of networks*

  VC_{SIS} incorporates the fraction of infected nodes at steady state for all possible infection strengths.

- *Increasing the probability of rewiring decreases the robustness of Watts-Strogatz (WS) networks*

  The initial regular network in the Watts-Strogatz (WS) model has the lowest value of VC_{SIS}, and therefore it is the most robust of any other obtained network given the probability of rewiring $0 < p \leq 1$.

- *$VC_{SIS-heuristic}$ is close to the exact value of VC_{SIS}*

  The proposed heuristic satisfies the basic requirements of simplicity and high accuracy in addressing solutions for any expensive computation quantity.

- *VC_{SIS-UB} and VC_{SIS-LB} effectively bound VC_{SIS} and $VC_{SIS-heuristic}$ from above and from below, respectively*

  Bounds give the feasible region in which the value of VC_{SIS} is predictable.

- *Our numerical results show that the regular structure of a network has a minimum VC_{SIS} value compared to any other structure*

  Given N nodes and L links, we can obtain $\binom{\binom{N}{2}}{L}$ different network structures. We believe that the regular structure of a network is the most robust to any spread of epidemic.

Chapter 3

Robustness of networks with respect to the spread of SIR epidemics

This chapter addresses the robustness of networks with respect to the spread of susceptible/infected/recovered (SIR) epidemics. Different SIR models are applied to some classes of contact networks [14–16, 37, 54, 70, 74, 79, 83, 99, 103, 105, 109] depending on the network characteristics. For example, an early SIR approach uses the homogeneous mean field approximation in which all individuals have the same probability of being infected and infectious. On the other hand, the SIR heterogeneous mean field approach has also been applied to structured networks considering the local connectivity of the network's individuals. For example, scale-free networks, which are networks owning power-law node degree distribution $P(k) \sim k^{-\nu}$, show a high level of vulnerability to the spreading of epidemics due to the highly heterogeneous node degree distribution property when the minimum node degree is greater than two [16]. In addition, the spread of epidemics has been studied on correlated networks and uncorrelated networks separately. However, the SIR epidemic approaches studied in the literature did not consider the whole structure of the network but only the local connectivity. For instance, the heterogeneous mean field approach only considers the node degree distribution, which is an aggregate representation of the network. However, it has been shown that networks with distinct topological properties can be characterized by the same node degree distribution [65]. Consequently, these approaches cannot distinguish among

individuals having the same node degrees because they neglect the centrality properties of the individuals. Therefore, the SIR epidemic approaches presented in the literature are not accurate for studying the effect of the network in the spreading process. Consequently these approaches cannot used to study the robustness of networks, and hence there is a need for a generalized SIR epidemic approach that considers not only the network properties (i.e. average node degree and node degree distribution), but also the whole network structure, and that represents every individual independently.

In this chapter, before we address the robustness metric VC_{SIR}, firstly, we propose an individual-based SIR approach [112], which is inspired by the continuous-time Markov chain model and which represents the network in the most accurate way. We separately study the state of each individual during the infection process, revealing the role of the individual's centrality properties in spreading the infection across the network. Although the continuous time SIR Markov model, based on the Markov chain stochastic process, describes the global change in the state probabilities of the network, it is limited to small networks due to the exponential divergence in the number of possible network states 3^N with the growth of network size N. Instead, our approach aims to reduce the complexity of the problem to $O(N)$ and to offer insights into the epidemic spreading mechanism. Through the new SIR approach, we study the spread of epidemics on any type of network regardless of its topological structure. We analytically derive the epidemic threshold for the new approach, which is inversely proportional to the spectral radius λ_{max} (the supremum eigenvalue within the eigenvalue spectrum) of the network.

Secondly, we propose a new robustness measure VC_{SIR}, which incorporates the fraction of infection sizes for different effective infection rates. The new measure is used to quantify the robustness of correlated preferential attachment networks.

In the last part of the chapter, a mitigation strategy is proposed for the spread of SIR epidemics in social contact networks. We consider a time variable weighted contact network

where the weight variation represents the variation of the level of contact; the initial value of the weight is the normal contact level, which can be reduced to represent a lower level of contact. We utilize the individual-based SIR model to determine the optimal adaptive contact network, in which both the total infection size and the variation in the contact weights of individuals are minimized. To this end, we formulated a continuous time optimal control problem, where the objective function is given by a weighted sum of the total size of the epidemic, and the network weight variation. The SIR system of differential equations, describing the evolution of the infection probability for each individual in the network, become the constraints of the optimal control problem. The optimal control formulation and two heuristics were developed and tested extensively on a large set of diverse contact networks, showing the effectiveness and implementability of the proposed methods.

This chapter is organized as follows: In Section 3.1, we review basic modeling approaches that are applied to SIR model and the mitigation strategies in the literature. In Section 3.2 we review the continuous time Markov chain model, and in Section 3.3, we present the individual-based approach in details and we show simulative and numerical results. The properties of the individual-based approach are discussed in Section 3.4. Moreover, the theoretical deviation between the individual-based approach and the continuous time Markov chain model is analyzed in Section 3.5. In Section 3.6 we introduce a new metric VC_{SIR} to quantify the robustness of networks with respect to the spread of SIR epidemics. Moreover, we present an optimal mitigation strategy approach for social contact networks in Section 3.7. Finally, the theoretical and numerical findings are summarized in Section 3.9.

3.1 Related work

The science of the spread of epidemics is based on compartmental models that assume individuals are classified into non-intersecting sets [8, 81]. Thus, the classical susceptible/infected/removed SIR model characterizes diseases that lead to either immunization or

death of individuals. The infected individuals are in the infected set, the healthy ones are in the susceptible set, and the cured or removed ones are in the removed set. Initially, a small number of infected individuals exist that try to infect their susceptible (healthy) neighbors. After receiving the infection, susceptible individuals become infected, and later they try to infect their susceptible neighbors. In such cases, infected individuals are infectious. Subsequently, every infected individual either is cured due to immunization or removed due to death.

3.1.1 SIR homogeneous mean field approach

In the SIR homogeneous mean field approach [8, 57], $s(t)$, $i(t)$, and $r(t)$ represent respectively the fraction of susceptible, infected, and recovered populations. As mentioned in the Introduction, we assume that the population is fixed i.e. $s(t) + i(t) + r(t) = 1$. In addition, the approach approximates the representation of the network and assumes that, on the average, every individual is connected with $< k >$ neighbors neglecting the heterogeneity of the node degrees. Depending on the fixed population assumption and the network representation, the homogeneous mean field approach describes the change in the susceptible, infected and recovered population fractions over time. The infected fraction $i(t)$ infects the fraction of susceptible neighbors $< k > s(t)$ with infection rate β, and simultaneously, a fraction of the infected population recovers $\delta i(t)$.

The rates of changes in s, i and r fractions are governed by the following continuous time differential equations:

$$\frac{ds(t)}{dt} = - < k > \beta i(t)s(t), \tag{3.1}$$

$$\frac{di(t)}{dt} = -\delta i(t) + < k > \beta i(t)s(t), \tag{3.2}$$

$$\frac{dr(t)}{dt} = \delta i(t). \tag{3.3}$$

These differential equations interpret the infection and cure processes. Initially, the

spreading process starts with a small infected fraction $i(0) \simeq 0$, a susceptible fraction of almost one $s(0) \simeq 1$, and the removed fraction of zero $r(0) = 0$. Every infected individual infects on average $< k >$ susceptible neighbors, each with an infection rate β, where $< k >= \sum_{k=1}^{N-1} kp(k)$ is the average node degree (average number of contacts), and $p(k)$ is the probability of having an individual with degree k. Following differential Eq. 3.3, an infected individual is removed at a rate δ. The removed fraction increases with time until it reaches a certain fraction level depending on the strength of the epidemic. A non-zero epidemic threshold exists and it is equal to $\delta(R_0 - 1)$ where R_0 is the reproductive number and equals $\frac{<k>\beta}{\delta}$. If R_0 is greater than 1, the epidemic prevails in the network. On the other hand, if R_0 is less than 1, the initially infected individuals totally recover without infecting other susceptible individuals [11]. Since on average, every infected individual infects a constant number of neighbors, the homogeneous approach does not account for heterogeneity in the node degrees of individuals in the network.

3.1.2 SIR heterogeneous mean field approach

Another approach in the literature is the heterogeneous mean field (also called heterogeneous mixing) SIR approach [16, 79, 109], which was proposed to overcome the shortcomings of the homogeneous approach. In this approach, individuals are classified according to their node degrees. Thus, for a given node degree k, the states' fractions $s_k(t)$, $i_k(t)$and $r_k(t)$ evolve with time t, and their sum is constant, such that $s_k(t) + i_k(t) + r_k(t) = 1$. The rates of changes in the three states for a given node degree k are governed by the following set of differential equations:

$$\frac{ds_k(t)}{dt} = -k\beta s_k(t)\theta_k(t), \tag{3.4}$$

$$\frac{di_k(t)}{dt} = -\delta i_k(t) + k\beta s_k(t)\theta_k(t), \tag{3.5}$$

$$\frac{dr_k(t)}{dt} = \delta i_k(t). \tag{3.6}$$

This approach was applied to both uncorrelated and correlated networks, leading to further analysis of the epidemic threshold. For uncorrelated networks, the epidemic threshold is $\tau^{ucr} = \frac{<k>}{<k^2>-<k>}$, where $< k^2 >$ is the second moment of the node degree distribution, and $\theta(t)$, representing the probability that a link is pointing to an infected individual, is found to be $\frac{\sum_k (k-1)p(k)i_k(t)}{<k>}$. On the other hand, the epidemic threshold for correlated networks is $\tau^{cr} = \frac{1}{\overline{\Lambda}_m}$, where $\overline{\Lambda}_m$ is the maximum eigenvalue of the connectivity matrix $\overline{C}_{kk'} = \beta \frac{k(k'-1)}{k'} p(k' \mid k)$, and $\theta_k^{cr}(t)$ equals $\sum_{k'} i_{k'}(t)(\frac{k'-1}{k'})p(k' \mid k)$.

Although this approach considers the heterogeneous connectivity in the networks, it does not reveal the state of each individual in the network. It only reflects the evolution of the fractions over time for a given node degree, while neglecting the states of individuals within that node degree.

3.1.3 Mitigation strategies for SIR epidemics

The work in [61] addressed the effect of quarantine strategy on the spread of SIR epidemics. The quarantine strategy assumed that the susceptible individuals disconnect their contacts with the infected neighbors, and they reconnect with other susceptible neighbors with a given probability. Using the rewiring approach, the authors found that there is a phase transition at a critical rewiring probability below which large number of individuals are infected. Our approach is different from [61] since individuals do not terminate their daily contacts with the infected neighbors and create new contacts with new individuals; However, susceptible individuals reduce their contact frequency with the infected neighbors.

Gao and Ruan in [39] studied the effect of human movement on the spread of infectious SIS epidemics. They confirmed the existence of reproductive number below which a disease does not spread out.

Gross and Blasius highlighted the research in adaptive networks as shown in [44] with two major lines. The first line is related to dynamics of networks that the topology evolves over time revealing many special characteristics. The second line of research is the dynamics on

networks in which the states of the individuals in the network change with time. Additionally, the relationship between the dynamics of networks and the dynamics on networks was studied showing that there is a feedback loop between the state of the individuals and the topology. Furthermore, surveys on ubiquity of adaptive networks across disciplines, robust self-organization in Boolean networks, leadership in coupled oscillator networks, cooperation in games on adaptive networks, and dynamics and phase transitions in opinion formation and epidemics were briefly discussed.

In addition, Gross et al.in [45] discussed the spread of epidemics on networks, and how the network can become adaptive by rewiring the links according to the state of the individuals. The authors proposed three types of rewiring processes. Given rewiring probability, every susceptible individual drops its link with an infected neighbor and rewires that link with a randomly selected susceptible individual. The first rewiring process is independent of the state of the individual. Such process leads to a random network with Poissonian node degree distribution. The second rewiring process takes place while the dynamics of the epidemic spread is turned off. In this case, the susceptible individuals form a cluster, which is disconnected from the cluster of infected individuals. The third rewiring process, both the dynamics of the epidemic spreading and link rewiring take place simultaneously resulting in the creation of two clusters (one for susceptible individuals and the other for the infected individuals) that are highly internally connected, while the clusters are weakly interconnected. This process leads to decrease the epidemic spreading chances due to the isolation of the infected individuals. However, the cluster of susceptible individuals is highly connected with large variance of node degrees, which increase the epidemic spreading chances. The authors concluded that the local effect of rewiring suppresses the epidemic, while the topological effect increases the chance of the epidemic spreading.

Jiang and Dong in [53] proposed an optimal concept of control measures to control the spread of SARS outbreaks in minimum time. They found that Bang-Bang controller is the optimal solution when the objective is to minimize the lifetime of an outbreak. They only

focused on a subsystem of the compartmental model, which only includes the exposed class E(t) and the infectious class I(t).

V. Marceau et al. studied the coevolution of SIS disease and the network topology simultaneously in [72]. They introduced an adaptive rewiring rule through which the topology changes with time. In their rewiring rule, susceptible individuals replace their links with infected neighbors by susceptible neighbors that are randomly chosen at a predefined rate. The results showed that during the initial phase of an epidemic, number of links that connect susceptible individuals with infectious individuals drop quickly, and hence the susceptible individuals compose a strong community, which is connected with very few infectious individuals. Eventually, the epidemic invades the well connected susceptible community causing a sharp drop in number of links in the susceptible community.

Prakash et al. studied the virus propagation on time-varying networks in [94]. They divided the time unit into periods. Each period has its own network adjacency matrix representing the binary contacts among the individuals. They found that the epidemic threshold is the reciprocal of the maximum eigenvalue of the multiplication of the two system matrices representing the two periods. They also extended the same concept to general alternative behaviors during the lifetime of the virus. In addition, the authors proposed different mitigation strategies that are based on topological characteristics, namely the average node degree, and the maximum node degree, and finally a greedy immunization strategy that causes the largest drop in the maximum eigenvalue of the system matrix.

Bondes *et al.* [18] studied the relationship between the communicable diseases and the evolution of social systems. They proposed a game theoretic approach to find the social behavior of a host during the prevalence of communicable diseases. In addition, the authors presented the relationship between the strength of the diseases and the contact rates of the hosts.

Reluga addressed the effect of social distancing on the spread of SIR epidemics. For example, he used the differential game theory to find the social distancing pattern in population-based model [95, 96]. The solution represents the equilibrium at which excess social dis-

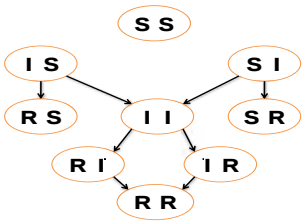


Figure 3.1: *The Markov chain state diagram for a simple network composed of two individuals (N=2) that are in contact. Number of states equals* $3^2 = 9$. *The state SS does not have any transition with other states because all the individuals are susceptible and there is no infection incidence presented in the network to cause the spreading of epidemics.*

tancing does not improve the solution.

3.2 Continuous time Markov chain SIR epidemic model

In this section, we discuss the exact continuous time Markov chain model and its complexity. The continuous time Markov chain model describes the epidemic spread process accurately, based on the fact that each individual in the network is either susceptible, infected, or recovered. An infected individual infects its susceptible neighbors, each with a rate β. Additionally, the infected individual cures itself with cure rate δ. Therefore, there are 3^N different states for any network with N individuals. In addition, there are $\binom{N}{j}2^{N-j}$ different states with j infected individuals. Figure 3.1 shows an example of the state diagram when the spread of epidemics takes place on a network with two individuals that are in contact. To analytically describes the exact Markov chain model, let us define the network state X with 3^N configurations, i.e. X_1, X_2, ... X_{3^N}. Each network state describes the state of each individual $X = \{x_1, x_2, ...x_N\}$ where $x = \{s, i, r\}$. Additionally, the probability that the network is in state X is given by $W_X = p(X = \{x_1, x_2, ...x_N\})$. To be more precise, the rate of change among the different states is described by the infinitesimal Q matrix, which has dimensions of $3^N \times 3^N$. Let $q_{X,Y}$ be an element in the Q matrix representing the rate of change from network state X to network state Y. The Q matrix is described as follows:

* $q_{X,Y} = \delta$ whenever network state X has an infected individual that is cured in network

state Y, while other individuals do not change their states.

* $q_{X,Y} = \beta \sum_{l=1}^{N} a_{m,l}\delta_{i_l,1}$ whenever a susceptible individual m in network state X is connected with infected neighbors $l \in N$ and it becomes infected in state Y, while other individuals do not change their states. Note that $a_{m,l}$ equals 1 when individuals m and l are connected, and $\delta_{i_l,1}$ is the Kronecker delta function and equals 1 if individual l is infected ($i_l = 1$).

* $q_{X,X} = -\sum_{Y \neq X} q_{X,Y}$.

* $q_{X,Y} = 0$ otherwise.

At any time t, the network is in any state X, with a given probability $W_X(t)$ such that $\sum_{X \in \{X_1,\ldots,X_{3^N}\}} W_X(t) = 1$. The rate of change of every network state is governs by the following differential equation:

$$\frac{dW^T(t)}{dt} = W^T(t)Q. \tag{3.7}$$

The solution of the differential equation is as follows:

$$W^T(t) = W^T(0)e^{Qt} \tag{3.8}$$

where $W^T(t)$ is the transpose of the state probability vector $W(t)$ at time t, and $W^T(0)$ is the transpose of the initial state probability vector.

3.3 Individual-based SIR approach

In this dissertation, we present an individual-based SIR approach to model the spread of epidemics in networks. A network is composed of nodes and links. A node represents an individual and the link represents the contact between a pair of individuals. The new approach overcomes the shortcomings of the homogeneous and heterogeneous mean field approaches. In the individual-based approach, each individual can be either susceptible S,

infected I or recovered R with a given probability for each state. The new approach is inspired by the continuous-time Markov chain SIR model, which is discussed in Section 3.2, and it aims to decrease the complexity of the solution from exponential $O(3^N)$ to polynomial $O(N)$. Therefore, instead of considering the combinatorial states of the individuals in the network, we study each individual specifically [77], by decomposing $Q_{3^N \times 3^N}$ matrix to N infinitesimal matrices, each with three states as follows:

$$q_v(t) = \begin{bmatrix} -\beta \sum_z a_{v,z}\delta_{i_z(t),1} & \beta \sum_z a_{v,z}\delta_{i_z(t),1} & 0 \\ 0 & -\delta & \delta \\ 0 & 0 & 0 \end{bmatrix}$$

where $a_{v,z}$ represents the contact level between individual v and individual z in a weighted or unweighted network, and the Kronecker delta function $\delta_{i_z(t),1} = 1$ represents the event that individual z is infected and zero otherwise. In exact Markov chain model, the infection event represents a condition given the neighbor individual is infected, and conditioning on every individual in the network leads to 3^N states in Markov chain (review Section 3.2). In individual-based approach, instead of conditioning on the state of every individual, we replace the actual random infection rate with its effective average infection rate,

$$E[\beta \sum_z a_{v,z}\delta_{i_z(t),1}] = \beta \sum_z a_{v,z} E[\delta_{i_z(t),1}] \tag{3.9}$$

where the infection rate β and the network topology are constant. Therefore $E[\delta_{i_z(t),1} = i_z(t) = 1] = p(i_z(t) = 1)$ is the probability that the neighbor individual z is infected. Replacing the actual random infection rate with its effective rate is basically a mean field approximation, and therefore the effective $q_v^{eff}(t)$ infinitesimal matrix is obtained and has the following expression:

$$q_v^{eff}(t) = \begin{bmatrix} -\beta \sum_z a_{v,z} p(i_z(t) = 1) & \beta \sum_z a_{v,z} p(i_z(t) = 1) & 0 \\ 0 & -\delta & \delta \\ 0 & 0 & 0 \end{bmatrix}.$$

For every individual v, we derive a system of differential equations by applying the general differential equation in 3.7 using the effective $q_v^{eff}(t)$ infinitesimal matrix as follows:

$$\frac{dState_v^T(t)}{dt} = State_v^T(t) q_v^{eff}(t) \tag{3.10}$$

where $State_v^T(t) = [S_v(t)\ I_v(t)\ R_v(t)]$ is the vector of the state probabilities of individual v. The obtained differential equations are as follows:

$$\frac{dS_v(t)}{dt} = -S_v(t)\beta \sum_{z \in N} a_{v,z} I_z(t), \tag{3.11}$$

$$\frac{dI_v(t)}{dt} = S_v(t)\beta \sum_{z \in N} a_{v,z} I_z(t) - \delta I_v(t), \tag{3.12}$$

$$\frac{dR_v(t)}{dt} = \delta I_v(t). \tag{3.13}$$

At any time t, each individual v will be in any of the states with total probability of 1, $S_v(t) + I_v(t) + R_v(t) = 1$. In addition, the sum of rates of changes in the state probabilities is zero $\frac{dS_v(t)}{dt} + \frac{dI_v(t)}{dt} + \frac{dR_v(t)}{dt} = 0$. Therefore, we only solve $2N$ simultaneous differential equations instead of $3N$. Figure 3.2 shows the time evolution of new infected individuals in assortative and disassortative preferential attachment (PA) networks with different $< k >$=4, 8, 12, 16 and 20 given $\beta = 0.1$ and $\delta = 0.2$.

3.3.1 Numerical evaluations

We performed Monte Carlo (MC) simulation to evaluate the accuracy of the individual-based SIR approach (numerical solution (NS)). In Monte Carlo approach, every infected individual tries to infect each of its susceptible neighbors with infection rate β, and it cures itself with rate δ. Figure 3.3 represents the total incidences with respect to the number of individuals in the networks when an epidemic spreads on random scale-free networks [10] with $N = 10^4$ and exponents ν equal 2.6, 2.9, 3.3 and 3.6 given different cure rate to infection rate $\frac{\delta}{\beta}$ for both MC and NS. The results are averaged over 100 runs. For every cure rate

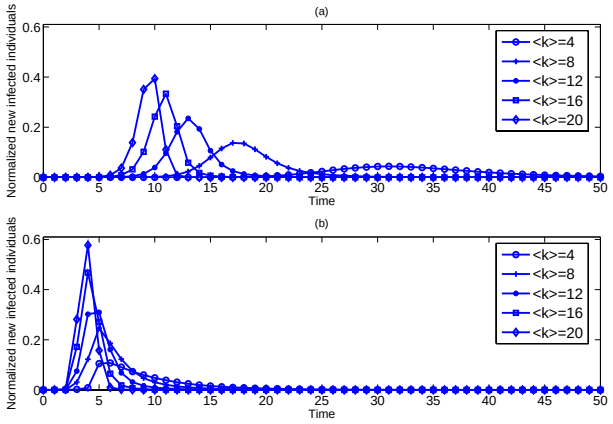


Figure 3.2: *Normalized new infected individuals as a function of time for* $\beta = 0.1$ *and* $\delta = 0.2$ *given correlated networks with* $N = 10^4$ *and different average node degree* $< k >$. *Two different types of correlated networks are simulated (a) assortative preferential attachment (PA) networks, and (b) disassortative preferential attachment (PA) networks. The peak of the new infected individuals in disassortative networks leads to the corresponding peak in assortative networks.*

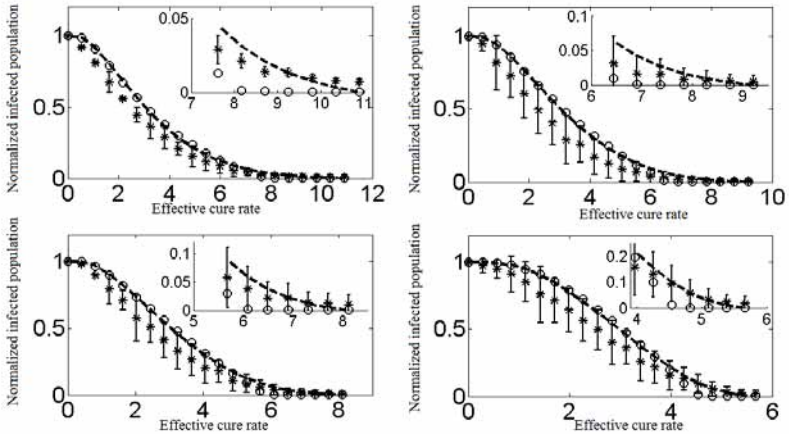


Figure 3.3: *Normalized total incidences of infection given different effective cure rate* $\frac{\delta}{\beta}$ *on random scale-free networks with different exponents* ν *(top left* $\nu = 2.6$, *top right* $\nu = 2.9$, *bottom left* $\nu = 3.3$, *bottom right* $\nu = 3.6$*). Each network has* 10^4 *individuals. The results are averaged over 100 runs. The dashed line represents the numerical solution (NS) of the individual-based SIR approach, the 'o' symbol represents the heterogeneous mean field approach results, and the '*' symbol represents the average of Monte Carlo (MC) results. The error bar is the standard deviation of MC results. The insets show how the individual-based approach outperforms the heterogeneous mean field approximation for the values* $\frac{\delta}{\beta}$ *that are close to the reciprocal of the epidemic threshold (*$\frac{1}{\tau}$*).*

δ and infection rate β, we compute the normalized total incidences ρ so that we obtain the relationship between $\frac{\delta}{\beta}$ and ρ. In MC, each simulation starts with a single infected incidence that is randomly chosen among the individuals, while in NS, each individual is initially infected with probability 10^{-4}. The error bar represents the standard deviation of MC trials. We notice that NS upper-bounds MC simulations for a large range of $\frac{\delta}{\beta}$. In Section 3.5, we analyze the deviation between the individual-based approach and Markov chain model and we show that the individual-based approach upper-bounds Markov model results. The insets in Figure 3.3 show that for $0.7\lambda_{max} \leq \frac{\delta}{\beta} \leq \lambda_{max}$, NS results approach MC simulations.

Additionally, we compare the individual-based SIR approach with the heterogeneous mean field approach discussed in Section 3.1 given the same range of cure rate to infection rate $\frac{\delta}{\beta}$ values. We use the same random scale-free networks with the same exponent values. We also assume that the initial probability of infection for every individual is 10^{-4}. In Figure 3.3, the results show that the heterogeneous mean field approach does not show any infection incidence for higher values of $\frac{\delta}{\beta}$ that are near the reciprocal of the epidemic threshold. On the other hand, the results of the individual-based approach are closer to MC simulation results, and both show the existence of incidences with non-zero values. Basically, the difference between the two approaches comes from the epidemic thresholds and the network representation in each approach. The range of $\frac{\delta}{\beta} \leq \frac{1}{\tau}$ in the individual-based approach is larger than the range in heterogeneous mean field approach, which is observed in Figures 3.3. Therefore, the individual-based approach captures more chances for the spread of epidemics with a broader range of $\frac{\delta}{\beta}$. Additionally, in the heterogeneous mean field approach, individuals are represented through their node degrees, which means that all individuals with the same node degree have the same probability of infection, while in the individual-based approach, every individual is studied separately. Thus the heterogeneous mean field approach neglects the centrality properties of the individuals in the network, while the individual-based approach not only considers the individual probability of infection, but

also distinguishes among the individuals' topological characteristics.

3.4 Properties of the individual-based SIR approach

In this section, we derive some useful properties from the individual-based SIR approach. Those properties describe the infection process more fully, and they relate the infection process with the contact network topology over which an outbreak takes place. First, we derive the epidemic threshold, and we show how it is related to the network topological properties. We also derive the condition under which the infection reaches its peak value with time, and how the peak infection value is related to the epidemic threshold. Moreover, we study the role of the network topology on the spread of epidemics.

3.4.1 Epidemic threshold

In epidemiology literature, the epidemic threshold is a function of a quantity called basic reproductive number R_0 [6, 8, 57, 67, 73]. The basic reproductive number is defined as the average number of secondary infected individuals when a single individual is infected initially. Mathematically, the reproductive number is $\frac{<k>\beta}{\delta}$ where $< k >$ is the average connectivity in the network. Therefore, if R_0 is greater than 1, the epidemic spreads on the network and vice versa. Many networks with individuals characterized by a wide range of contacts (i.e. some individuals have few contacts and other individuals have many contacts) can have the same average level of contact [30, 68, 69]. The information on the average contact does not provide any indication of the contact level distribution. Being R_0 only dependent on the average level of contact, it will be the same for all networks with the same average level of contact, independently of their contact level distribution. Recent works showed that the epidemic threshold is a function of non-trivial network characteristics.
Below, we confirm that the epidemic threshold is a function of the maximum eigenvalue of the matrix representing accurately the contact level among the individuals [77, 106]. To compute the threshold, we follow the analysis presented in [16]. We assume that the

initial fraction of infected individuals is very small and therefore $S_v(0) \simeq 1$. The differential Eq. 3.36 is written as follows:

$$\frac{dI_v(t)}{dt} \simeq \sum_z \tilde{L}_{v,z} I_z(t) \tag{3.14}$$

where the element $\tilde{L}_{v,z} = \beta a_{v,z} - \delta\delta_{v,z}$ is the entry of the Jacobian matrix $\tilde{L} = \{\tilde{L}_{v,z}\} = \beta A - \delta I_{N\times N}$, and $\delta_{v,z}$ is the Kronecker delta function and equals 1 forall $v = z$. Since any element $a_{v,z}$ of the symmetric adjacency matrix A of the binary contact network is either 0 or 1, and according to Frobenius theorem, the maximum eigenvalue $\lambda_{max,A}$ of A is positive and real, the eigenvalues of the matrix $\tilde{L}$ have the form of $\beta\lambda_{i,A} - \delta$, and the eigenvectors are the same as those for the adjacency matrix A. Thus, the stability condition of the solution $I = 0$ of the differential Eq. 3.14 is $-\delta + \beta\lambda_{max,A} < 0$, and the SIR individual-based epidemic threshold τ for any undirected network becomes:

$$\frac{1}{\lambda_{max,A}} = \tau \tag{3.15}$$

As a consequence, an outbreak will occur if and only if $\frac{\beta}{\delta} > \tau$.

3.4.2 The existence of a maximum number of infected individuals

The number of infected individuals increases in time following a certain profile [11] depending on the infection strain. Below, we derive the condition for which a maximum number of infected individuals occurs, and how the condition is related to the epidemic threshold. Let $u^T I(t) = \sum_v I_v(t)$ be the total number of infected individuals in the network, where u^T is the transpose of a vector of ones $u^T = [1\ 1\ ...1]$. The existence of a maximum value for $I(t)$ is determined through $\frac{du^T I(t)}{dt} = \sum_v \frac{dI_v(t)}{dt} = 0$, and we obtain:

$$\sum_v \left[S_v(t)\beta \sum_z a_{v,z} I_z(t) - \delta I_v(t) \right] = 0 \tag{3.16}$$

By rewriting Eq. 3.16 in the matrix form, we obtain the following equation:

$$\left[\beta S^T(t)A - \delta u^T\right] I(t) = 0 \tag{3.17}$$

There are two possible solutions for Eq. 3.17: 1) $I(t)$ equals zero, which happens when the network becomes cured, or 2) $\beta S^T(t)A - \delta u^T$ equals zero. The second solution derives a condition for the existence of a positive maximum value of $I(t)$. Consequently, the second solution $AS(t) = \frac{\delta}{\beta}u$ has the form of $Mx = \rho x$, where x and ρ are an eigenvector and an eigenvalue of the arbitrary matrix M, respectively. The vector $S(t)$ is equal to the vector u only if $\frac{\delta}{\beta}$ is equal to the maximum eigenvalue $\lambda_{max,A}$ of A, which follows Frobenius theorem and takes place for $t \to 0$ and $S(0) \to 1$. Moreover, this solution proves the existence of the epidemic threshold shown in inequality 3.15 whenever $\frac{\delta}{\beta} < \lambda_{max,A}$, and therefore the epidemic spreads in the network, $S_v(t) \leq 1$ forall v, and the maximum number of incidences takes place before the network becomes cured.

3.4.3 The effect of the network spectrum

This subsection addresses the effect of the network topology on the spread of epidemics. Rigorously, we show that the eigenvalues and the corresponding eigenvectors reveal the role of the centrality properties of the individuals in spreading the epidemics in networks. Below, we mathematically derive the effect of the centrality properties of the networks on the spread of epidemics followed by an interpretation for the final mathematical formula 3.21, and we show how the eigenvector corresponding to the maximum eigenvalue can predict the probability of infection of the individuals in the network.

Recall that u^T is the transpose of a vector of ones ($u^T = [1\ 1\ ...1]$), and A to be the adjacency matrix of a binary contact network. Thus, we can write the total rate of change of infected individuals $u^T I(t)$ as follows:

$$\frac{du^T I(t)}{dt} = \beta(u^T - I^T(t) - R^T(t))AI(t) - \delta u^T(t)I(t). \tag{3.18}$$

If we denote the vector of node degrees $D = u^T A$, and the eigenvalue decomposition of the adjacency matrix $A = U\Lambda U^T$, we can rewrite the differential Eq. 3.18 as follows:

$$\begin{aligned}\frac{du^T I(t)}{dt} = \quad & (\beta D - \delta u)^T I(t) - \beta (U^T I(t))^T \Lambda (U^T I(t)) \\ & -\beta (U^T R(t))^T \Lambda (U^T I(t))\end{aligned} \tag{3.19}$$

Let x_z be the z^{th} element in the vector $U^T I(t)$, and let y_z be the z^{th} element in the vector $U^T R(t)$. Now we can rewrite the differential equation as follows:

$$\frac{du^T I(t)}{dt} = (\beta D - \delta u)^T I(t) - \beta \sum_{z=1}^{N} \lambda_z x_z^2 - \beta \sum_{z=1}^{N} \lambda_z x_z y_z \tag{3.20}$$

To relate I_{max} with the spectrum λ_z and the eigenvectors U, let $\frac{du^T I(t)}{dt}$ equal zero, and therefore we can obtain the following equation:

$$\sum_{v=1}^{N} (d_v - \frac{\delta}{\beta}) I_{v_{max}} = \sum_{z=1}^{N} \lambda_z x_z^2 - \sum_{z=1}^{N} \lambda_z x_z y_z \tag{3.21}$$

The fact that the matrix A is symmetric and therefore λ_{max} is a positive eigenvalue and elements of the corresponding eigenvector are positive as well is used to understand Eq. 3.21 as follows: As β increases, both the LHS and the RHS increase too. We can also see that the vector $I(t)$ is proportional to the eigenvectors of the adjacency matrix A, while the coefficients of the proportion are the eigenvalues; however, the dominant term in the RHS is $\lambda_{max} x_{\lambda_{max}}^2 = \lambda_{max} (U_{\lambda_{max}}^T I(t))^2$ since both the maximum eigenvalue λ_{max} and the elements of the corresponding eigenvector $U_{\lambda_{max}}^T$ are positive. Therefore, the vector $I(t)$ is increasingly more aligned with the eigenvector corresponding to the maximum eigenvalue on the RHS. Thus, the elements of the eigenvectors corresponding to the maximum eigenvalue provide an estimate of the probability of infection for each individual in the network.

To conclude, this property reveals the role of the centrality properties of the individuals in spreading the infection across the contact network.

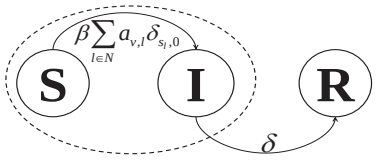


Figure 3.4: *The state transition diagram of an individual v from susceptible state $s_v = 1$ to infection state $s_v = 0$.*

3.5 Markov chain model and mean field approximation

In this section, we analytically show the differences between the continuous time Markov chain epidemic model that is discussed in Section 3.2 and the individual-based SIR approach, and the effect of the mean field approximation on the outcome of the model.

Let us first consider the state transition of an individual from susceptible to infection. Let $s_v = 1$ if individual v is susceptible, and $s_v = 0$ if it is infected. In the SIR model, the transition happens from susceptible state to infection state, while the opposite does not happen as shown in Figure 3.4. Thus a susceptible individual v becomes infected with rate $\beta \sum_l a_{v,l} \delta_{s_l,0}$ (see Figure 3.4), where $\delta_{s_l,0}$ is the Kronecker delta function and it equals 1 if the neighbor l is infected, i.e. $s_l = 0$. Therefore, the susceptible individuals do not increase with time. In other words, the change in the susceptible state $s_v = 1$ over a small time interval Δt is defined as follows:

$$\frac{s_v(t + \Delta t) - s_v(t)}{\Delta t} = -\delta_{s_v(t),1} \beta \sum_{l=1}^{N} a_{v,l} \delta_{s_l(t),0} \tag{3.22}$$

By taking the average on both sides, and considering $E[s_v(t)] = p(s_v = 1) = S_v(t)$, we obtain the following equation:

$$\frac{S_v(t + \Delta t) - S_v(t)}{\Delta t} = -E[\delta_{s_v(t),1} \beta \sum_{l=1}^{N} a_{v,l} \delta_{s_v(t),0}] \tag{3.23}$$

The above equation is only applied for any $v \neq l$. Since we assume that the infection

rate does not change with time, and the topology is static, we process the expectation only on the Kronecker delta functions as follows:

$$\begin{aligned} E[\delta_{s_v(t),1}.\delta_{s_l(t),0}] &= E[(s_v(t)=1)\cap(s_l(t)=0)] \\ &= p(s_v(t)=1, s_l(t)=0) \\ &= p(s_v(t)=1 \mid s_l(t)=0)p(s_l(t)=0). \end{aligned}$$

In Eq. 3.23, let $\Delta t \to 0$, and we obtain the following differential equation:

$$\frac{dS_v(t)}{dt} = -\beta \sum_{l=1}^{N} a_{v,l} p(s_v(t)=1 \mid s_l(t)=0)p(s_l(t)=0) \tag{3.24}$$

The conditional probability represents the probability that individual v is susceptible given that the neighbor individual l is infected. Recalling Eq. 3.11,
which is $\frac{dS_v(t)}{dt} = -S_v(t)\beta \sum_{l=1}^{N} a_{v,l} I_l(t)$ in the individual-based approach, and comparing it with Eq. 3.24, we notice that the mean field theory assumes that events are independent as follows:

$$\begin{aligned} p(s_v(t)=1 \mid s_l(t)=0)p(s_l(t)=0) &= p(s_v(t)=1, s_l(t)=0) \\ &= p(s_v(t)=1)p(s_l(t)=0) \\ &= S_v(t)I_l(t). \end{aligned}$$

Thus, we address the relationship between the conditional probability $p(s_v(t)=1 \mid i_l(t)=1)$ and the probability $p(s_v(t)=1)$ for an individual v and its neighbor l as follows: we assume that the state transition takes place over the time interval Δt. We also assume that individual v is susceptible at time t. Hence, the conditional probability can be rewritten as $p(s_v(t+\Delta t) \mid s_v(t), i_l(t))$, while the probability $p(s_v(t)=1)$ can be rewritten as $p(s_v(t+\Delta t) \mid s_v(t))$. Using the total probability theory, $p(s_v(t+\Delta t) \mid s_v(t))$ is as follows;

$$
\begin{aligned}
& p(s_v(t+\Delta t) \mid s_v(t)) = \\
= \quad & p(s_v(t+\Delta t) \mid s_v(t), s_l(t+\Delta t))p(s_l(t) \mid s_v(t)) \\
& + p(s_v(t+\Delta t) \mid s_v(t), r_l(t))p(r_l(t) \mid s_v(t)) \\
& + p(s_v(t+\Delta t) \mid s_v(t+\Delta t), i_l(t))p(i_l(t) \mid s_v(t)) \\
= \quad & p(s_l(t) \mid s_v(t)) + p(r_l(t) \mid s_v(t)) + p(s_v(t+\Delta t) \mid s_v(t), i_l(t))p(i_l(t) \mid s_v(t)) \\
= \quad & 1 - p(i_l(t) \mid s_v(t)) + p(s_v(t+\Delta t) \mid s_v(t), i_l(t))p(i_l(t) \mid s_v(t))
\end{aligned}
\tag{3.25}
$$

Recall the following inequality:

$$p(s_v(t+\Delta t) \mid s_v(t), i_l(t)) \leq 1$$

Multiply both sides by $(1 - p(i_l(t) \mid s_v(t)))$, after rearranging, we obtain the following inequality

$$p(s_v(t+\Delta t) \mid s_v(t), i_l(t)) \leq 1 - p(i_l(t) \mid s_v(t)) + p(s_v(t+\Delta t) \mid s_v(t), i_l(t))p(i_l(t) \mid s_v(t))$$

which is as follows:

$$p(s_v(t+\Delta t) \mid s_v(t), i_l(t)) \leq p(s_v(t+\Delta t) \mid s_v(t)).$$

The last inequality is valid since an infected neighbor l does not increase the probability of an individual v to remain susceptible. Therefore in the Markov chain model, the absolute rate at which a susceptible individual changes its state to infection state is lower than the corresponding rate in the individual-based approach.

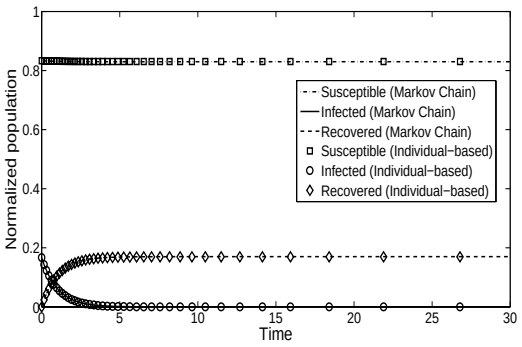
We conclude that for any individual v, the susceptible probability in the individual-based SIR approach lower-bounds the susceptible probability in the Markov chain model. Therefore, we also conclude that the infection probability in the individual-based SIR approach upper-bounds the infection probability in the Markov chain model. Consequently, the probability of recovery in the individual-based SIR approach also upper-bounds the probability of recovery in the Markov chain model.

3.5.1 Numerical evaluations

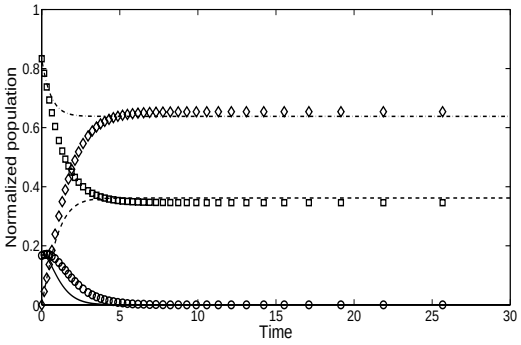
We compare the individual-based approach with the continuous time Markov chain model to study the deviation between the two approaches. In Figures 3.5(a)- 3.5(e), we numerically evaluate the spread of epidemics on a 6-node ring network where initially a single individual is infected and $\frac{\delta}{\beta}$ equals 10^2, 1.6667, 1.25, 0.5, and 10^{-2}, respectively. Note: the Markov chain model can only be modeled on smaller networks since the number of states exponentially increases with the number of individuals in the network. In Figure 3.5(a), we notice that the initially infected individuals recovered and the epidemic dies out, and there is no new infection since $\frac{\delta}{\beta}$ is above the reciprocal of the epidemic threshold ($\lambda_{max}^{ring} = 2$). However, a deviation is noticeable in Figures 3.5(c) and 3.5(d) between the individual-based approach and the Markov model, and it is due to the mean field approximation used in the individual-based approach. In particular, we observe in the same Figures that the susceptible population in the individual-based approach lower-bounds the corresponding population in the Markov chain model, while the infected population in the individual-based approach upper-bounds the corresponding population in the Markov chain model. In Figure 3.5(e), the individual-based approach almost coincides with the continuous time Markov chain model. The simulations in Figure 3.3 show that total incidences in the individual-based approach upper-bounds the total incidences in Monte Carlo simulations. Such a result agrees with the theoretical prediction that the probability of infection in the individual-based approach upper-bounds the probability of infection in Markov chain model.

3.6 Viral conductance VC_{SIR}

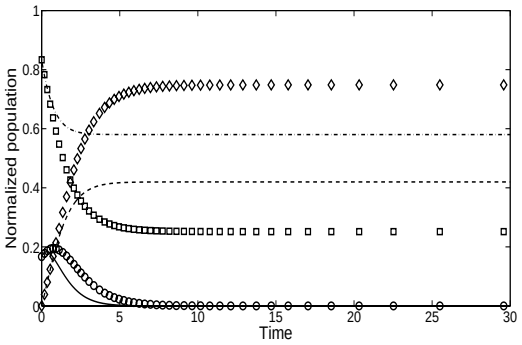
We employed the individual-based SIR model to assess the robustness of a complex network such that the total number of new infected individuals reflects the vulnerability of the network to the spread of epidemics given any infection strength. In chapter 2, the measure VC_{SIS} is introduced that takes into account the number of infected individuals at steady state for the susceptible/infected/susceptible compartmental model. In this chap-



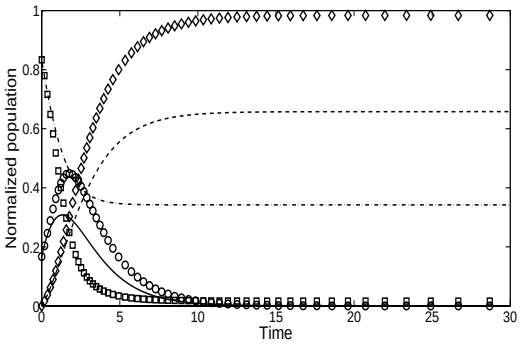
(a) $\frac{\delta}{\beta} = 10^2$



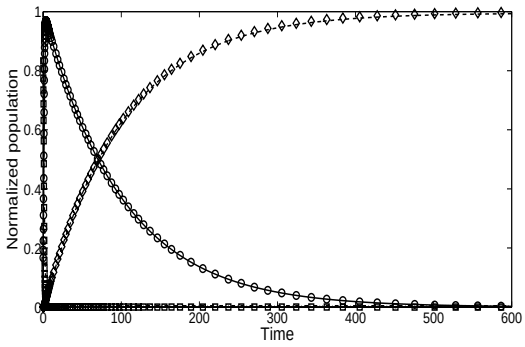
(b) $\frac{\delta}{\beta} = 1.6667$



(c) $\frac{\delta}{\beta} = 1.25$



(d) $\frac{\delta}{\beta} = 0.5$



(e) $\frac{\delta}{\beta} = 10^{-2}$

Figure 3.5: *Comparison between the continuous time Markov chain model and the individual-based approach given a 6-node ring network and* $\frac{\delta}{\beta} = 10^2, 1.6667, 1.25, 0.5,$ *and* 10^{-2}.

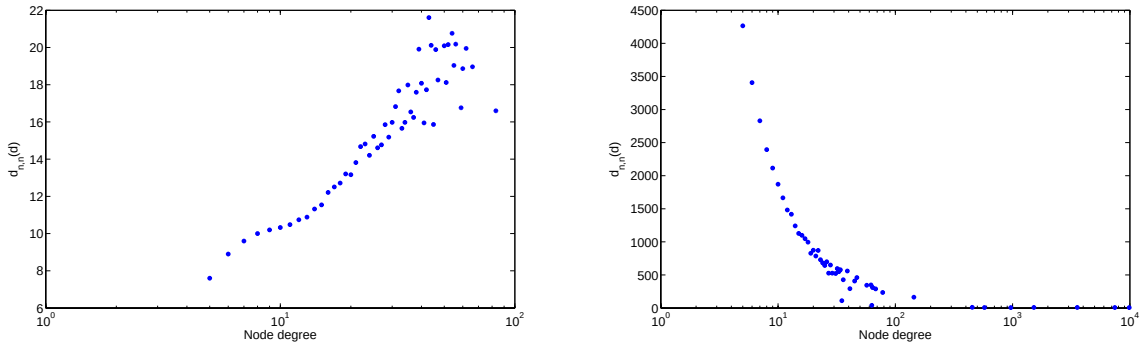
ter, we extend the concept of the viral conductance, and we introduce a new robustness assessment measure VC_{SIR} with respect to the spread of SIR epidemics. We define the new metric as the ability of a network to resist the spread of an SIR epidemic given all possible effective cure rates. The smaller the measure VC_{SIR}, the robust is the network to resist the spread of the epidemics. Mathematically, we define the assessment measure VC_{SIR} by fixing $\beta = \frac{1}{\lambda_{max,A}}$ and for a given cure rate δ, the total number of new infected individuals is $\int_0^{t^*} \sum_k S_k(t)\beta \sum_j a_{k,j} I_k(t,\delta)dt$. By integrating over the defined range of cure rate $0 \leq \delta \leq 1$, we obtain VC_{SIR} as follows:

$$VC_{SIR} = \int_0^1 \int_0^{t^*} \frac{1}{N} \sum_k S_k(t)\beta \sum_j a_{k,j} I_j(t,\delta)dtd\delta \tag{3.26}$$

The integration term represents the vulnerability of the network, and hence the reciprocal of the vulnerability measure is the robustness measure VC_{SIR}.

3.6.1 Numerical evaluations

We use the new measure VC_{SIR} to evaluate the robustness of correlated networks, these in which node degree correlation is observed. They are also classified as assortative and disassortative networks. For example, social networks are classified as assortative networks, while technological and biological networks are classified as disassortative networks [84]. In assortative networks, individuals of small node degree are connected with other individuals of small node degree, while individuals with large node degree are connected with other individuals with large node degree. On the other hand, the opposite is true for disassortative networks. Pearson assortativity coefficient [29, 84] was proposed to characterize the node degree correlation numerically. However, it does not give an accurate measure for networks with complicated degree correlation functions. To accurately describe the degree correlations, we evaluate the average connectivity of the neighbors of an individual k by following the technique presented in [11, 89, 104]:



(a) Assortative PA network

(b) Disassortative PA network

Figure 3.6: *Node degree as a function of average neighbors connectivity $d_{n,n}(d)$ of individuals with the same node degree for a sample of correlated PA network with $N = 10^4$ and $< k >= 8$.*

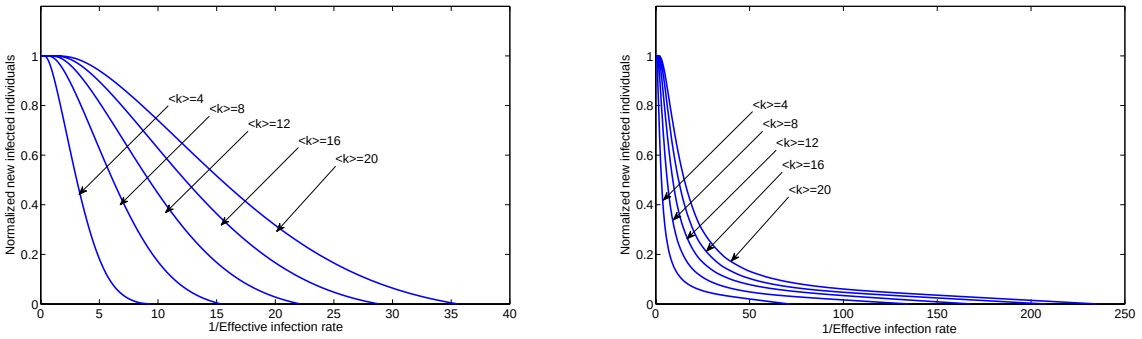
$$d_{n,n,k} = \frac{1}{d_k} \sum_{j \in neighbors(k)} d_j \tag{3.27}$$

The average connectivity of neighbors of an individual is averaged overall of all individuals for a given node degree d,

$$d_{n,n}(d) = \frac{1}{N_d} \sum_{k/d_k=d} d_{n,n,k} \tag{3.28}$$

where N_d is the number of individuals of degree d. Figures 3.6(a) and 3.6(b) show two examples for correlated networks, one for an assortative network and the other for a disassortative network, respectively.

We focus on the robustness assessment of correlated PA networks. We generate assortative and disassortative PA networks using the algorithm in [47]. The algorithm starts with a connected graph with $m_0 \ll N$ individuals. Every new individual is connected to the already existing individuals through two stages: In the first stage, a new individual is connected to an existing individual k with probability $\pi_k = \frac{d_k}{\sum_j d_j}$; in the second stage, a new link between the new individual and one of the neighbors s of the chosen individual k in the first stage is added with probability $p_s = \frac{d_s^\alpha}{\sum_{v \in \Gamma_k} d_v^\alpha}$, where α is an assortative tunning coefficient, and Γ_k is the set of neighbors of individual k chosen in the first stage.



(a) Assortative PA network

(b) Disassortative PA network

Figure 3.7: *Normalized total of new infected individuals as a function of the effective cure rate for correlated preferential attachment networks given* $N = 10^4$ *and different average node degree* $< k >$. *The curve starts from the point where* $\frac{\delta}{\beta} = 0$, *and the normalized total new infected cases is 1, and then it decreases until it reaches the value zero when the value of* $\frac{\delta}{\beta}$ *equals the spectral radius of the network.*

To simplify the evaluation of numerical results, both the constructed assortative and disassortative networks have the same number of individuals N and links L with average node degrees $< k >$=4, 8, 12, 16 and 20. Next, we apply the new measure VC_{SIR} in Eq. 3.26 to quantitatively assess the robustness of both assortative and disassortative networks. All the simulations are averaged over 10 runs. Figures 3.7(a) and 3.7(b) show the numerical simulations of the spread of an epidemic for $0 \leq \frac{\delta}{\beta} \leq \lambda_{max,A}$ on assortative and disassortative preferential attachment PA networks given different average node degrees $< k >$, where $\frac{\delta}{\beta}$ is the inverse of the effective infection rate.

Table 3.1 summarizes the values of VC_{SIR} for both types of networks for different average node degrees. We notice that the disassortative networks have lower values of robustness measure VC_{SIR} than those of assortative networks regardless of the average node degree value. In addition, the VC_{SIR} value decreases with increases in L (i.e.$< k >$) due to the increase in the effective spreading rate of any infected individual for its susceptible neighbors. Moreover, in Fig. 3.2, we observe that the peaks of normalized new infected individuals in disassortative networks are greater than the peaks in assortative networks; meanwhile, the

$< k >$	Assortative networks	Disassortative networks
4	3.32	6.54
8	6.54	12.58
12	9.76	17.65
16	12.98	23.47
20	16.22	28.68

Table 3.1: *Robustness measure VC_{SIR} for assortative and disassortative PA networks given different average node degrees $< k >$. The network size is $N = 10^4$.*

peaks in disassortative networks lead the corresponding peaks in assortative networks. In other words, an epidemic widely spreads in disassortative networks, and it spread faster than in assortative networks. Fig. 3.2 also reveals insights about any future immunization strategy that could be applied to both networks. For example, we can assume that immunization strategies on assortative and disassortative networks are different. Therefore, in assortative PA networks, mitigation strategies are going to be more effective than in disassortative PA networks.

3.7 Optimal mitigation of epidemics in weighted networks

We adapt the individual-based SIR approach to find the optimal mitigation strategy to minimize the final size of infection in social networks through the optimal control theory. Thus, to apply the optimal control theory in social networks, we consider the spread of SIR epidemic in the network as a system, and the total number of infection cases as the state of the system, while the controller variable is the reduction in the weights of the social network leading to slow/reduce the spread of epidemics. Unlike the trivial methods in which infected individuals are isolated to mitigate the epidemic, the new approach preserves a global minimum contact level, which mainly depends on the social activities of individuals within the social community. Meanwhile, our approach minimizes the reduction in the contact weights. Hence, the optimal control formulation addresses the trade-off between minimization of total infection cases and minimization of contact weights reduction. Consequently, the solution

represents an optimal adaptive weighted contact network.

3.7.1 Weighted SIR epidemic approach

Traditionally, any social network is a weighted graph representing the contact frequency and the proximity among the individuals. We denote $w_{m,n}(t)$ to be the contact weight between individuals m and n at time t such that $0 \leq w_{m,n}(t) \leq 1$. The spread of infectious disease takes place in social networks due to the contacts among susceptible and infected individuals. Therefore, the actual infection rate from an infected individual n towards a susceptible individual m at time t becomes $\beta w_{m,n}(t)$. Using the system of differential equations presented in 3.3, we obtain the following SIR epidemic approach for a weighted network:

$$\frac{dS_m(t)}{dt} = -S_m(t)\beta \sum_{n \in N} w_{m,n}(t) I_n(t), \tag{3.29}$$

$$\frac{dI_m(t)}{dt} = S_m(t)\beta \sum_{n \in N} w_{m,n}(t) I_n(t) - \delta I_m(t), \tag{3.30}$$

$$\frac{dR_m(t)}{dt} = \delta I_m(t). \tag{3.31}$$

Differential equations in 3.29 - 3.31 represent the system of epuations in the optimal control formulation to minimize both the total infected cases as well as the reduction in the contact weights.

3.7.2 Optimal dynamical weights

In this section, we formulate a continuous time optimal control problem to minimize the total size of infection by properly reducing the contact weights among individuals within a finite time interval $t \in [0, T_{final}]$. Meanwhile, our objective is to minimize the weight reduction with respect to original epidemic-free contact weights ($t = 0$). The weighted contact network is composed of N individuals and L directed links. Initially, we assume that the contact network is undirected weighted graph at time $t = 0$; However, for $t > 0$,

Data input	Definition
T_{final}	Final time
β	Infection rate
δ	Cure rate
α	Global minimum social contact coefficient
$w_{m,n}(0)$	Initial weight $\forall\ m, n \in N$
$S_m(0)$	Initial susceptible probability of individual m
$I_m(0)$	Initial infection probability of individual m
Variables	**Definition**
$w_{m,n}(t)$	Link weight at time t
$S_m(t)$	The susceptible probability of individual m at time t
$I_m(t)$	The infection probability of individual m at time t

Table 3.2: *Definitions of the data inputs and the variables*

the contact network can become directed weighted graph in which $w_{m,n}(t)$ can be different from $w_{n,m}(t)$. During the spread of an epidemic $t > 0$, weights are reduced from their initial epidemic-free values $w_{n,m}(0)$. In particular, we impose two bounds on each weight (m, n): a) $\alpha w_{m,n}(0) \leq w_{m,n}(t)$ and b) $w_{m,n}(t) \leq w_{m,n}(0)$, where $w_{m,n}(0)$ is the original epidemic-free weight value between individual m and individual n, and α is a global minimum social level coefficient $0 < \alpha < 1$. These constraints have direct implications on the network as follows: First, to preserve a minimum contact level among individuals even during epidemics, we introduce a lower bound on the weights, $\alpha w_{m,n}(0) \leq w_{m,n}(t)$. Second, during an epidemic, the level of contact between two individuals can not increase beyond the original level $(w_{m,n}(0))$.

3.7.3 Optimal control formulation

For every individual m, the infection probability $I_m(t)$ as well as the susceptible probability $S_m(t)$ are the state variables, while the weight reduction $(w_{m,n}(0) - w_{m,n}(t))$ is the controller variable. The data inputs and the variables are summarized in Table 3.2. The objective function is given by a weighted sum of the weight reduction cost function, and the total infection size as shown in the following equation:

$$\text{Minimize} \int_0^{T_{final}} A \sum_{m,n \in N} [f(w_{m,n}(0) - w_{m,n}(t))] + B \sum_{m \in N} \beta S_m(t) \sum_{n \in N} I_n(t) w_{m,n}(t) dt \tag{3.32}$$

The first term $f(.)$ is a non negative convex function representing the weight reduction cost function, while the second term represents the new infection cases at time t. The constants A and B are the coefficients for of the weighted sum objective function. The problem constraints are composed of the system of differential equations 3.29 and 3.30, and the following weight constraint:

$$\alpha w_{m,n}(0) \leq w_{m,n}(t) \leq w_{m,n}(0) \; \forall \; m, n \; \in N \tag{3.33}$$

Case study: Homogeneous networks

Homogeneous networks are used to model the spread of infectious diseases in large population [9]. In addition, homogeneous network model is used to model the within-household contact patterns [13, 41, 71]. In a homogeneous network, every individual has the same number of contacts with other individuals in the population. Therefore, the spatial index m is reduced from the susceptible probability, infection probability and the contact weight, and these variables become $S(t)$, $I(t)$, and $w(t)$, respectively. Consequently, the objective function becomes as follows:

$$\text{Minimize} \int_0^{T_{final}} f(w(0) - w(t)) + \beta S(t) I(t) w(t) dt \tag{3.34}$$

The function $f(w(0) - w(t))$ represents the weight reduction cost such that $f(0) = 0$. In addition, we assume that $f(w(0) - w(t))$ is a strictly convex function where $f''(w(0) - w(t)) > 0$. The second term represents the fraction of new infection at time t. Therefore, the main objective is to minimize the weight reduction and the total fraction of new infection for all

$0 < t \leq T_{final}$.

For homogeneous networks, the constraints become as follows:

$$\frac{dS(t)}{dt} = -\beta S(t)I(t)w(t), \tag{3.35}$$

$$\frac{dI(t)}{dt} = \beta S(t)I(t)w(t) - \delta I(t), \tag{3.36}$$

$$\alpha w(0) \leq w(t) \leq w(0). \tag{3.37}$$

Applying Pontryagin's minimum principle [93], we obtain the Hamiltonian function H as follows:

$$\begin{aligned} H(S(t), I(t), w(t), \lambda_I(t), \lambda_S(t)) &= f(w(0) - w(t)) \\ &+\beta S(t)I(t)w(t) - \delta\lambda_I(t) \\ &+(\lambda_I(t) - \lambda_S(t))\beta S(t)I(t)w(t). \end{aligned} \tag{3.38}$$

$$\tag{3.39}$$

The co-state equations and the transversality conditions are as follows:

$$\lambda_S^{'}(t) = -\tfrac{\partial H}{\partial S(t)} = -\beta I(t)w(t) - (\lambda_I - \lambda_S)\beta I(t)w(t), \tag{3.40}$$

$$\lambda_I^{'}(t) = -\tfrac{\partial H}{\partial I(t)} = -\beta S(t)w(t) - (\lambda_I - \lambda_S)\beta S(t)w(t) + \lambda_I(t)\delta. \tag{3.41}$$

$$\lambda_I(T_f) = 0 \tag{3.42}$$

$$\lambda_S(T_f) = 0 \tag{3.43}$$

We proceed with the optimality condition as follows:

$$H(w^*(t), S(t), I(t)) \leq H(w(t), S(t), I(t)) \tag{3.44}$$

where $w^*(t)$ is the optimal weight value at time t such that $H(w^*(t), S(t), I(t))$ is minimized. After substituting the Hamiltonian Eq. 3.38 in the optimality condition, we obtain the following inequality:

$$f(w(0) - w^*(t)) + \psi(t)w^*(t) \leq f(w(0) - w(t)) + \psi(t)w(t) \tag{3.45}$$

where $\psi(t) = \beta S(t)I(t)(1 + \lambda_I(t) - \lambda_S(t))$.

Let $y(t) = w(0) - w(t)$ represents the weight reduction such that $0 \leq y(t) \leq w(0)(1 - \alpha)$, and the inequality becomes as follows:

$$f(y^*(t)) - \psi(t)y^*(t) \leq f(y(t)) - \psi(t)y(t). \tag{3.46}$$

Since $y(t) = 0$ is an admissible point and $f(y(t) = 0) = 0$, therefore, the following inequality holds for all time t:

$$f(y^*(t)) - \psi^*(t)y^*(t) \leq 0 \tag{3.47}$$

Since $f(y(t)) \geq 0$ and $0 \leq y(t) \leq w(0)(1 - \alpha)$, then:

$$f(y^*(t)) \leq \psi^*(t)y^*(t) \tag{3.48}$$

showing that $\psi^*(t)$ is a positive function

$$\beta S^*(t)I^*(t)(1 + \lambda_I^*(t) - \lambda_S^*(t)) > 0. \tag{3.49}$$

Since $S^*(t)$ and $I^*(t) > 0$, then:

$$1 + \lambda_I^*(t) - \lambda_S^*(t) > 0. \tag{3.50}$$

By following the analysis in [59], we proposed two lemmas that are used to show the dynamic behavior of the weights in networks.

Lemma 3.7.1. *The co-state variable $\lambda_I^*(t)$ is strictly positive, $\lambda_I^*(t) > 0$ for $0 \leq t < T_{final}$.*

Proof. First, we apply the following function property: *for any continuous and differentiable function $g(t)$ and $g(t_1) = x$ such that for any $t > t_1$ we have $g(t) > x$, then $\lim_{t \to t_1^+} g'(t) \geq 0$.*

Secondly, recall that $\lambda_I^*(T_{final}) = \lambda_S^*(T_{final}) = 0$ and $\lambda_I^{*'}(T_{final}) = -\beta S^*(T_{final})w(T_{final})$, $\lambda_I^{*'}(T_{final}) < 0$. Thirdly, we derive the proof by contradiction: Let t_a be the time before T_{final} at which $\lambda_I^*(t_a) \leq 0$, and $\lambda_I^*(t) > 0$ for $t_a < t < T_{final}$ as shown in the following two cases:

- $\lambda_I^*(t_a) = 0$

$$\lim_{t \to t_a^+} \lambda_I^{*'}(t) = \lim_{t \to t_a^+} -\beta S^* w^*(1 + \lambda_I^* - \lambda_S^*) + \lambda_I^* \delta \tag{3.51}$$

Recall that $(1 + \lambda_I^* - \lambda_S^*) > 0$, and hence

$\lim_{t \to t_a^+} \lambda_I^{*'}(t) < 0$, which contradicts the above property.

- $\lambda_I^*(t_a) < 0$

$$\lim_{t \to t_a^+} \lambda_I^{*'}(t) = \lim_{t \to t_a^+} -\beta S^* w^*(1 + \lambda_I^* - \lambda_S^*) + \lambda_I^* \delta \tag{3.52}$$

Therefore $\lim_{t \to t_a^+} \lambda_I^{*'}(t)$ is strictly negative, and hence it contradicts the above property.

The above two cases show that t_a does not exist. Therefore $\lambda_I^*(t) > 0$ for $0 < t < T_{final}$. □

Lemma 3.7.2. *The function $\psi^*(t)$ is a positive concave function in time.*

Proof. The first derivative of $\psi^{*'}$ is as follows:

$$\psi^{*'} = \ (\lambda_I^{*'} - \lambda_S^{*'})\beta S^* I^* + (1 + \lambda_I^* - \lambda_S^*)\beta S^{*'} I + (1 + \lambda_I^* - \lambda_S^*)\beta S^* I^{*'} \tag{3.53}$$

After rearrangement, it becomes as follows:

$$\psi^{*'} = -(1 - \lambda_S^*)\beta\delta S^* I^*.$$

Equation 3.40 can be rewritten as $\lambda_S^{*'}(t) = -(1+\lambda_I^*-\lambda_S^*)\beta I^*(t)w^*(t)$. Since $1+\lambda_I^*-\lambda_S^* > 0$ and $\lambda_S^*(T_{final}) = 0$, the co-state variable $\lambda_S^*(t)$ is a positive decreasing function for time $0 < t < T_{final}$. In addition, $\lambda_s^*(t)$ function may equal 1 since $1 + \lambda_I^* - \lambda_S^* > 0$ and $\lambda_I^* > 0$. Therefore, the function $-(1 - \lambda_S^*)$ is a decreasing function in time with positive values $(\lambda_S^* > 1)$, then zero $(\lambda_S^* = 1)$ then negative values $(\lambda_S^* < 1)$. Also, the term $\beta\delta S^* I^*$ is

positive, therefore, we concluded that $\psi^{*'}$ is a decreasing function in time from positive to negative values and hence, $\psi^*(t)$ is a positive concave function with an inflection point in time at $\lambda_S^* = 1$. □

Based on inequality 3.47, and the fact that $\psi^*(t)$ is a concave function in time, we state the following theory, which shows the optimal dynamic weights in homogeneous networks.

Theorem 3.7.3. *The optimal dynamic weight reduction during the spread of an SIR epidemic in homogeneous networks is as follows:*

$$y^*(t) = \begin{cases} 0 & if \quad \frac{\partial f}{\partial y}|_{y(t)=0} > \psi(t), \\ (\frac{\partial f}{\partial y})^{-1}(\psi^*(t)) & if \quad \frac{\partial f}{\partial y}|_{y(t)=0} \leq \psi(t). \end{cases} \tag{3.54}$$

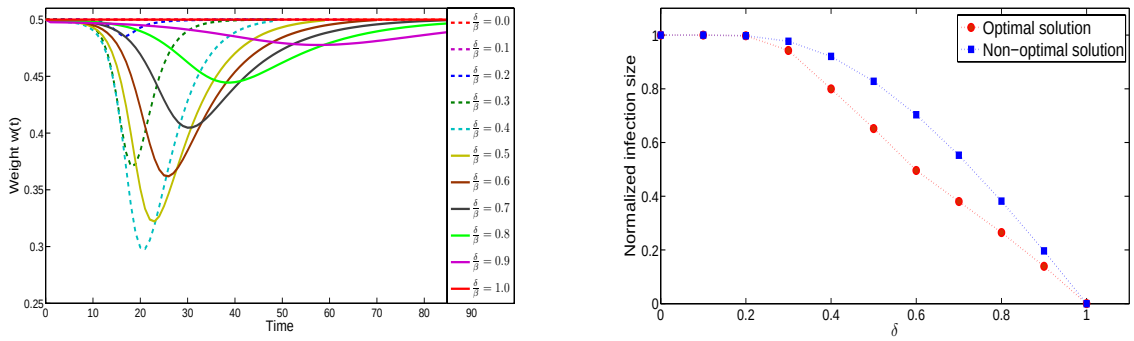
Proof. From Eq. (3.47), $\frac{\partial(f(y(t))-\psi(t)y(t))}{\partial y(t)}|_{y(t)=y^*(t)} = 0$ is evaluated to find the optimal $y^*(t)$. Therefore, we obtain $y^*(t)$ as follows:

$$y^*(t) = (\frac{\partial f}{\partial y})^{-1}(\psi^*(t)) \tag{3.55}$$

Since inequality (3.47) has to be preserved for all the time t, Eq. 3.55 is applied if and only if $\frac{\partial f}{\partial y}|_{y(t)=0} \leq \psi(t)$ is true. Hence, Eq. 3.54 is obtained. Consequently, optimal dynamical weight $w^*(t) = w(0) - y^*(t)$ is obtained as well.

□

Figure 3.8(a) shows the evaluation of the dynamical weights in time for a homogeneous network in which every individual is in contact with two other individuals composing a regular weighted graph. The simulation settings are as follows. 1) initial weight $w(0)$ equals 0.5, 2) infection rate β equals 1, and 3) the weight reduction cost function is convex with the form $f(z) = z^2$. Every curve in the figure represents the evaluation of the optimal control problem for a given cure rate δ value ($0 \leq \delta \leq 1$). The Figure shows that the amount of weight reduction increases ($w(0) - w(t)$ increases) for $0 \leq \delta \leq 0.4$, and then the weight reduction decreases ($w(0) - w(t)$ decreases) for $0.5 \leq \delta \leq 1$. In addition, Figure 3.8(b)



(a) Optimal dynamical weight

(b) Optimal normalized infection size

Figure 3.8: *Numerical evaluation of optimal dynamical weight and optimal normalized infection size for a regular network given different values of cure rate* $0 \leq \delta \leq 1$ *and constant infection rate* $\beta = 1$.

shows the numerical evaluations for the total fraction of infection cases given different cure rate values $0 \leq \delta \leq 1$, and a constant infection rate $\beta = 1$ for both the non-optimal and optimal solution. The optimal solution has lower fraction of infection than non-optimal solution. We also notice that for small and large values of cure rate, the optimal solutions coincide with the corresponding non-optimal solutions. To clarify, there are two factors leading to reduce the total infection size, which are the cure rate and the weight reduction. Meanwhile, there is a trade off between their roles in reducing the infection size. Large cure rate leads to reduce number of infections, and hence the weight reduction becomes less effective; However, very small cure rate leads to large number of infection, because the infection process is stronger than the cure process. Hence, the weight reduction becomes less effective. For intermediate values of cure rates, we find that the weight reduction is very effective. Both the weight reduction and the cure process, together, minimize the total infection size. For a given cure rate in Figure 3.8(a), the dynamical behavior of the weight is described as follows. During the early phase of the epidemic spreading in the population, few individuals are infected, and hence the weight reduction is not effective; However, when the infection size increases, the weight reduction becomes more effective. When the infection process becomes less dominant, the weight reduction decreases due to the exponential decay

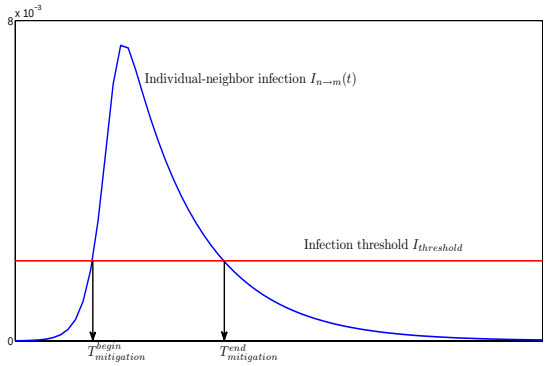


Figure 3.9: *The mitigation strategy is effective during the time period* $[T^{begin}_{mitigation}\ T^{end}_{mitigation}]$ *in which* $I_{threshold} < I_{n\to m}(t)$.

of infection size, which is a property of the SIR model. Such behavior of the dynamical weight is expect due to the concavity nature of $\psi(t)$ in Eq. 3.57.

3.8 Computational heuristics for the optimal mitigation strategies

Since the optimal dynamical weight behavior is obtained through the centralized optimal control framework, and the obtained solution requires solving simultaneous nonlinear equations in a centralized way, it is beneficial to propose heuristics to find near optimal solutions in a decentralized way with reduced complexity. Therefore, we proposed two different heuristics that are based on the results found from the homogeneous networks. For each heuristic, we study the infection between every pair of individuals that have contact. We denote $I_{n\to m}(t)$ as the probability that individual m becomes infected because of being in contact with an infected neighbor n. Mathematically, $I_{n\to m}(t)$ is described as follows:

$$\frac{dI_{n\to m}(t)}{dt} = \beta S_m(t) w_{m,n}(t) I_n(t) - \delta I_{n\to m}(t) \tag{3.56}$$

The weight reduction mitigation strategy takes place in a time duration $[T^{begin}_{mitigation}\ T^{end}_{mitigation}]$ during which $I_{n\to m}(t)$ exceeds a predefined infection threshold $I_{threshold}$ as shown in Fig-

ure 3.9. Analytically, the weight dynamic is described as follows:

$$w_{m,n}(t) = \begin{cases} \alpha w_{m,n}(0) \leq w_{m,n}(t) < w_{m,n}(0) & \text{if} \quad I_{n\to m}(t) > I_{threshold}, \\ w_{m,n}(0) & \text{if} \quad I_{n\to m}(t) \leq I_{threshold}. \end{cases} \tag{3.57}$$

The weight function in time $w_{m,n}(t)$ is either a Bang-Bang function that changes between two weight levels ($\alpha w_{m,n}(0)$ and $w_{m,n}(0)$), or a piecewise nonlinear function in time.

3.8.1 Bang-Bang controller heuristic

In this heuristic, once the infection probability of individual m due to the infection from neighbor n, $I_{n\to m}(t)$, exceeds the predefined infection threshold $I_{threshold}$, the initial weight $w_{m,n}(0)$ changes to the minimum social level $\alpha w_{m,n}(0)$. As long as $I_{n\to m}(t) \leq I_{threshold}$, the contact weight from individual m to individual n, $w_{m,n}(t)$ does not change. Therefore, weights are reduced to $\alpha w_{m,n}(0)$ during the time period
$[T^{begin}_{mitigation}\ T^{end}_{mitigation}]$ in which $I_{n\to m}(t) > I_{threshold}$. The Bang-Bang controller reflects the fact that once an individual receives the infection from one or more of its neighbors, the individual reduces its contact weight to the minimum social level. This heuristic also represents a type of homogeneous mitigation strategy in which the mitigation function (weight reduction profile) is the same among all the individuals; However, every pair of individuals decides about the start time and the time duration at which the contact weight is sharply reduced to its minimum level.

3.8.2 Piecewise nonlinear controller heuristic

During the time period $[T^{begin}_{mitigation}\ T^{end}_{mitigation}]$ in which $I_{n\to m}(t) > I_{threshold}$, the weights follow a proposed nonlinear decreasing function, which mainly depends on the infection rate β, the cure rate δ, and $\alpha w(0)$. The individual decreases the contact weight with its infected neighbor(s) according to the following rate of change controller differential equations $\frac{dw_{m,n}(t)}{dt}$:

* $\frac{dw_{m,n}(t)}{dt} = 0$ **if** $I_{threshold} \geq I_{n\to m}(t), t < T^{begin}_{mitigation}$

* $\frac{dw_{m,n}(t)}{dt} = -\beta^2(w_{m,n}(0)(1-\alpha))(e^{\beta(t-T^{begin}_{mitigation})})$ **if** $I_{threshold} < I_{n\to m}(t),\ T^{begin}_{mitigation} < t \leq T^{end}_{mitigation}$

* $\frac{dw_{m,n}(t)}{dt} = 0$ **if** $I_{threshold} < I_{n\to m}(t),\ T^{begin}_{mitigation} < t \leq T^{end}_{mitigation}$

* $\frac{dw_{m,n}(t)}{dt} = (w_{m,n}(0) - w_{m,n}(t))(1 - e^{-\delta(t-T^{end}_{mitigation})})$ **if** $I_{threshold} \geq I_{n\to m}(t), t > T^{end}_{mitigation}$

The motivation of proposing this specific nonlinear weight (controller) function is as follows:

- During the early phase of the spread of epidemic, the infection size is very small, and therefore, individuals do not change their contact levels $\frac{dw_{m,n}(t)}{dt} = 0 \Rightarrow w_{m,n}(t) = w_{m,n}(0)$.

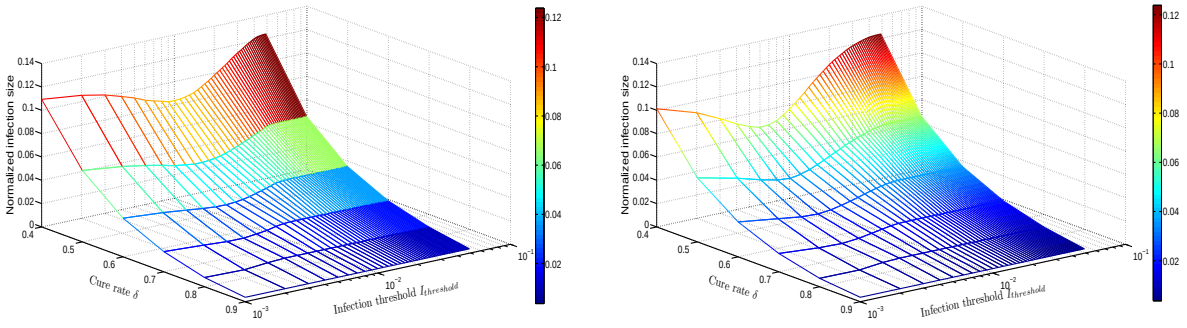
- Due to the presence of large susceptible population, the spread of the epidemic strengthens, and the infection size starts to increase. Consequently, every individual m decreases its contact level with each neighbor n according to the infection level $I_{n\to m}(t)$.

- If a neighbor n persistently is highly infected($I_{threshold} < I_{n\to m}(t)$) for longer time, and the contact weight $w_{m,n}(t)$ reaches its minimum level $\alpha w_{m,n}(0)$, the contact weight remains constant ($\frac{dw_{m,n}(t)}{dt} = 0 \Rightarrow w_{m,n}(t) = \alpha w_{m,n}(0)$) until the infection level $I_{n\to m}(t)$ becomes lower than the infection threshold.

- When the mitigation time ends ($t > T^{end}_{mitigation}$), individuals recover their contacts following a nonlinear increasing function, which is proportional to the cure rate δ until their initial contact levels are retrieved ($w_{m,n}(0)$).

3.8.3 Numerical evaluations

We perform extensive simulations on two types of networks to evaluate the proposed heuristics. The first type is the preferential attachment networks [10], and the second type is a survey-based contact social network [99]. In addition, we briefly discuss the differences between our approaches and social distancing mitigation methods.



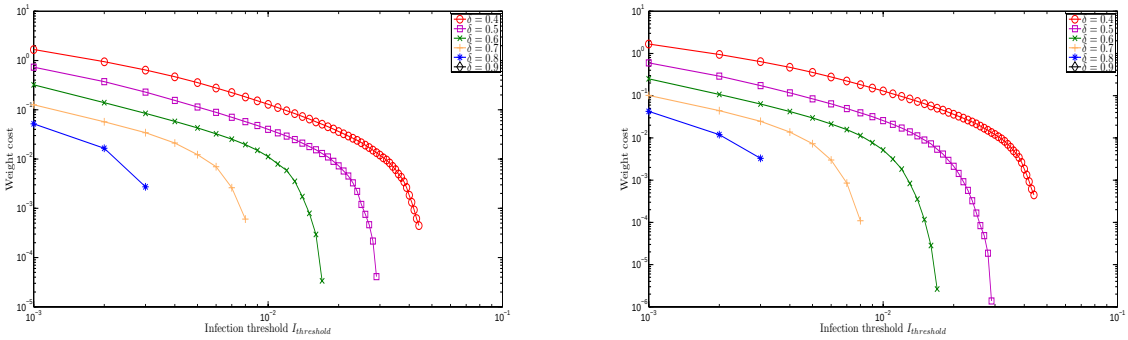
(a) Bang-Bang controller

(b) Piecewise nonlinear controller

Figure 3.10: *Numerical evaluation of the normalized number of infection size for both the Bang-Bang controller heuristic and the piecewise nonlinear controller heuristic given preferential attachment networks, different cure rate* $0.4 \leq \delta \leq 0.9$ *and different infection threshold* $I_{threshold}$*. The numerical evaluation is averaged over 20 runs. The best* $I_{threshold}$ *is the value at which the normalized infection size is minimal.*

Preferential attachment networks

The proposed heuristics are applied on preferential attachment networks. Different cure rates δ, and infection threshold $I_{threshold}$ are used with every heuristic, while the infection rate β is set to be equal to the reciprocal of the epidemic threshold $\frac{1}{\lambda_1}$ [112], where λ_1 is the maximum eigenvalue of the original weighted contact network. Figures 3.10(a) and 3.10(b) show the normalized infection size for different infection thresholds $I_{threshold}$ and cure rates $0.4 \leq \delta \leq 0.9$ for Bang-Bang and piecewise nonlinear controller heuristics, respectively. All results are averaged over 20 different preferential attachment networks, and each network has 10^4 individuals with initial contact weight value equals $w_{m,n}(0) = 0.5$ and minimum contact weight equals $\alpha w_{m,n}(0) = 0.05$. For Bang-Bang controller, as shown in Figure 3.10(a), for a given δ and small $I_{threshold}$, the infection size decreases until it reaches a minimum value, and it increases for higher values of $I_{threshold}$ until it reaches its highest value at which the mitigation strategy is no longer effective. Therefore, small values as well as high values of $I_{threshold}$ are not effective to reduce infection cases. In addition, as shown in Figure 3.11(a), small values as well as high values of $I_{threshold}$ incur extreme high and low weight reduction



(a) Bang-Bang controller

(b) Piecewise nonlinear controller

Figure 3.11: *Numerical evaluation of the weight reduction cost for both the Bang-Bang controller heuristic and the piecewise nonlinear controller heuristic given preferential attachment networks, different cure rate* $0.4 \leq \delta \leq 0.9$ *and different infection threshold* $I_{threshold}$. *The numerical evaluation is averaged over 20 runs.*

cost, respectively. Hence, for a given δ, there is a best $I^{best}_{threshold}$ value at which the infection size has the lowest value as shown in Figure 3.10(a) with a moderate weight reduction cost as shown in Figure 3.11(a). The above observations are analogous with the finding by Bondes *et al.* [18] that the optimal immunization investment is maximized for intermediate values of infection probabilities, while the immunization investment is less effective for low and high infection probabilities values. The same observations are obtained in the piecewise nonlinear controller heuristic as shown in Figures 3.10(b) and 3.11(b).

Below, we summarize the numerical evaluation obtained from both heuristics as follows:

- As the cure rate δ increases, the best infection threshold $I^{best}_{threshold}$ decreases. To clarify, high cure rate represents short infection time, and hence the probability that a susceptible individual receives the infection from an infected neighbor decreases. Therefore it requires smaller values of $I_{threshold}$ for the mitigation strategy to be effective.

- The normalized infection size for the piecewise nonlinear controller heuristic is smaller than the normalized infection size for Bang-Bang controller heuristic at every cure rate. Furthermore, the incurred weight reduction cost from the piecewise nonlinear

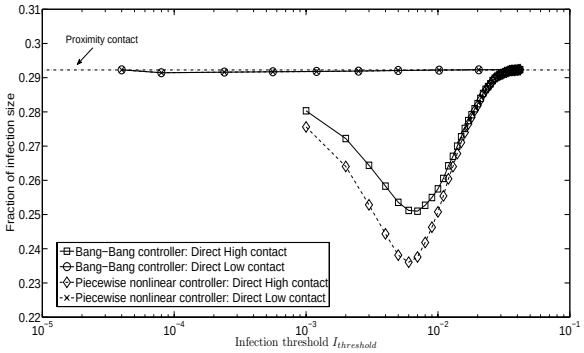


Figure 3.12: *Numerical evaluation of the Bang-Bang and the piecewise nonlinear heuristics for the three types of contacts in the survey-based contact network.*

controller heuristic is lower than the incurred cost from the Bang-Bang controller heuristic as shown in Figures 3.11(a) and 3.11(b). Therefore, the piecewise nonlinear controller heuristic outperforms the well-known Bang-Bang controller in both the infection size and the weight cost. However, every heuristic has different applications. For example, in Bang-Bang controller, individuals simply change their activities to the minimum level towards other individuals who are suspected to hold the infection. In the piecewise nonlinear controller, individuals prefer to maintain their contacts to the levels that not only allow them to interact with other individuals in the community, but also become cautious about receiving the infection from their infected neighbors.

Survey-based weighted social network

Bang-Bang and piecewise nonlinear heuristics are applied to a survey-based social network presented in Figure 2.6.6 to study the effect of the mitigation strategies on the total infection size.

We applied Bang-Bang and piecewise nonlinear heuristics on every type of contact between individuals i and j. For every heuristic applied to a certain type of contact, say $w_{x,i,j}$, the minimum contact level is $\alpha w_{x,i,j}$. The total contact rate between individuals i and j is as follows:

$$\frac{w_{proximity,i,j} + w_{direct-low,i,j} + w_{direct-high,i,j}}{3}.$$

For example, when Bang-Bang heuristic is applied to proximity contact, only the proximity contact rate is reduced. Similarly, the mitigation strategy is applied to the direct-low, and direct-high contacts. The same mitigation strategy process is used with the piecewise nonlinear controller. For different values of infection thresholds, Figure 3.12 shows that the mitigation of every type of contact has a different impact on the total size of infection. To clarify, when the mitigation strategy is applied on direct high contacts, the largest impact on the total infection size is obtained; however, the mitigation strategy has no impact on the proximity contact. In addition, the piecewise nonlinear heuristic is more effective on the direct high contact than the Bang-Bang heuristic.

Social distancing

We discuss the differences between the proposed heuristic approaches and social distancing mitigation strategy. Consider that the normal contact weights among the individuals have the same value, say $w(0)$. In social distancing, the contact weight is static and it is reduced to its lowest level αw_0 all the time. In this case, the epidemic threshold increases from $\frac{1}{w(0)\lambda_1}$ to $\frac{1}{\alpha w(0)\lambda_1}$, which in turn reduces number of infection cases for any effective infection rate $\frac{\beta}{\delta}$ that is greater than the epidemic threshold. Therefore, static social distancing mitigation strategy reduces number of infection cases; However, it incurs the maximum weight reduction cost. On the other hand, in the proposed approaches, the contact weight is dynamically changed between its maximum and minimum values $w(0)$ and $\alpha w(0)$, respectively. Hence, the contact network does not incur the maximum weight reduction cost value, and meanwhile number of infection cases is reduced. Therefore, our approaches outperform the static social distancing mitigation strategy in balancing between minimization of number of infection cases and minimization of weight reduction cost.

3.9 Summary

We summarize our findings as follows:

- *The epidemic threshold is the reciprocal of the spectral radius of the network*

  The theoretical analysis shows the role of the spectral radius in the epidemic spreading in networks, and the simulations verify the accuracy of the theoretical epidemic threshold.

- *A condition exists for the occurrence of a peak value of the infection*

  The theoretical analysis provides a condition for the existence of a peak value for the infection incidence, which is related to the epidemic threshold.

- *The individual-based SIR approach well approximates the continuous time Markov chain model*

  The individual-based approach is accurate for all $\frac{\delta}{\beta} > \frac{1}{\tau}$ and $\frac{\delta}{\beta} \ll \frac{1}{\tau}$. For values of $\frac{\delta}{\beta} < \frac{1}{\tau}$, the individual-based approach provides an upper-bound for the infection incidence.

- *The accuracy of the individual-based approach increases as the number of individuals increases*

  For large networks, the Markov chain closely approximates any stochastic process given sufficiently large numbers of states (i.e. 3^N different states in our case).

- *The range of effective infection rates leading to an epidemic is larger for the individual-based approach than for the heterogeneous mean field approach*

  For $0.7\lambda_{max} \leq \frac{\delta}{\beta} \leq \lambda_{max}$, the individual-based approach agrees with Monte Carlo simulations showing the existence of infection incidence, while the heterogeneous mean field approach does not show any infection incidence.

- *The numerical evaluations show that the new measure VC_{SIR} distinguishes between the robustness of assortative and disassortative preferential attachment networks*

  According to the robustness measure VC_{SIR}, the assortative preferential attachment networks is more robust than the disassortative networks.

- *In homogeneous networks, the optimal contact weights are adaptive*

  To balance between the minimization of the weight reduction cost and number of infection cases, weights decrease nonlinearly from their normal values when the epidemic spreading process overwhelms the network, and then the normal contact weight values are retrieved when the curing process overcomes the infection process

- *The optimal control problem addresses the relationship between the effective cure rate and the optimal control function*

  For intermediate values of effective cure rate, the value of weights are reduced dramatically from their normal values.

- *The piecewise nonlinear controller outperforms the Bang-Bang controller*

  Number of infected cases and the incurred weight reduction cost obtained by the nonlinear heuristic are lower than those obtained by the Bang-Bang controller.

- *The most effective strategy is the mitigation of direct-high contact in social networks*

  Using nonlinear controller, the mitigation strategy of the direct-high contact is the most effective strategy comparing with the proximity and direct-low contacts.

Chapter 4

Robustness of power grids with respect to cascading failures

In this chapter, we address a main question, *how robust is the electric power grid to resist cascading failures?* A cascading failure takes place when a single or multiple faults happen in the grid, and the stress on the transmission lines increases. The stress on the transmission line is the number of transmission loading relief procedures (TLR) in which the loads on the faulty lines is shifted to other lines [80].

The North American Electric Reliability Corporation (NERC) [115] introduced *Adequate Level of Reliability* (ALR) in which six characteristics of a power system are defined to guarantee a certain level of reliability under any contingency situation. NERC also introduced *Transmission System Standards: Normal and Emergency Conditions* [116] in which four different categories of events that take place in a power system are defined. The four categories are as follows:

- Category A: No Contingencies
- Category B: Event resulting in the loss of a single element.
- Category C: Event(s) resulting in the loss of two or more (multiple) elements.
- Category D: Extreme event resulting in two or more (multiple) elements removed or Cascading out of service.

These categories are used to study the relationship between the likelihood and the severity of the events as shown in [63].

Any power grid is characterized by its topology and the power flows on it. The topology represents the connectivity of substations (generation and distributions), while the power flows represent the dynamics that are controlled by the electrical characteristics of the grid, and are delivered from the generators to the distribution substations. The electrical characteristics of the grid are the capacity and the inductances of the transmission lines, the voltage values, the difference between the voltage phase angles at the terminals of each transmission line, and the loads at the distribution substations. The cascading failure takes place in the grid because some transmission lines becomes overloaded. Hence, the robustness of the grid can not be assessed only through the topology of the grid. Therefore, the power flows in the grid have to be involved as well in assessing the robustness of power grids.

In this chapter, we introduce a robustness metric η for the electric grid against cascading failures. We consider the electric grid as a network, and the electric power represents the dynamic that flows in the network. A power grid is composed of three different functioning parts, namely, generation, transmission and distribution. The electric power flows from the generators to the distribution stations through the transmission lines. Any transmission line can experience different types of faults causing the circuit breakers that are connected at the transmission line terminals to trip the line and to shift the electric power flow from the faulty line to another neighbor line. If the neighbor line becomes overloaded, it reaches its thermal limits and becomes disconnected as well. Such a process causes a contingency situation, which may lead to cascading failure that spreads across the electric power grid. Due to the hazardous situation that an electric power grid can experience due to fault(s) in the transmission lines, we focus our studies on the transmission system network, which consists of nodes representing the power substations (buses) and links representing the transmission lines. The new metric quantifies the robustness of the electric power grid with respect to

cascading failures. It is mainly based on two essential measures, namely, the frequency of link survival, and the depth of a cascading failure. The frequency of link survival is the number of events that a given link does not fails during a cascading failure. Therefore, it is based on the frequency of link failure, which is computed through the condition that there exists a link failure in the grid and whether other links fail consequently. The second quantity is the depth of the cascading failure, which represents number of failure stages in a grid given the start location of the failure before the grid totally collapses. Therefore, a grid that experiences a cascading failure with larger cascading depth is more robust than a grid that experience a cascading failure with shorter depth causing the power grid to collapse quickly. Also, when a cascading failure with long depth takes place in a grid, it allows the mitigation strategies to protect the grid before a black out takes place.

Differently from the literature work, in which the power flow is not considered, the new robustness measure relies on the power flow model to compute the flows in the transmission lines. It depends on the transmission line parameters, the voltage angles and the power/load at every bus in the grid. The transmission line parameters are the inductance and the capacities. The voltage angle controls the flow direction of the electric power, and the power/load represents the generated/consumed power at each node.

We apply the new robustness measure on three real power grid topologies given the power/load at each bus and the transmission lines parameters. To extend our numerical simulations, we generate synthetic topologies that are similar to the three real grids. The numerical results show that the power grid with large cascading depth preserves a high level of total power during a cascading failure comparing to the power grids with smaller cascading depth.

In addition to the robustness measure, we propose an optimal islanding mitigation strategy to the power grids with respect to cascading failures. The islanding approach finds the structures of the islands such that the total load shedding is minimized. To evaluate the mitigation strategy, real power grids are used to find the islands, the amount of load shedding, and the power generation reduction.

This chapter is organized as follows: In Section 4.1, we summarize the literature review. In Section 4.2, we explain the DC power flow model in details, and we propose a new robustness measure. Next, the computational algorithm of the robustness metric and the numerical evaluations are addressed in Section 4.3. The optimal islanding mitigation strategy is discussed in Section 4.4. Finally the chapter is summarized in Section 4.5.

4.1 Related work

Many studies have studied the cascading failure models in the electric grid using the complex networks approach. Motter and Lai in [80] studied the cascade failures in the complex network by introducing a capacity with tolerance parameter for each node, and they assumed that every pair of nodes exchanges homogeneous flow that is routed along the shortest path connecting them. They simulated the cascading failures through random and targeted removal of nodes based on the topological characteristics and the load distribution. Their measure of robustness is the largest connected component of the network.

Lai et al. in [62] introduced an efficiency measure of the complex networks while analyzing cascading failure. The efficiency measure is inversely proportional to the shortest paths between pairs of nodes. They showed that scale-free networks are more vulnerable to attacks on short-range links than attacks on long-range links, where the range of a link connecting two nodes is the shortest path between them after removing that link. The vulnerability of the Italian GRTN power grid to the cascading failure was studied in [27] where the authors introduced an efficiency measure for each link. The link efficiency is initially homogeneous, and it is updated by the removal of a node and redistribution of the loads. The power flow between any pair of nodes follows the most efficient path separating them. They concluded that the Italian GRTN power grid is robust to many failures, while it is vulnerable to the removal of nodes with highest betweenness.

Albert et al. in [2] studied the robustness of North American power grid with respect to

failures leading to cascade. They presented a vulnerability measure that is proportional to the average of the normalized number of generators feeding every distribution substation. Their conclusions coincided with those in [27].

The vulnerability of the European power grid was studied in [97] and it included 24 countries. The authors found that for the removal of nodes, the reaction of power grid is similar to the reaction of scale-free networks. They concluded that the fragility of the power grid increases with the growth of the network size.

All the above methods are generic and simplified. The dynamic quantity that is carried on the complex network is routed depending on its nature and its applications. For example, the dynamic quantity in the power grid is the electric current, which flows on a link with voltage difference greater than zero. In addition, the inductances of the transmission lines govern the amount and the direction of power that flow from the power generation nodes to the distribution substations. Therefore, these methods can not reflect the robustness of the power grids since they do not consider the amount of power that flow in the grid as well as the electrical characteristics of the grid. The authors of [17] addressed the obstacles for applying complex network theory to assess the vulnerability of power systems. The authors defined a network performance concept to overcome these obstacles.

Different mitigation strategies for the power grids were addressed in the literature. The work in [35] addressed a mixed integer programming approach to construct islands in the power grid. The approach mainly depends on integer optimizationin which the islands topologies are found giving number of islands, the transmission line capacities and inductances. Our optimal islanding mitigation strategy is different from the one in [35] in that the size of the islands, the load shedding for all nodes in each island, and the load shedding for all nodes in the topological complement of each island are minimized.

4.2 Problem definition

We consider the power grid as a network given that the generation and transmission substations are the nodes, while the transmission lines are the links, and we denote number of nodes as N and number of links as L. The DC Power Flow model, a simplification or linearization of the full AC power flow model, is used for the network analysis of the power grid in this work. The work [100] suggests that the term *DC Power Flow* comes from an old DC network analyzer [48], [19], in which the network branch was represented by a resistance proportional to its series reactance and each DC current was proportional to a Mega Watt (MW) flow. However, in the digital era this model became a set of simple, real (non-complex), nodal admittance matrix equations in terms of bus voltage angles and MW injections. In the AC model, the relation between real power, complex voltages and line impedance is expressed through the following equation describing the amount of real power that flows through a transmission line [49]:

$$P_{ij} = \frac{|V_i||V_j|}{z_{ij}} sin(\delta_{ij}) \tag{4.1}$$

where V_i and V_j are the voltages at nodes i and j, δ_{ij} is the phase angle between these voltages and z_{ij} is the line impedance. The above equation is modified to make it suitable for the linearized analysis by making the following assumptions:

- Voltage angle differences are small, i.e. $sin(\delta_{ij}) \approx \delta_{ij}$.
- Flat Voltage profile: All voltages are considered 1 $p.u.$
- Line resistance is negligible i.e. $R << X$.

Thus, the above equation can now be written as:

$$P_{ij} = \frac{\delta_{ij}}{x_{ij}} \tag{4.2}$$

In terms of matrices, P is an $N \times N$ matrix of power flows between each node i and j in the network, δ is the $N \times 1$ vector of phase angles and X is the $N \times N$ weighted adjacency matrix, and each matrix element represents the reactance of a transmission line. It is a real number if a line is present between two nodes, and 0 otherwise.

In matrix form,

$$[P] = [b][\delta] \tag{4.3}$$

The matrix $[b]$ represents the imaginary part of the Y_{bus} matrix of the power grid, where $b_{ij} = -\frac{1}{x_{ij}}$ and $b_{ii} = \sum_{i \in N} -b_{ij}$ for $i \neq j$. We usually assume that there is a reference node with 0 voltage angle.

The power handled by each node is the net sum of all the ingoing and outgoing power flows at that node as follows:

$$P_i = \sum_{j=1}^{N} P_{ij} = \sum_{j=1}^{N} (-b_{ij}\delta_{ij}) \tag{4.4}$$

The phase angles are computed given the total load at each bus using the following equation:

$$[\delta] = [b]^{-1}[P] \tag{4.5}$$

We assume that the power flow is below the capacity of the transmission lines [64] in normal operation. Whenever a link is removed, the power flowing through that link is redistributed to neighbor links. These neighbor links then have to carry their own power as well as the additional power which was redistributed due to the failure of a link. This redistribution may overload some links and these links become disconnected from the grid. The power from these newly failed links is again redistributed to their neighbors and more failures may happen in the grid. The redistribution of the power flow ia called the transmission loading relief procedure (TLR). If a sequence of failures takes place in the grid, a cascade of overloading failures may be triggered. Every time the redistribution of power takes place,

all the required quantities are recomputed using the DC power flow model. This step is repeated until the cascade stops due to either fragmentation of the grid causing a black out or there is no more any overloaded transmission line and the grid operates at a different operating point. In this dissertation, we assume the presence of circuit breakers to trip the overloaded transmission lines.

Two main quantities are used to obtain a new robustness measure, which are the average rank of a link, and the link survival frequency. The two quantities are explained in details in Section 4.3. The product of the average rank of a link and its survival frequency represents the average depth that a link can survives before its removal due to a cascading failure that happens in the grid. The larger the average rank, the more robust the grid is to resist the cascading failure before the power grid totally collapses causing a black out. Therefore, we propose a new metric η to quantify the robustness of electric power grids with respect to the cascading failures. The new metric mainly depends on the frequency of link survival, and on the depth of the cascading failure in the grid as follows:

$$\eta = \frac{1}{L} \sum_{j=1}^{L} f(l_j \ survives) r_{l_j} \tag{4.6}$$

Based on the measure in Eq. 4.6, we can compare the robustness of different power grids numerically. In addition, we can investigate four possible cases as follows:

1. *The frequency of survival is high and the average rank is also high.*

   This is the best case. It means that a given link can survive in most of the cases when failure occurs in the grid. Whenever a failure takes place, the link failure happens at late stage of the cascade. This makes the product of the two terms higher and hence contributes to the robustness of the network.

2. *The frequency of survival is high but the average rank is low.*

   In this case the link is quite robust to failures but when it fails its failure happens at an

earlier stage of cascade. This causes the product of the two terms to have intermediate values. Intuitively, this case is also good for the robustness of the grid since these links can resist failure most of the times.

3. *The frequency of survival is low but the average rank is high.*
   This case will again give intermediate values for the product of the two terms. However, this case is worse than the previous case because it is an indication of the presence of a weak link in the network. Although this link fails at later stages, it is almost sure that this link will fail as a result of other failures in the network.

4. *The frequency of survival is low and the average rank is also low.*
   This is the worst case. This is a set of most vulnerable links in the network and these links always fail when there is a cascade in the grid.

4.3 Computation algorithm of the robustness metric

The frequency that link l_j survives a cascade is computed as follows: We initially assume that there is a single link failure, say link l_i, and consequently, other links will fail due to the transmission loading relief procedure (TLR). Next, we identify the failed links due to the initial removal of link l_i. We repeat this process for every link l_i $i \in L$, and we list all the links that fail due to the removal of every initial link l_i. Therefore, the frequency that a link l_j fails $p(l_j\ fails)$ is the total number of times that link l_j fails due to the initial removal of each link independently, divided by the overall number of failed links. Therefore, the frequency of link failure reveals the most critical links in the grid. A link with the highest failure frequency is the weakest link in the grid. On the other hand, the frequency of link survival $f(l_j\ survives)$ is equal to $1 - p(l_j\ fails)$.

To address the depth of cascading failure, we define the link rank $r_{l_j|l_i\ removed}$ for link l_j as the cascading stage at which link l_j fails when link l_i is initially removed ($l_i\ removed$)

causing the cascading failure to take place in the grid. For instance, a link, say l_j, with rank 0 ($r_{l_j|l_i\ removed}$=0) means it does not fail due to the initial removal of link l_i, while a link with rank 1 ($r_{l_j|l_i\ removed}$=1) means that the link fails in the first stage of the cascade due to the removal of link l_i and so on. The average rank of link l_j is the summation of all ranks of that link, divided by number of links L. Mathematically, the average rank of link l_j is $r_{l_j} = \frac{1}{L}\sum_{i=1}^{L} r_{l_j|l_i\ removed}$.

The robustness measure η basically depends on the probability of survival and the average link rank. Algorithm 1 introduces a computational heuristic for the robustness metric η.

Algorithm 1 Computational heuristic for the robustness measure η

```
for i = 1 to L do
  Remove link l_i
  while Cascade failure takes place do
    Recompute P, and δ
    for l_j to L do
      if P_{l_j} ≥ Capacity_{l_j} then
        Compute the rank r_{l_j|l_i removed}
        Remove link l_j
      end if
    end for
  end while
  Recreate the original power grid
end for
Compute the average rank r_{l_i}, and the survival frequency of each link f(l_i survives) ∀i ∈ L
```

Using algorithm 1, we can assess the robustness of a power grid.

4.3.1 Power grid networks

We use three different realistic power grids to evaluate the new robustness measure. For each realistic power grid, the load at each bus is given as well as the inductances of the transmission lines. The realistic grids are described as follows:

- IEEE 247 bus test system [117] with 355 links
- IEEE 118 bus test system [117] with 179 links

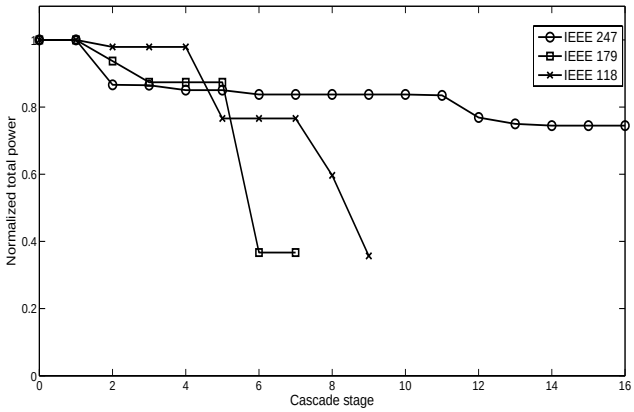


Figure 4.1: *Normalized total power at each cascading stage in case of the maximum cascading failures take place in the real three grids.*

- WSCC 179 bus equivalent system with 222 links

Since number of available power grid topologies are very limited, we use the network generator presented in [88] to generate synthetic power grids having the same number of nodes, the same number of links, and the same maximum node degree.

4.3.2 Numerical evaluations of the robustness metric

We apply the new metric η on the different power grids to measure their robustness with respect to the cascading failures. To extend the numerical simulations to larger number of grids, we generate four synthetic grids for each real power grid. We use the power flow simulator in [88], which is based on the DC power flow model, to evaluate the power flow on each link. Table 4.1 summarizes the robustness measure η of the four real grids and their synthetic ones. The table shows that the synthetic grids are generally more robust than their corresponding real grids. We also report the maximum number of cascading stages for both real and synthetic grids. We notice that the robustness measure is proportional to maximum number of cascade stages a power grid can experience. The real IEEE 247 is the most robust grid among the other real grids, while the real IEEE 118 grid is more robust than the real IEEE 179 grid.

Network	η	Maximum cascade stage
IEEE 247		
Real network	142.58	16
Synthetic network 1	160.03	21
Synthetic network 2	133.66	23
Synthetic network 3	138.71	26
Synthetic network 4	130.12	18
IEEE 179		
Real network	31.53	7
Synthetic network 1	114.71	15
Synthetic network 2	71.16	12
Synthetic network 3	127.31	17
Synthetic network 4	62.77	11
IEEE 118		
Real network	54.82	9
Synthetic network 1	75.42	11
Synthetic network 2	132.98	16
Synthetic network 3	138.63	17
Synthetic network 4	149.11	16

Table 4.1: *The robustness measure η and the maximum cascade stage for the real grids IEEE 247, IEEE 179 and IEEE 118 bus test systems, and four synthetic grids for each real grid*

To emphasize the role of the cascading depth in assessing the robustness of the power grids, we plot the normalized total power at each stage of the cascading failure. Figure 4.1 shows how the normalized total power decreases when the cascading failure takes place in the grid. We consider the case when the maximum number of cascading stages happens on each real grid. The figure shows that the IEEE 247 grid has the largest cascading depth, while it has the lowest power loss. On the opposite, the other two real grids has larger power losses and smaller cascading depth. Therefore, the IEEE 247 grid preserves a high level of total power during a cascading failure with larger depth before a total collapse takes place in the grid.

4.4 Optimal islanding mitigation strategy for the power grids

The main objective of the optimization problem is to find the optimal islands structure when a fault takes place at any part of the grid. In the presence of the fault, an island is created to isolate the spreading of the fault to the remaining part of the grid. Hence, the created island should be small in size, while the delivered power from the generators is maximized to satisfy the loads. In other words, the objective to minimize the load shedding at each load bus. Assume that number of islands is $n_{islands}$, and we denote k to be the island index such that $k = 1 \dots n_{islands}$. In addition, we assume that the power grid topology is G, the island topology is g_k, the group of links that interconnects the island with its topological complement is l_k, and the topological complement of the island is $T_k = G \backslash \{l_k \cup g_k\}$. Moreover, we denote s to be the index that distinguishes between the two partitions of the power grid i.e. $s = 1$ denotes the island topology (g_k) and $s = 2$ denotes the island topological complement (T_k) for island k. Figure 4.2 shows an example of three islands ($k = 1,\ 2,$ and 3) for a given power grid topology with 12 nodes, and 21 links. Every island k is denoted by $k, s = 1$, while the topological complement is denoted by $k, s = 2$. When an island is created, links are classified to three categories as shown in Figure 4.2: 1) links belong to the island (black line), 2) Links belong to the topological complement T of the island g (gray line), and 3) Links interconnect the island g with its topological complement T (gray dash line).

The objective function aims to minimize both the amount of generation/load change from normal operation, and the island size.

$$\begin{aligned} \text{Minimize} \quad & A \sum_{i=1}^{i=N} \sum_{k=1}^{k=n_{islands}} \mid power_i - d_i^k \mid \\ & + B \sum_{i=1}^{i=N} \sum_{j=1}^{j=N} \sum_{k=1}^{k=n_{islands}} \mu_{i,j}^{k,1} \end{aligned} \tag{4.7}$$

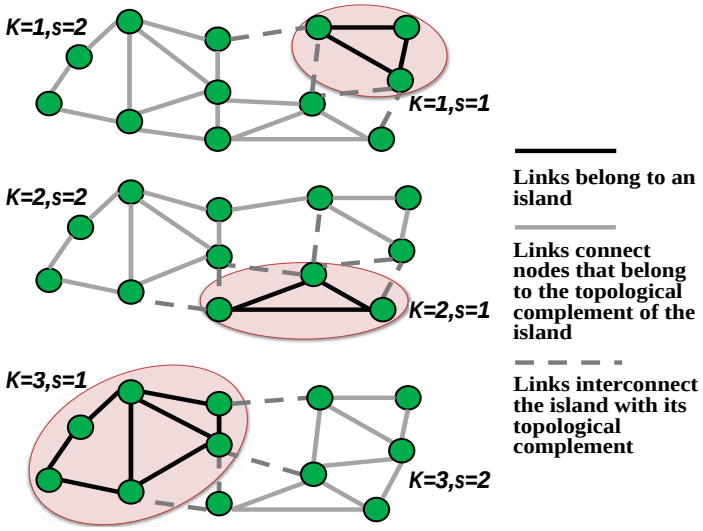


Figure 4.2: *The structure of three islands that belong to a power grid consisting of 12 nodes and 21 links. In the creation of island k, links are classified as follows: 1) links belong to island $k, s = 1$ (black line), 2) Links connect nodes that belong to the topological complement of the island $k, s = 2$ (grey line), and 3) Links interconnect island $k, s = 1$ with its topological complement $k, s = 2$ (gray dash line).*

Below, the islanding topology constraints are discussed in details.

$$\sum_{k=1}^{k=n_{islands}} \mu_{i,j}^{k,1} \leq a_{i,j} \quad \forall\ i,j = 1 \ldots N \tag{4.8}$$

$$\mu_{i,j}^{k,1} + \mu_{i,j}^{k,2} \leq 1 \quad \forall\ i,j = 1 \ldots N, k = 1 \ldots n_{islands} \tag{4.9}$$

$$\sum_{i=1}^{i=N} \sum_{j=1}^{j=N} \mu_{i,j}^{k,s} \geq 4 \quad \forall\ k = 1 \ldots n_{islands}, s = 1,2 \tag{4.10}$$

$$\sum_{j=1}^{j=N} (\mu_{i,j}^{k,1} + \mu_{i,j}^{k,2}) \geq 1 \quad \forall\ i = 1 \ldots N, k = 1 \ldots n_{islands} \tag{4.11}$$

$$\sum_{k=1}^{k=n_{islands}} (\mu_{i,j}^{k,1} + \mu_{i,j}^{k,2}) \leq a_{i,j} n_{islands} \quad \forall\ i,j = 1 \ldots N \tag{4.12}$$

$$\sum_{j=1}^{j=N} \sum_{k=1}^{k=n_{islands}} \mu_{i,j}^{k,1} \geq 1 \quad \forall\ i = 1 \ldots N \tag{4.13}$$

Table 4.2: *Definitions of data inputs and decision variables*

Data input	Definition
$a_{i,j}$	Binary adjacency matrix entry, $a_{i,j} = 1$ if there is a transmission line between substations i and j
$b_{i,j}$	Admittance matrix entry of the transmission lines
$power_i$	Normal power/load at substation i
α	Coefficient of minimum delivered load
$c_{i,j}$	Capacity of transmission line (i, j)
$n_{islands}$	Total number of islands
$n_{generators}$	Total number of generator nodes
n_{loads}	Total number of load nodes
Decision variables	**Definition**
$\mu_{i,j}^{k,s=1}$	Binary variable equals 1 if link (i, j) belongs to island k
$\mu_{i,j}^{k,s=2}$	Binary variable equals 1 if 1) link (i, j) belongs to the topological complement of island k, and 2) link (i, j) is not an interconnecting link between two islands
d_i^k	Load decision variable at node i, which belongs to island k or its topological complement. The variable d_i^k is negative if the node i represents a generator, while it has a positive value for load nodes.
δ_i^k	Voltage angel variable at node i,which belongs to island k or its topological complement
$f_{i,j}^k$	Network flow variable for link (i, j) that belongs to island k or its topological complement. It can have positive or negative value depending on the power flow direction.

$$\Big(\sum_{j=1}^{j=N}(\mu_{i,j}^{k,1}+\mu_{j,i}^{k,1}) \quad \geq 1\Big)$$
$$\Rightarrow \Big(\sum_{j=1}^{j=N}(\mu_{i,j}^{k,2}+\mu_{j,i}^{k,2}) \quad = 0\Big) \quad \forall\ i = 1\ldots N, k = 1\ldots n_{islands} \tag{4.14}$$

$$\Big(\sum_{j=1}^{j=N}(\mu_{i,j}^{k,2}+\mu_{j,i}^{k,2}) \quad \geq 1\Big)$$
$$\Rightarrow \Big(\sum_{j=1}^{j=N}(\mu_{i,j}^{k,1}+\mu_{j,i}^{k,1}) \quad = 0\Big) \quad \forall\ i = 1\ldots N, k = 1\ldots n_{islands} \tag{4.15}$$

$$\Big(\sum_{j=1}^{j=N} \mu_{i,j}^{k,1} \geq 1\Big)$$

$$\Rightarrow \Big(\sum_{j=1}^{j=N} \sum_{k'=1,k'\neq k}^{k'=n_{islands}} \mu_{i,j}^{k',1} = 0\Big) \ \forall\ i = 1 \ldots N, k = 1 \ldots n_{islands} \tag{4.16}$$

$$\mu_{i,j}^{k,s} = \mu_{j,i}^{k,s} \ \ \forall\ i,j = 1 \ldots N, k = 1 \ldots n_{islands}, s = 1,2 \tag{4.17}$$

$$\mu_{i,i}^{k,s} = 0 \ \ \forall i = 1 \ldots N, k = 1 \ldots n_{islands}, s = 1,2 \tag{4.18}$$

The group of inequalities in 4.8 - 4.18 describes the topological constraints in creating the islands. Inequality 4.8 imposes that a transmission line can belong to one island. In inequality 4.9, a transmission line at most can belong to either the island $\mu_{i,j}^{k,s=1} = 1$ or the complement topology of the island $\mu_{i,j}^{k,s=2} = 1$. In addition, the transmission line can interconnect two islands, and it does not carry power flow, and hence, both $\mu_{i,j}^{k,s=1}$ and $\mu_{i,j}^{k,s=2}$ equal 0 since the transmission line belongs to neither the island $k, s = 1$ nor the complement of the island $k, s = 2$. Inequality 4.10 imposes the number of transmission lines in every island $k, s = 1$ and its topological complement $k, s = 2$ to be at least 2. Notice that $\mu_{i,j}^{k,s} = \mu_{j,i}^{k,s}$ and therefore the right-hand side of inequality 4.10 is 4. Inequality 4.11 ensures that every node i in the power grid is assigned to either an island $k, s = 1$, or to the topological complement $k, s = 2$. In addition, in inequality 4.12, every link (i, j) at most can belong to an island $k, s = 1$ or the topological complement $k, s = 2$, and it can belong to any combination of them at most $n_{islands}$ times. Inequality 4.13 guaranties that every node i is assigned to an island $k, s = 1$. Inequality constraints 4.14, 4.15 and 4.16 are formulated using the logical constraint programming, which is composed of two constraints, the conditional constraint, and the actual constraint. If the conditional constraint is true, the actual constraint is applied to the problem. Therefore, the conditional constraint implies the actual constraint. As shown in logical constraint 4.14, if node i belongs island $k, s = 1$,

it can not belong to the topological complement $k, s = 2$. In addition, if node i belongs to the topological complement $k, s = 2$, it can not belong to the island $k, s = 1$ as shown in 4.15. Logical constraint 4.16 ensures that if node i belongs to a certain island $k, s = 1$, then it can not belong to any other island $k' \neq k$. Inequality 4.17 is a link symmetricity constraint, and inequality 4.18 avoids self connecting link.

4.4.1 Power flow model constraints

$$\delta^k_{generator(1)} = 0 \quad \forall k = 1 \dots n_{islands} \tag{4.19}$$

$$\begin{aligned} &\left(\mu^{k,1}_{i,j} + \mu^{k,2}_{i,j} \quad = 1\right) \\ \Rightarrow &\left(\mid b_{i,j}(\delta^k_i - \delta^k_j) \mid \quad \leq c_{i,j}\right) \quad \forall\ i, j = 1 \dots N, k = 1 \dots n_{islands} \end{aligned} \tag{4.20}$$

$$d^k_{generator(g)} \leq 0 \quad \forall\ g = 1 \dots n_{generators}, k = 1 \dots n_{islands} \tag{4.21}$$

$$d^k_{generator(g)} \geq power_{generator(g)} \quad \forall\ g = 1 \dots n_{generators}, k = 1 \dots n_{islands} \tag{4.22}$$

$$d^k_{load(l)} \leq power_{load(l)} \quad \forall\ l = 1 \dots n_{loads}, k = 1 \dots n_{islands} \tag{4.23}$$

$$d^k_{load(l)} \geq \alpha power_{load(l)} \quad \forall\ l = 1 \dots n_{loads}, k = 1 \dots n_{islands} \tag{4.24}$$

The group of constraints 4.19 - 4.24 represents the DC power flow model equations, the upper bounds, and the lower bounds of the generated power and loads. For DC power flow model, we assume that bus 1 is the slack bus with voltage angle equals 0 regardless the island scheme k as shown in equation 4.19. In logical constraint 4.20, if the link (i, j) belongs to either the island topology $k, s = 1$ or the complement topology $k, s = 2$, the DC power flow

model is applied such that the amount of power flow on link (i, j) does not exceed the link capacity. Inequalities 4.21 and 4.22 represent the bounds on the amount of generated power from each generator $generator(g)$ for each island scheme k, while inequalities 4.23 and 4.24 represent the bounds on the delivered loads. We impose the lower bound on the delivered power at each node to be a fraction α of the total load.

4.4.2 Network flow model constraints

$$\begin{aligned} &\left(\mu_{i,j}^{k,1} + \mu_{i,j}^{k,2} \quad = 1\right) \\ \Rightarrow &\left(b_{i,j}(\delta_i^k - \delta_j^k) \quad = f_{i,j}^k\right) \quad \forall\ i, j = 1 \ldots N, k = 1 \ldots n_{islands} \end{aligned} \tag{4.25}$$

$$\begin{aligned} &\left(\mu_{i,j}^{k,1} + \mu_{i,j}^{k,2} \quad = 1\right) \\ &\Rightarrow \left(\mid f_{i,j}^k \mid \quad \leq c_{i,j}\right) \quad \forall\ i, j = 1 \ldots N, k = 1 \ldots n_{islands} \end{aligned} \tag{4.26}$$

$$\sum_{i=1}^{i=N} f_{i,j}^k = d_j^k \quad \forall j = 1 \ldots N, k = 1 \ldots n_{islands} \tag{4.27}$$

The network flow model plays an important role in the islanding problem formulation. It separates the total load computation of the island from its topological complement. In addition, it ensures that links, which interconnect any island with the rest of the topology (topological complement), do not affect the load computation of nodes belonging to either the island or the topological complement. Therefore, we consider the network flow model in the islanding problem as shown in inequalities 4.25 -4.27. Logical constraint 4.25 represents the condition that if the link (i, j) is either a part of island $k, s = 1$ or the topological complement $k, s = 2$, the amount of flow equals $f_{i,j}^k$, which is constrained by the link capacity as shown in the logical constraint 4.26. The power flow at node j is the sum of all the flows on its attached links as shown in 4.27.

4.4.3 Numerical evaluation for the optimal islanding

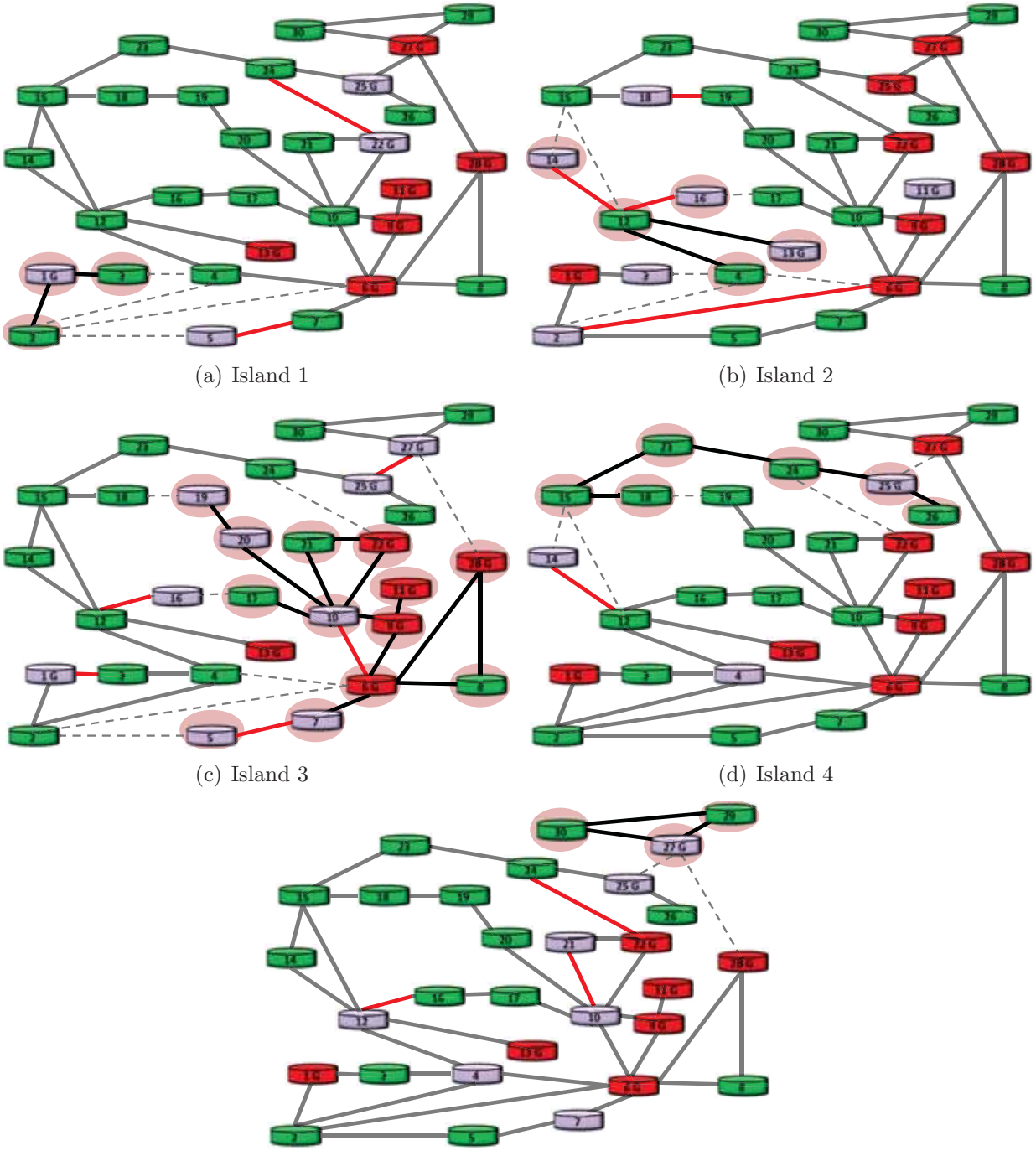
We apply the optimal islanding mitigation strategy on a power system with 30 nodes and 42 links. We assume that number of islands is 5. The topology of the islands and their topological complements are shown in 4.3. The dotted lines (- - -) are the links that interconnect every island with its topological complement. The black lines belong to an island, while the gray lines belong to the topological complement of an island. Nodes that are surrounded with elliptical shade belong to an island. The green color represents the load nodes that do not experience any load shedding. The purple color represents both the generators that generate power less than the normal operation and the loads that experience load shedding. The red color represents both the links that are operated at their capacity and the generators that generate the same power in the normal operation.

As shown in Figure 4.3, every island has at least one generator and two transmission lines to guarantee the island connectivity. In addition, Figure 4.3 shows that there are many loads do not experience load shedding (green load nodes), however, due to link capacities, few load nodes experience load shedding as shown in island 2 4.3(b) (red links and purple load nodes), and hence the power generation is reduced in some generators (purple generators). Another reason that a generator reduces its output power is that the total load in the island becomes less than its normal output generation as shown in island 1 4.3(a), island 4 4.3(d), and island 5 4.3(e).

4.5 Summary

The chapter is summarized as follows:

- *The new metric η measures the robustness of power grids respect to cascading failure* The new metric is a function of the survival frequency of the transmission lines, and the cascading depth when cascading events take place before a blackout happens in the grid.



(a) Island 1

(b) Island 2

(c) Island 3

(d) Island 4

(e) Island 5

Figure 4.3: *The optimal islanding solution for a power grid with 30 nodes and 42 links. The power grid has 9 generators and 31 load nodes.*

- *The role of the link survival frequency and the depth of the cascading failure is outlined*

  The power degradation curve shows the importance of the cascading failure depth. In addition, both the link survival frequency and the cascading depth represent four different cases, which can describe the state of the power grids due to a single link failure.

- *Optimal islanding strategy are proposed to mitigate the power grids from cascading failure events*

  The problem of finding the islands are formulated using the Mixed Integer Non Linear Programing MINLP. The solution finds the group of transmission lines that are disconnected to form the islands, the load shedding at the load nodes, and generation reduction at the generators.

Chapter 5

Conclusions and future research work

In this Chapter, we summarize our conclusions for robustness of complex networks, and we present some directions for future work.

5.1 Conclusions

Different metrics were introduced to quantify the robustness of complex networks with respect to disturbing dynamics. Firstly, we introduced Viral Conductance VC to quantify the robustness of complex networks with respect to the spread of susceptible/infected/susceptible SIS and susceptible/infected/recovered SIR epidemics. The new metric mainly integrates the effective cure rate with the corresponding infected population. For instance, the infected population due to the prevalence of SIS epidemics in networks represents the persistent fraction of infected nodes, while in the spreading of SIR epidemics, the infected population represents the accumulation of the new infected cases until the epidemic prevalence dies out. The new metrics were applied to different types of synthetic and real networks revealing their robustness dissimilarity due to the distinct topological structures. In addition, viral conductance was applied to different networks to measure the efficiency of the mitigation strategies.

Secondly, we proposed a new individual-based approach to model the spread of SIR epidemics in networks. The new model outperforms the well-known heterogeneous mean field approach in modeling the spread of epidemics with effective infection rates that are close to the epidemic threshold. Moreover, the new approach reveals the effect of the network structure on the spreading process through the eigenvalues and their corresponding eigenvectors of the networks.

Thirdly, an optimal mitigation strategy was proposed to reduce the total infected population in social contact networks due to the spread of SIR epidemics. The new mitigation strategy is based on the optimal control theory in which, the contact weights are optimally reduced from their normal rates to slow/reduce the spread of infection. Therefore, the optimal control problem aims to minimize both the total infected cases and the reduction in the contact weights. We analytically found the optimal adaptive weighted contact networks for homogeneous networks given different effective cure rates. Since the optimal solution was found based on centralized methods, we introduced two heuristics to find the near optimal solution in a decentralized way. The two heuristics are based on Bang-Bang controller and a piecewise nonlinear controller, respectively. In Bang-Bang heuristic, weights can switch between the normal level and the minimum level, while in the piecewise nonlinear controller, weights are reduced nonlinearly from their normal values. The results showed that the piecewise nonlinear controller resulted in lower infected cases and less reduction in the contact weights than the Bang-Bang controller.

Finally, to extend the robustness measure concept, a new metric was proposed to measure the robustness of power grids with respect to cascading failures. The new metric basically depends on two quantities, which are the survival probability and the link rank. The survival probability represents the probability that a link does not fail during a cascading process, while the link rank represents the average stage at which a link failure takes place in the grid. The new metric was applied to different

real and synthetic power grids to measure their robustness with respect to cascading failures caused by an initial link fault. Moreover, a new strategy is proposed to mitigate the power grid from cascading failure events. The new strategy depends on the grid islanding in which, the optimal island topologies are found such that the maximum electric loads are satisfied in the island and in the remaining part of the grid.

5.2 Future work

The presented work in this dissertation can be extended to different research directions. Below, we summarize the possible future work:

- The viral conductance is mainly based on probabilistic epidemic models, which deterministically evaluate the infection size in time. However, the actual mechanism of epidemic outbreaks differ from the probabilistic models. Therefore, an extensive evaluation for the viral conductance based on actual disease outbreaks in real networks becomes very vital.
- Viral conductance can be analytically expressed using the mathematical series aiming to provide a fast computational heuristic that is featured with high accuracy and less complexity. Van Mieghem [75] has proposed a Laurent series to expand the mathematical expression of viral conductance in the case of SIS epidemics, however, numerical evaluations are still required for validations.
- In the optimal mitigation strategies, the optimal control theory can be applied to minimize the total infection sizes by controlling the total contact weight of an individual. The total contact weight is the sum of weights of an individual with the neighbors. Instead of studying the individual-neighbor contact weight, the overall weight of an individual will be the main controller. Therefore, this approach will reduce the problem complexity.

- There are many challenges related to measuring the robustness of dynamic networks. The work in [94] addresses the spread of epidemics on dynamic topologies, and the authors found that the epidemic threshold is the product of individual epidemic thresholds for the different topologies in discrete time domain. However, the dynamic of topologies can be modeled in continuous time. The relative change between the network and the dynamics plays a major role in quantifying the robustness of networks. On one hand, if the epidemic spreading process is faster than the dynamic of the topology over a long time duration, the topology is considered static. On the other hand, if the changes in both the topology and the epidemic spreading have the same pace, there are some challenges concerning the change in the epidemic threshold, and hence, the range of effective infection rate will change accordingly.
- Robustness of power grids with respect to cascading failures is a very challenging research topic. Analytical and simulative approaches are needed to address the behavior of the power grids when rare events take place causing blackouts. Therefore, stochastic as well as deterministic models can be derived to describe the cascading failures, and extensive numerical evaluations are also required for accuracy and validation purposes.

Bibliography

[1] M. Ajelli et al., Comparing large-scale computational approaches to epidemic modeling: agent-based versus structured metapopulation models, *BMC Infectious Diseases 10, 190*, 2010.

[2] R. Albert, I. Albert, and G. Nakarado, Structural vulnerability of the North America power grid, *arXiv, 0401084*, 2004.
SIAM Review 45, 167256, 2003.

[3] R. Albert, H. Jeong, and A. Barabási, Error and attack tolerance of complex networks, *Nature*, vol. 406, pp. 278–82, 2000.

[4] L. A. Amaral, A. Scala, M. Barthelemy and H. E. Stanley, Classes of small-world networks, *Proc Natl Acad Sci U S A*, 97, 11149–11152, 2000.

[5] R. Anderson, Population ecology of infectious disease agents, *Theoretical ecology*, 1981.

[6] R. Anderson, Population ecology of infectious disease agents, *Theoretical ecology*, 1981.

[7] R. Anderson and R. May, Infectious diseases of humans: Dynamics and control, *Oxford Press*, 1991.

[8] R. M. Anderson and R. M. May,Infectious diseases in humans, *Oxford University Press*, Oxford, 1992.

[9] P. Bajardi et al.,Modeling vaccination campaigns and the Fall/Winter 2009 activity of the new A(H1N1) influenza in the Northern Hemisphere, *Emerging Health Threats Journal, 2:e11 doi: 10.3134/ehtj.09.011*, 2009.

[10] A-L. Barabási and A. Réka, Emergence of scaling in random networks, *Science*, vol. 286, pp. 509–512, 1999.

[11] A. Barrat, M. Barthélemy and A. Vespignani, Dynamical processes on complex networks, *Cambridge University Press*, Cambridge, 2008.

[12] C. Barrett, K. Bisset, S. Eubank, X. Feng and M. Marathe, EpiSimdemics: an efficient and scalable framework for simulating the spread of infectious disease on large social networks, *In Proceedings of SuperComputing 08 International Conference for High Performance Computing, Networking Storage and Analysis. Austin, Texas, November 15-21*, 2008.

[13] C. Barrett, K. Bisset, J. Leidig, A. Marathe and M. Marathe, An integrated modeling environment to study the co-evolution of networks, individual behavior, and epidemics, *AI Magazine 31: 7587*, 2009.

[14] M. Barthélemy, A. Barrat, R. Pastor-Satorras, A. Vespignani, Dynamical patterns of epidemic outbreaks in complex heterogeneous networks, *J Theor Biol.*, 235, 275-88, 2005.

[15] A. J. Black, A. J. McKane, A. Nunes, and A. Parisi, Stochastic fluctuations in the susceptible-infective-recovered model with distributed infectious periods, *Phys. Rev. E*, 80, 021922, 2009.

[16] M. Boguñá, R. Pastor-Satorras and A. Vespignani, Epidemic spreading in complex networks with degree correlations, *Lect. Notes Phys*, 625, 127-147, 2003.

[17] E. Bompard, M. Masera, R. Napoli and F. Xue , Assessment of Structural Vulnerability for Power Grids by Network Performance Based on Complex Networks, *CRITICAL INFORMATION INFRASTRUCTURE SECURITY LNCS*, Vol. 5508, 144-154, 2009.

[18] M. Bondes, D. Keenan, A. Leidner and P. Rohani, Higher disease prevalence can induce greater sociality: A game theoretic coevolutionary model, *The society for the study of Evolution: International journal of organic evolution, Vol. 59, No. 9*, September 2005.

[19] J.A. Casazza, and W.S. Ku, The co-ordinated use of A-C and D-C network analyzers, *Proceedings of American Power Conference, Vol. 16*, 1954.

[20] D. Chakrabarti, J. Leskovec, C. Faloutsos, S. Madden, C. Guestrin, and M. Faloutsos, Information Survival Threshold in Sensor and P2P Networks, in *proceedings of IEEE INFOCOM.07*, 2007.

[21] Z. Chen and C. Ji, Spatial-Temporal Modeling of Malware Propagation in Networks, *IEEE TRANSACTIONS ON NEURAL NETWORKS*, vol. 16, pp. 1291–1303, 2005.

[22] KC. Claffy, T. E. Monk, and D. McRobb, Internet tomography, *Nature*, 1999, http://www.caida.org/tools/measurement/skitter/.

[23] F. Coffman, Z. Ge, V. Misra, and D. Towsley, Network resilience: exploring cascading failures within BGP, in *proceedings 40th Annual Allerton Conference on Communications, Computing and Control*, 2002.

[24] R. Cohen, K. Erez, D. ben-Avraham, and S. Havlin , Resilience of the Internet to Random Breakdowns, *Phys. Rev. Lett.*, 85, 46264628, 2000.

[25] V. Colizza et al., Estimate of Novel Influenza A/H1N1 cases in Mexico at the early stage of the pandemic with a spatially structured epidemic model, *PLoS Currents: Influenza*, Nov. 2009.

[26] V. Colizza, A. Barrat, M. Barthelemy and A. Vespignani, Epidemic predictability

in meta-population models with heterogeneous couplings: the impact of disease parameter values, *Int. J. Bifurcation and Chaos 17, 2491-2500*, 2007.

[27] P. Crucitti, V. Latora, and M. Marchiori, A topological analysis of the Italian electric power grid, *Elsevier, Physica A, 338*, 2004.

[28] D. Cvetkovic, M. Doob, and H. Sachs, *Spectra of graphs, Theory and Applications*, Johan Ambrosius Barth Verlag, Heidelberg, 3rd edition, 1995.

[29] S.N. Dorogovtsev, Networks with given correlations, *arXiv:cond-mat/0308336v1*, 2003.

[30] JC Doyle, DL Alderson, L. Li, S. Low, M. Roughan, S. Shalunov, R. Tanaka and W. Willinger The robust yet fragile nature of the Internet, *PNAS USA, 11;102(41):14497-502*, October 2005.

[31] R. Durrett, Some features of the spread of epidemics and information on a random graph, *PNAS*, vol. 107, 10, 2010.

[32] P. Erdös and A. Rényi, On Random Graphs, *Publicationes Mathematicae*, vol. 6, pp. 290–297, 1959.

[33] S. Eubank et al., Modelling disease outbreaks in realistic urban social networks, *Nature 429(6988): 180-184*, 2004.

[34] P. Eugster, R. Guerraoui, A. Kermarrec, , and L. Massoulié, From Epidemics to Distributed computing, *IEEE Computer*, vol. 37, pp. 60–67, 2004.

[35] N. Fan, D. Izraelevitz, F. Pan, P. M. Pardalos, and J. Wang, A mixed integer programming approach for optimal power grid intentional islanding, *Energy Syst*, Springer-Verlag, 3:77-93,2012.

[36] D. A. Fell and A. Wagner, The small world of metabolism, *Nature Biotechnology*, 18, 1121-1122, 2000.

[37] B. Frank, S. David, and T. Pieter, Threshold behaviour and final outcome of an epidemic on a random network with household structure, *Adv. in Appl. Probab.*, 41, 3, 765-796, 2009.

[38] A. Ganesh, L. Massoulié,
and D. Towsley, The Effect of Network Topology on the Spread of Epidemic, in *proceedings IEEE INFOCOM.05*, Miami, FL, 2005.

[39] D. Gao and S. Ruan, An SIS patch model with variable transmission coefficients, *Mathematical Biosciences, Volume 232, Issue 2, Pages 110-115*, August 2011.

[40] R. Gehring, P. Schumm, M. Youssef and C. Scoglio, A Network-based Approach for Resistance Transmission in Bacterial Populations, *J. Theor. Bio.*, 262, 97–106, 2010.

[41] T. Germann, K. Kadau, I. Longini and C. Macken, Mitigation strategies for pandemic influenza in the United States, *Proceedings of the National Academy of Sciences 103: 59355940*, 2006.

[42] M. Girvan and M. E. J. Newman, Community structure in social and biological networks, *Proc. Natl. Acad. Sci. U S A*, 99, 7821–7826, 2002.

[43] S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni and Y. Moreno, Discrete-time Markov chain approach to contact-based disease spreading in complex networks, *Europhys. Lett*, 89, 38009, 2010.

[44] T. Gross and B. Blasius, Adaptive coevolutionary networks: a review, *Journal of the Royal Society Interface, 5(7): e11569*, 2010.

[45] T. Gross, C. J. Dommar DLima, and B. Blasius, Epidemic Dynamics on an Adaptive Network, *Phys. Rev. Lett. 96, 208701*, 2006.

[46] B. Gungor, Power Systems, *Technology Publications*, 1988.

[47] Q. Guo, T. Zhou, J.-G. Liu, W.-J. Bai, B.-H. Wang, and M. Zhao, Growing scale-free small-world networks with tunableassortative coefficient, *Physica A*, vol. 371, pp.814–822, 2006.

[48] W.C. Hahn, Load studies on the D-C calculating table, *General Electric Review, Vol. 34*, 1931.

[49] D.V. Hertem, J. Verboomen, K. Purchala, R. Belmans, and W.L. Kling, Usefulness of DC Power Flow for Active Power Flow Analysis with Flow Controlling Devices, *In proceedings of 8th IEEE International Conference on AC and DC Power Transmission*, London, March 2006.

[50] A. Jamakovic, R.E. Kooij, P. Van Mieghem, and E. van Dam, "Robustness of networks against the spread of viruses: the role of the spectral radius," in *the 13th Annual Symposium of the IEEE/CVT Benelux*, Liége,Belgium, 2006.

[51] H. Jeong, S. Mason, A.-L. Barabási, and Z. N. Oltvai,Lethality and centrality in protein networks, *Nature*, 411, 41-42, 2001.

[52] H. Jeong, B. Tombor, R. Albert, Z. N. Oltvai, and A.-L. Barabási,The large-scale organization of metabolic networks, *Nature*, 407, 651-654, 2000.

[53] C. Jiang and M. Dong, Optimal Measures for SARS Epidemics Outbreaks, *IEEE Intelligent Control and Automation WCICA*, October 2006.

[54] E. Kenah, JM. Robins,Network-based analysis of stochastic SIR epidemic models with random and proportionate mixing, *J Theor Biol.*, 249, 4, 706-22, 2007.

[55] J.O. Kephart and S.R. White, " Direct-graph epidemiological models of computer viruses," *IEEE Computer Society Symposium on Research in Security and Privacy*, pp. 343–359, 1991.

[56] W. O. Kermack and A. G. McKendrick, A Contribution to the Mathematical Theory of Epidemics, *Proc. Roy. Soc. Lond.*, A 115, 700-721, 1927.

[57] W. O. Kermack and A. G. McKendrick, A Contribution to the Mathematical Theory of Epidemics, *Proc. Roy. Soc. Lond.*, A 115, 700-721, 1927.

[58] J.O. Kephart, G.B. Sorkin, and M. Swimmer, An immune system for cyberspace, in *Proceedings of Conference on Systems, Man, and Cybernetics IEEE International: Computational Cybernetics and Simulation*, 1997, vol. 1.

[59] M.H.R Khouzani, S. Sarkar, and E. Altman,Optimal Control of Epidemic Evolution, *Proceedings of IEEE INFOCOM 2011*, Shanghai, China, April 10-15, 2011.

[60] R. E. Kooij, P. Schumm, C. Scoglio, and M. Youssef, "A New Metric for Robustness with Respect to Virus Spread," *Proceedings of the 8th International IFIP-TC 6 Networking Conference*, L. Fratta et al. (Eds): NETWORKING 2009, LNCS 5550, pp. 562 -572, 2009.

[61] C. Lagorio et al., Quarantine Generated Phase Transition in Epidemic Spreading, *Phys. Rev. E 83, 026102*, 2011.

[62] Y.-C Lai, A. Motter, and T. Nishikawa, Attacks and cascades in complex networks, *Lct. Notes Phys., 650, 299-310*, 2004.

[63] S. Lee, Probabilistic Reliability Assessment for transmission planning and operation including cascading outages, *IEEE PES Power Systems Conference and Exposition*, 2009.

[64] E. Lerner, What's wrong with the electric grid?, *American Institute of Physics, The Industrial Physicist*, October/November 2003.

[65] L. Li, D. Alderson, W. Willinger, J. Doyle, R. Tanaka, and S. Low, A First

Principles Approach to Understanding the Internet's Router Technology, *Proc. Sigcomm, ACM*, 2004.

[66] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley and Y.øAberg, The web of human sexual contacts, *Nature*, 411, 907–908, 2001.

[67] G. Macdonald, The analysis of equilibrium in malaria, *Trop Dis. Bull.*, 49, 1952.

[68] P. Mahadevan, D. Krioukov, M. Fomenkov, B. Huffaker, X. Dimitropoulos, Kc claffy and A. Vahdat, Lessons from Three Views of the Internet Topology, *arXiv:cs/0508033v1*, August 2005.

[69] P. Mahadevan, D. Krioukov, B. Huffaker, X. Dimitropoulos, Kc claffy and A. Vahdat, Comparative analysis of the Internet AS-level topologies extracted from different data sources, Datapage,http://www.caida.org/analysis/topology/astopocomparisons/.

[70] M. Manzano, J. Segovia, E. Calle, P. Vilá and J. Marzo, Modeling spreading of failures in GMPLS-based Networks, *proceedings of SPECTS*, Ottawa, Canada, 2010.

[71] A. Marathe, B. Lewis, J. Chen and S. Eubank, Sensitivity of Household Transmission to Household Contact Structure and Size,*PLoS ONE 6(8): e22461*, 2011.

[72] V. Marceau et al., Adaptive networks: coevolution of disease and topology, *Phys. Rev. E 82, 036116*, 2010.

[73] R. May and R. Anderson, Population biology of infectious diseases, *Part II Nature*, 280, 1979.

[74] R. M. May and A. L. Lloyd, Infection dynamics on scale-free networks, *Phys. Rev. E*, 64, 066112, 2001.

[75] P. Van Mieghem, The viral conductance of a network, *To appear in Compter Communications, Elsevier*, 2012.

[76] P. Van Mieghem and J. Omic, In-homogeneous Virus Spread in Networks, *Delft University of Technology, report 2008081*, 2008.

[77] P. Van Mieghem, J. Omic, and R.E. Kooij, "Virus Spread in Networks," *IEEE/ACM Transactions on Networking*, vol. 17, pp. 1–14, 2009.

[78] J. M. Montoya and R. V. Solé, Small world patterns in food webs, *J. Theor. Bio.*, 214, 405-412, 2002.

[79] Y. Moreno, R. Pastor-Satorras and A. Vespignani, Epidemic outbreaks in complex heterogeneous networks, *Eur. Phys. J. B*, 26, 521-529, 2002.

[80] A. Motter, and Y. Lai, Cascade-based attackes on complex networks, *Physical Review E, 66.065102*, 2002.

[81] J. D. Murray, Mathematical Biology, *Springer Verlag*, Berlin, 1993.

[82] J. D. Murray, Mathematical biology: An introduction, *Springer*, 2002.

[83] M. E. J. Newman, Spread of epidemic disease on networks, *Phy. Rev. E*, 66, 16128, 2002.

[84] M. E. J. Newman, Assortative Mixing in Networks, *Physical Review Letter*, vol. 89,208701, 2002.

[85] M. E. J. Newman, The structure and function of complex networks, *SIAM Review 45, 167256*, 2003.

[86] M. Newman, A. Barabási, and D. Watts, *The Structure and Dynamics of Networks*, Princeton University Press,Princeton, NJ 2006.

[87] V. Nikiforov, Eigenvalues and degree deviation in graphs, *Elsevier Linear Algebra and its Applications*, vol. 414, 2006.

[88] S. Pahwa, A. Hodges, C. Scoglio, and S. Wood, Topological Analysis of the Power Grid and Mitigation Strategies Against Cascading Failures, *In proceedings of 4th Annual International IEEE Systems Conference*, San Diego CA, April 2010.

[89] R. Pastor-Satorras, A. Vázquez and A. Vespignani, dynamical and correlation properties of the Internet, *Phys. Rev. Lett.*, 87, 258701, (2001).

[90] R. Pastor-Satorras and A. Vespignani, Epidemic Spreading in Scale-Free Networks, *Review Letters*, vol. 86, pp. 3200–3203, 2001.

[91] R. Pastor-Satorras and A. Vespignani, Epidemic dynamics and endemic states in complex networks, *Physical Review E*, vol. 63,066117, 2001.

[92] R. Pastor-Satorras and A. Vespignani, Immunization of complex networks, *Phys. Rev. E* 65, 2002.

[93] L.S. Pontryagin, L.S. et al., The Mathematical Theory of Optimal Processes, *Interscience, vol. 4*, September 1962.

[94] B. Prakash, H. Tong, N. Valler, M. Faloutsos and C. Faloutsos, Virus Propagation on Time-Varying Networks: Theory and Immunization Algorithms, *ECML-PKDD 2010*, Barcelona, Spain 2010.

[95] T. Reluga, Game theory of social distancing in response to an epidemic, *PLOS Computational Biology*, 6 (5): e1000793, 2010.

[96] T. Reluga and A. Galvani, A general approach to population games with application to vaccination, *Mathematical Biosciences*, 230 (2): 67-78, April, 2011.

[97] M. Rosas-Casals, S. Valverde, and R. Sole, Topological vulnerability of the European power grid under errors and attacks, *International Journal of Bifurcation and Chaos, Vol. 17, No. 7*, 2007.

[98] P. Schumm, C. Scoglio, T. Easton, and D. Gruenbacher, Epidemic Spreading on Weighted Contact Networks, in *proceedings of BIONETICS'07*, Budapest, Hungary.

[99] C. Scoglio, W. Schumm, P. Schumm, T. Easton,S. Roy Chowdhury, A. Sydney and M. Youssef, Efficient Mitigation Strategies for Epidemics in Rural Regions, *PLoS ONE*, 5(7): e11569, 2010.

[100] B. Stott, J. Jardim, and O. Alsac, DC Power Flow Revisited, *IEEE Transactions on Power Systems, Vol. 24, No. 3*, 2009.

[101] S. H. Strogatz, Exploring complex networks, *Nature*, 410, 268-276 2001.

[102] S. Towers and Z. Feng, Pandemic H1N1 influenza: predicting the course of a pandemic and assessing the efficacy of the planned vaccination programme in the United States., *Euro Surveill.*, vol. 14(41), 2009.

[103] H. C. Tuckwell, L. Toubiana, and J.-F. Vibert, Spatial epidemic network models with viral dynamics, *Phys. Rev. E*, 57, 2163, 1998.

[104] A. Vázquez, R. Pastor-Satorras and A. Vespignani, Large-scale topological and dynamical properties of the Internet, *Phys. Rev. E*, 65, 066130, 2002.

[105] E. Volz and L. A. Meyers , Epidemic thresholds in dynamic contact networks, *J. R. Soc. Interface*, 6, 233–241, 2009.

[106] Y. Wang, D. Chakrabarti, C. Wang, and C. Faloutsos, Epidemic spreading in real

networks: An eigenvalue viewpoint, in *22nd Symposium in Reliable Distributed Computing*, Florence, Italy, 2003.

[107] D. Watts and S. Strogatz, Collective dynamics of small-world networks, *Nature*, vol. 393, pp. 440–442, 1998.

[108] R. J. Williams and N. D. Martinez, Simple rules yield complex food webs, *Nature*, 404, 180-183, 2000.

[109] R. Yang, B.-H. Wang, J. Ren, W.-J. Bai, Z.-W. Shi, W.-X. Wang, T. Zhou, Epidemic spreading on heterogeneous networks with identical infectivity, *Physics Letters A*, 364, 189.193, 2007.

[110] M. Youssef, R. E. Kooij, and C. Scoglio, "Viral conductance: Quantifying the robustness of networks with respect to spread of epidemics," *Elsevier Journal of Computational Science*, Volume 2, Issue 3, Pages 286-298, August 2011.

[111] M. Youssef and C. Scoglio, On graph-based characteristics of optimal overlay topologies, *The International Journal of Computer and Telecommunications Networking, COMNET*, vol. 53, pp. 913–925, 2009.

[112] M. Youssef and C. Scoglio, An individual-based approach to SIR epidemics in contact networks, *JTB: Journal of Theoretical Biology, Elsevier, Volume 283, Issue 1 , pp 136-144*, 2011.

[113] University of oregon routeviews project, http://www.routeviews.org/.

[114] Internet routing registries, http://www.irr.net/.

[115] NERC: North American Electric Reliability Corporation, http://www.nerc.com/

[116] Standard TPL-001-0.1: System Performance Under Normal Conditions, Transmission System Standards: Normal and Emergency Conditions, http://www.nerc.com/files/TPL-001-0_1.pdf/

[117] PSTCA: Power Systems Test Case Archive, http://www.ee.washington.edu/