INTER-FACULTY MASTER PROGRAM on
COMPLEX SYSTEMS and NETWORKS
SCHOOL of MATHEMATICS
SCHOOL of BIOLOGY
SCHOOL of GEOLOGY
SCHOOL of ECONOMICS
ARISTOTLE UNIVERSITY of THESSALONIKI

Master Thesis

Title:

Network Robustness Analysis

Ανάλυση Ανθεκτικότητας Δικτύων

Chatziefstratiou Ioannis

SUPERVISOR: Professor Antoniou Ioannis, School of Mathematics AUTH

Thessaloniki, June 2018

ΔΙΑΤΜΗΜΑΤΙΚΟ ΠΡΟΓΡΑΜΜΑ ΜΕΤΑΠΤΥΧΙΑΚΩΝ ΣΠΟΥΔΩΝ στα
ΠΟΛΥΠΛΟΚΑ ΣΥΣΤΗΜΑΤΑ και ΔΙΚΤΥΑ
ΤΜΗΜΑ ΜΑΘΗΜΑΤΙΚΩΝ
ΤΜΗΜΑ ΒΙΟΛΟΓΙΑΣ
ΤΜΗΜΑ ΓΕΩΛΟΓΙΑΣ
ΤΜΗΜΑ ΟΙΚΟΝΟΜΙΚΩΝ ΕΠΙΣΤΗΜΩΝ
ΑΡΙΣΤΟΤΕΛΕΙΟ ΠΑΝΕΠΙΣΤΗΜΙΟ ΘΕΣΣΑΛΟΝΙΚΗΣ

Μεταπτυχιακή Διπλωματική Εργασία

Τίτλος Εργασίας:

Network Robustness Analysis

Ανάλυση Ανθεκτικότητας Δικτύων

Χατζηευστρατίου Ιωάννης

ΕΠΙΒΛΕΠΩΝ: Αντωνίου Ιωάννης, Καθηγητής, Τμήμα Μαθηματικών, ΑΠΘ.

Εγκρίθηκε από την Τριμελή Εξεταστική Επιτροπή την 26η Απριλίου 2018.

.....	.....	.....
Ι. Αντωνίου	Π. Μωυσιάδης	Β. Καραγιάννης
Καθηγητής Α.Π.Θ.	Ομ. Καθηγητής Α.Π.Θ.	ΕΔΙΠ ΑΠΘ

Θεσσαλονίκη, Ιούνιος 2018

……
Χατζηευστρατίου Ιωάννης
Πτυχιούχος Μαθηματικός Α.Π.Θ.

ABSTRACT

The use and study of networks has become more and more relevant in the recent years. Their functionality, as well as their ability to preserve it, is highly important in the world of today. This work is an attempt to collect the research on the various methods of measuring network robustness to malicious attempts to disconnect them, as well as the effectiveness of such attempts. Furthermore, we compare the robustness of four different directed networks both in name, using indices of robustness, and in practice, observing the effects of various deconstruction attempts on them. This step is performed using the 3.4.3 version of the R programming language on a 64-bit windows platform. Our findings agree, in general, with previous research, but bring up a few points that require further exploration.

Chapter 1 contains the introduction. In chapter 2 we provide the necessary background definitions on network theory. In chapter 3 we define and classify the various indices of network robustness and present some of their properties. In chapter 4 we present the methods of attack, as well as tables with the studies they have been tested on. In chapter 5 we present an example of network analysis on four directed networks. In chapter 6 we present a summary of the known results in the literature that has been presented in chapters 3 and 4, and compare them with our own findings in chapter 5. In chapter 7 we draw our conclusions and present the discussion of our work

KEY WORDS

Network robustness, Network deconstruction, Vertex ranking, Network connectivity.

Table of contents

Σύνοψη

Synopsis in Greek

Η παρούσα διπλωματική εργασία πραγματεύεται την ανάλυση της αντοχής των δικτύων. Αυτό απαιτεί διερεύνηση τριών κυρίως τομέων. Πρώτον, την μελέτη της δομής του εκάστοτε δικτύου. Δεύτερον, τον ορισμό κατάλληλων εκτιμητών, οι οποίοι να μπορούν να αντιρποσωπεύσουν αριθμητικά την ανθεκτικότητα ενός δικτύου. Τρίτον, την ανάλυση των διαφορετικών μεθόδων με τις οποίες μπορεί να επιτεθεί κάποιος στο δίκτυο.

Σκοπός της εργασίας είναι η παρουσίαση και ταξινόμηση των διαφόρων δεικτών που εκτιμούν την ανθεκτικότητα των δικτύων, καθώς και η παρουσίαση των διαφόρων μεθόδων επίθεσης σε δίκτυα που έχουν μελετηθεί εώς τώρα. Αυτό γίνεται από την μία πλευρά με μία εκτεταμένη βιβλιογραφική ανασκόπηση και από την άλλη με ένα παράδειγμα υπολογισμού της αντοχής ορισμένων κατευθυνόμενων δικτύων, και στη συνέχεια μελέτη διαφόρων επιθέσεων πάνω σε αυτά.

Η εργασία είναι δομημένη σε επτά κεφάλαια, τα οποία περιγράφονται αναλυτικά:

Στο πρώτο κεφάλαιο παρατίθενται οι βασικές ερωτήσεις του κειμένου, οι οποίες είναι

1. Πως μπορούμε να εκτιμήσουμε την ανθεκτικότητα, η την ευαισθησία ενός δικτύου όσον αφορά τις διάφορες μεθόδους επίθεσης;
2. Πως οι διάφοροι μέθοδοι επίθεσης επηρεάζουν ένα δίκτυο;
3. Πως ανταποκρίνονται διάφορες δικτυακές δομές απέναντι σε διάφορες επιθέσεις;

Στη συνέχεια αναλύονται οι λόγοι που αυτές οι ερωτήσεις είναι σημαντικές και ορίζονται οι πολύ βασικές έννοιες. Εν συντομία, οι ερωτήσεις αυτές είναι σημαντικές τόσο για την καλύτερη προστασία “καλών” δικτύων (όπως κοινωνικά δίκτυα, δίκτυα ηλεκτροδότησης ή αερομεταφορών) απέναντι τόσο σε προβλήματα που μπορεί να προκύψουν τυχαία (βλάβες, κακοκαιρία κλπ.) όσο και σε εσκεμμένες επιθέσεις (τρομοκρατία, διάδοση ψευδών ειδήσεων κλπ). Έπειτα, παρουσιάζονται οι σχετικές απόπειρες συλλογής και σύγκρισης δεικτών αντοχής που έχουν γίνει ως τώρα και τέλος, διατυπώνεται η δομή της υπόλοιπης εργασίας.

Στο δεύτερο κεφάλαιο διατυπώνονται όλοι οι βασικοί ορισμοί, όπως τι είναι γράφος, δίκτυο κλπ. Ορίζονται επίσης οι διάφορες "κλασσικές" δικτυακές δομές (Scale free, Erdos-Renyi κλπ.) καθώς και κάποιες βασικές δομικές μετρικές (assortativity, συντελεστής σύμπλεξης κλπ). Τέλος, διατυπώνονται κάποιοι μηχανισμοί που διέπουν ορισμένα δίκτυα, σύμφωνα με τις αντίστοιχες μοντελοποιήσεις που έχουν γίνει, όπως δυνατότητα επανασύνδεσης ακμών, ή δευτερεύουσες εξαφανίσεις κόμβων.

Στο τρίτο κεφάλαιο ορίζονται και ταξινομούνται οι διάφοροι εκτιμητές της ανθεκτικότητας των δικτύων, παρουσιάζονται οι συλλογισμοί που τους παρήγαγαν, καθώς και κάποιες ιδιότητές τους. Η ταξινόμηση γίνεται σε δύο στάδια. Αρχικά διαχωρίζονται οι εκτιμητές που εξαρτώνται από την επίθεση που μελετάται, από τους εκτιμητές που εξαρτώνται μόνο από την δομή του δικτύου. Στη συνέχεια, οι πρώτοι διαχωρίζονται περεταίρω σε αυτούς που εκτιμούν την χειρότερη δυνατή περίπτωση για το δίκτυο (όπως συνδετικότητα ακμών/κόμβων, integrity-ακεραιότητα, isoperimetric number, κλπ.) και σε αυτούς που μελετάνε την επίδραση που έχει μία οποιαδήποτε επίθεση στο δίκτυο. Στην δεύτερη περίπτωση, αναδιατυπώνουμε και επεκτείνουμε τη χρήση μίας μετρικής που έχει χρησιμοποιηθεί υπό τον όρο robustness, ως *a*-fragmentation threshold (α-κατώφλι κατακερματισμού). Αναδιαμορφώνουμε επίσης την μελέτη της γενικής αποδοτικότητας (global efficiency) του δικτύου όταν αυτό υπόκειται σε κάποια επίθεση, με τον να την απλοποιήσουμε, με σκοπό να χρησιμεύσει ως μέτρο σύγκρισης ανάμεσα στην αντοχή διαφορετικών δικτύων στην ίδια μορφή επίθεσης. Η δεύτερη κατηγορία δεικτών διαχωρίζεται στους δείκτες που προκύπτουν απ' ευθείας από το δίκτυο και σε αυτούς που προκύπτου από την αλγεβρική επεξεργασία του πίνακα γειτνίασης, ή του αντίστοιχου Λαπλασιανού.

Στο τέταρτο κεφάλαιο ορίζονται οι διάφορες επιθέσεις των δικτύων και ταξινομούνται οι έρευνες στις οποίες έχουν μελετηθεί. Η ταξινόμηση των ερευνών γίνεται σε πίνακες για κάθε είδος επίθεσης, σύμφωνα με:

1. Τα δίκτυα που μελετήθηκαν, στα οποία παρατίθενται οι εξής πληροφορίες:
   i. Η δομή τους.
   ii. Η ύπαρξη κατευθυνόμενων ακμών ή/και βαρών.

iii. Οι περεταίρω μηχανισμού που τα διέπουν (δομικοί, ή διαχείρησης ροής).

2. Η προέλευση του δικτύου (αν είναι κατασκευασμένο από κάποιο μοντέλο, ή τι δίκτυο είναι, πχ. Αερογραμμών, διακίνησης ναρκοτικών κλπ).
3. Τα μέτρα που χρησιμοποιήθηκαν για να εκτιμήσουν την αντοχή του δικτύου ή το μέγεθος της ζημιάς που δέχτηκε από την επίθεση.
4. Την σχετική αναφορά της έρευνας που έγινε.

Στο πέμπτο κεφάλαιο παρουσιάζεται ένα παράδειγμα ανάλυσης αντοχής δικτύων. Επιλέγονται τέσσερα κατευθυνόμενα διατροφικά δίκτυα (food webs), τα οποία παρουσιάζουν σχέσεις θηρευτή-θηράματος σε οικοσυστήματα της νότιας Φλόριδας. Υπολογίζονται κάποιοι δείκτες ανθεκτικότητας οι οποίοι προσαρμόστηκαν από μη κατευθυνόμενα δίκτυα. Συγκεκριμένα υπολογίζεται το natural connectivity, το οποίο λάβαμε το περεταίρω βήμα να το απλοποιήσουμε ώστε να μπορούμε να συγκρίνουμε δίκτυα διαφορετικής τάξης. Η διαφορά μεταξύ των μέτρων της πρώτης και της δεύτερης ιδιοτιμής, στηριζόμενοι στην βασική αρχή της απόκλισης του δικτύου από τον βέλτιστο χαρακτήρα good expansion (μέτρου ευπάθειας). Καθώς και το assortativity coefficient του δικτύου, προς εξέταση των αμφιβολιών για το αν μπορεί όντως να χαρακτηριστεί ως μέτρο ανθεκτικότητας, καθώς έχει συσχετιστεί θετικά και αρνητικά με την αντοχή των δικτύων σε διαφορετικές έρευνες. Τα μέτρα αυτά υπολογίστηκαν τόσο για το σύνολο του δικτύου, όσο και για τις ισχυρά συνδετικές γιγαντιαίες συνιστώσες τους. Σε αυτά τα δίκτυα πραγματοποιήθηκαν οκτώ διαφορετικές επιθέσεις. Συγκεκριμένα οι κόμβοι τους ταξινομήθηκαν και αφαιρέθηκαν με φθίνουσα σειρά σύμφωνα με τους εξής δείκτες:

1. Έσω βαθμός.
2. Έξω βαθμός.
3. Συνολικός βαθμός.
4. Ενδιαμεσότητα.
5. Ιδιοκεντρικότητα.
6. Συνολική επηροή 2 βημάτων (collective influence).
7. Συνολική επηροή 3 βημάτων.
8. Τυχαία (πραγματοποιήθηκαν πέντε διαφορετικές τυχαίες επιθέσεις και παρατίθενται οι μέσες τιμές των αποτελεσμάτων).

Η αποτίμηση του αποτελέσματος της επίθεσης έγινε υπολογίζοντας το σχετικό μέγεθος της ασθενώς και της ισχυρά συνδετικής γιγαντιαίας συνιστώσας και της συνολικής αποδοτικότητας (global efficiency), κανονικοποιημένης σύμφωνα με την αρχική τάξη του δικτύου. Παρατίθενται επίσης οι τιμές του εμβαδού αυτών των καμπυλών, καθώς και τα α-κατώφλια κατακερματισμού, για $\alpha = 0.25, 0.50$ και 0.75, για την ισχυρή και την ασθενή γιγαντιαία συνιστώσα κάθε δικτύου, που αντιπροσωπεύουν το πλήθος κόμβων που πρέπει να αφαιρεθεί για να μειωθεί η εκάστοτε συνιστώσα στο αντίστοιχο ποσοστό (ίσο με α) του αρχικού της μεγέθους. Τέλος, δίνονται τα συμπεράσματα του παραδείγματος. Το πιο ενδιαφέρον εξ αυτών είναι ότι το natural connectivity προβλέπει σχεδόν ακριβώς την διάταξη των δικτύων ως προς την μείωση της αποδοτικότητάς τους για κάθε επίθεση.

Στο έκτο κεφάλαιο, παρουσιάζονται τα σημαντικότερα συμπεράσματα που έχουν εξαχθεί από την βιβλιογραφία, ενώ στη συνέχεια συγκρίνονται με τα αποτελέσματα που έδειξε το παράδειγμα του πέμπτου κεφαλαίου. Αναφέρονται ενδεικτικά κάποια από τα πιο σημαντικά ή ενδιαφέρονται συμπεράσματα της βιβλιογραφίας.

- Τα δίκτυα εγκληματιών έχουν δειχθεί ότι είναι εξαιρετικά ανθεκτικά, ειδικά αν συνυπολογιστούν δυνατότητες επανασύνδεσης ακμών.
- Τα αεροπορικά δίκτυα είναι ευάλωτα σε μεγάλης κλίμακας φυσικές καταστροφές.
- Τα δίκτυα που έχουν εξελιχθεί φυσικά, έμφανίζουν πολύ πιο συχνά χαρακτήρα good expander από αυτά που έχουν κατασκευαστεί καθολικά από τον άνθρωπο.
- Διάφορα μέτρα ανθεκτικότητας ταξινομούν βασικές δικτυακές δομές με πολύ διαφορετικούς τρόπους.

Κατά την σύγκριση των αποτελεσμάτων του παραδείγματος με τα γνωστά συμπεράσματα της βιβλιογραφίας βρήκαμε ελάχιστες διαφορές, κυρίως όσον αφορά την αποδοτικότητα των μέτρων συνολικής επηροής (collective influence), τα οποία ήταν πολύ λιγότερο αποδοτικά από το αναμενόμενο.

Στο έβδομο κεφάλαιο παρατίθενται τα συμπεράσματα που έχουμε εξάγει από το σύνολο της βιβλιογραφικής αναδρομής, καθώς και από την σύγκριση του παραδείγματός μας με αυτήν. Τονίζεται η έλλειψη εκτεταμένης έρευνας στους περισσότερους δείκτες, και της διερεύνησης του τι αντιμετωπίζει ο κάθε δείκτης ως ανθεκτικότητα, καθώς οι

περισσότερες έρευνες γίνονται σε σχετικά λίγες δικτυακές δομές (πολλά δίκτυα, λίγα είδη μοντέλων). Ερμηνεύονται επίσης κάποιες αντιφάσεις που έχουν εμφανιστεί, στη βιβλιογραφία γενικά, αλλά και σε σχέση με το παράδειγμα του κεφαλαίου πέντε. Έπειτα, παρουσιάζονται γενικότερα τα κενά που εμφανίζονται στο σύνολο της έρευνας της ανθεκτικότητας των δικτύων. Παρατίθενται επίσης οι καινοτόμες ιδέες και η συμβολή της εργασίας. Τέλος, παρουσιάζεται ένα σύνολο ιδεών για περεταίρω έρευνα στον τομέα της ανάπτυξης δεικτών ανθεκτικότητας και βελτιστοποιημένων επιθέσεων.

Acknowledgements

This thesis was compiled to conclude the Inter-Faculty Master Program in Complex Systems and Networks at the Aristotle University of Thessaloniki.

I thank my supervisor, professor Antoniou I. for all of his precious time that he devoted to our cooperation, as well as all the rest of the teaching staff of the master program, for the care, effort and time they have invested in me and my colleagues. I am also grateful to professor Moysiadis Ch. and Dr. Karagiannis V. for their valuable teaching and insight.

Finally, I thank my friends and family for their unfailing support. I would never have made it without you.

Thessaloniki, June 2018

Chatziefstratiou Ioannis

1. Introduction

1.1. Questions-Research Subject

There are three questions addressed in this thesis.

Question 1: How can we measure the robustness or the fragility of a network with respect to different kinds of attacks.

Question 2: How do different methods of attack affect a network?

Question 3: How do different network structures respond to various attacks.

1.2. Why are these questions important?

Today's world is a highly networked one, and as such, we can find a lot of important reasons to study network robustness. The very structure of every organization, including society, is a network. Design of the structure of a corporation requires consideration of random failures, such as people falling ill or having accidents. Flow networks, such as public transport (Wilkinson et al. 2011, Zhang et al. 2011, Dunn and Wilkinson 2015, Yin et al. 2016, Zhang et al. 2018), water and electricity distribution (Motter and Lai 2002, Albert et al. 2004, Schneider et al. 2010, Schieber et al. 2015), are a cornerstone of modern life, and as such should be, and are, protected by optimizing their robustness to the most common dangers they face. For the sake of comprehension, we give some working definitions for the robustness and the vulnerability of networks.

Definition 1.2.1: *Robustness*

Robustness is the ability of a network to continue fulfilling its objectives adequately, when subjected to failures or attacks.

Definition 1.2.2: *Fragility/Vulnerability*

Fragility or vulnerability is the sensitivity of a network to failures or attacks (The opposite of robustness).

Equally important are the reasons to study how to more effectively damage, or completely deconstruct a network. For example, in law enforcement (Baker and Faulkner 1993, Krebs 2002, Raab and Milward 2003, Duijn et al. 2014), operations such as common fraud, drug trafficking, arms trafficking and terrorist groups such as Al Qaeda, constitute "dark" social networks. "Dark" networks also exist as parts of larger networks in the political level via connections to other networks, such as each other, or even legitimate ones.

Definition 1.2.3: *Attack on a network*

Attack on a network is any modification of the network resulting in degradation of its performance. For example, adding a node with negative edges or with false beliefs, or removing a node with the intent to disconnect parts of the network.

This thesis is concerned with the removal of nodes or edges with the intent to disconnect a network. Therefore, when the term "attack on a network" is used, it refers to node or link removals.

Definition 1.2.4: *Random removals*

Random removals on a network are removals of a set of nodes or links, realized by a random distribution. Usually, this distribution is chosen to be the uniform distribution. Random removals are simulated by random number generators.

Definition 1.2.5: *Targeted attacks*

Targeted attacks on a network are removals of a set of nodes or links, resulting from a specific strategy. For example, removals of Nodes with high Centrality, articulations points or bridges.

Definition 1.2.6: *Repeated/Simultaneous/Salvo attacks.*

Attacks are distinguished in two categories. Repeated attacks, where nodes or links are removed one by one until a certain number of attacks or a certain result is reached. Simultaneous or salvo attacks, where the set of nodes or links is removed once.

Due to the nature of the process of an attack or failure, as well as the way most of the studies are conducted, we will refer to both intended attacks and failures as attacks.

Numerous proposals have been made to assess network robustness. From the simplest ones, such as vertex or edge connectivity, to the more complex ones such as the scattering number of a graph. Such measures vary highly in many regards. Some are computational such as their complexity, others mathematical such as monotonicity, while others differ on their intent. Variables like the order of the network and the expected methods of attack against it, are paramount in considering a method of measuring its ability to withstand such attacks.

1.3. Relevant studies

Research on network robustness and network deconstruction has been carried out by several groups from various fields. As a result, many approaches have been proposed.

The presently available reviews are mainly focused on comparing few selected measures such as effective resistance with a few others (Ellens 2011, Ellens and Kooij 2013), or proposing a new, possible improved, measure (Wu et al. 2008, Wu et al. 2010, Schieber et al. 2015).

1.4. Methodology of the research

The methodology proposed and used in this paper, is an extensive review of all relative literature, and the classification of the methods to study network robustness as well as the methods of attack. We define and present some properties for the indices of robustness, and the various methods of measuring the impact of an attack on a network. Subsequently, we present the methods of attack that have been studied, and order them according to the types of networks they have been inflicted upon, as well as by the methods of estimating the robustness, or the impact of the attacks on these networks. Finally, we present an example of network robustness analysis in directed networks, and we compare our results to the known results of the literature.

The example was performed using the 3.4.3 version of the R programming language, in the R-studio software suite, version 1.1.423, and the following packages: igraph, Matrix and ggplot2.

1.5. Thesis outline

Chapter 1 contains the introduction. In chapter 2 we provide the necessary background definitions on network theory. In chapter 3 we define and classify the various indices of network robustness and present some of their properties. In chapter 4 we present the methods of attack, as well as tables with the studies they have been tested on. In chapter 5 we present an example of network analysis on four directed networks. In chapter 6 we

present a summary of the known results in the literature that has been presented in chapters 3 and 4, and compare them with our own findings in chapter 5. In chapter 7 we draw our conclusions and present the discussion of our work, as well as its original aspects and draw the next steps to be taken.

2. Networks

2.1. Introduction

In this chapter we will define the various network structures that have been used and studied in the literature, as well as the relevant limitations and properties that the networks in these studies have.

2.2. Basic concepts

Definition 2.2.1: *Graph, Node/Vertex, Edge, Directed, Undirected, Weighted, Unweighted, Simple graph, Size and Order of a graph*

- An undirected, unweighted graph is an ordered pair $G = (V, E)$ comprising a set V of vertices or nodes or points together with a set E of edges or arcs or lines or links, which are 2-element subsets of V (i.e. an edge is associated with two vertices, and that association takes the form of the unordered pair comprising those two vertices).
- If the set of edges E is comprised of ordered pairs, then the graph is a directed graph.
- If each element of the set of edges E has a real number associated to it, then the graph along with these values is called a weighted graph.
- If each edge $(u, v) \in E$ is unique in E and $\nexists\, v \in V : (v, v) \in E$, then the graph is a simple graph.
- The size of the graph is the number of its edges.
- The order of the graph is the number of its nodes

Definition 2.2.2: *Adjacency matrix, Weighted adjacency matrix*

- Let $G = (V, E)$ be a simple graph and with $V = v_1, v_2, \dots v_n$. Then it can be represented by a binary matrix A_{ij} with elements $a_{ij} = 1$ if there is an edge between v_i and v_j or $a_{ij} = 0$ if there isn't. This matrix is the adjacency matrix of the graph.
- If the graph is weighted, we can replace each a_{ij} that equals 1 with the corresponding edge weight, thus creating the weighted adjacency matrix.

Definition 2.2.3: *Network*

A network is a representation of relations between discrete objects as a graph. If such relations are symmetric, then the network is undirected and if they are asymmetric, the network is directed.

Definition 2.2.4: *Walk, Trail, Path, Length of a path, Distance between nodes, Diameter of a graph, Geodesic*

- A walk is an alternating sequence of vertices and edges, starting and ending at a vertex, in which each edge is adjacent in the sequence to its two endpoints. In a directed graph the ordering of the endpoints of each edge in the sequence must be consistent with the direction of the edge.
- A trail is a walk with no repeated edges.
- A path is a trail in which all vertices (except possibly the first and last) are distinct.
- The length of a path is the number of edges it contains.
- The distance between two nodes of a graph is the length of the shortest path between them.
- Diameter of the graph is the largest distance between two of its nodes inside it.
- Geodesic between two nodes is the distance of the shortest path between them.

Definition 2.2.5: *Connected, Weakly/Strongly connected*

- Let $G = (V, E)$ be an undirected graph. If for every pair $v_i, v_j \in V$ there exists a path connecting those vertices, then the network is connected.
- If G is directed and there exists a path from v_i to v_j or from v_j to v_i for every $v_i, v_j \in V$, then the network is weakly connected.
- If G is directed and there exists a path from v_i to v_j and from v_j to v_i for every $v_i, v_j \in V$, then the network is strongly connected.

Definition 2.2.6: *Tree*

An undirected graph is called a *tree* if there is only one path connecting each pair of nodes.

Definition 2.2.7: *Density*

The density of a network is the number of existing edges on it, divided by the number of possible edges. Alternatively, it is the probability that a possible edge of the network exists.

Definition 2.2.8: *Giant component, Giant bicomponent/biconnected component, Strong giant component, Weak giant component*

- The giant component of a network is the largest of its connected components. In the cases when there are multiple components of the same order of magnitude as the largest one, we can say that the network has more than one giant components.
- The giant biconnected component or giant bicomponent of the network is the largest subgraph where for every pair of nodes i, j there are at least two distinct paths connecting them.
- In the case of directed networks, we can distinguish the strong and the weak giant components, as the largest strongly connected and weakly connected components of the network.

Definition 2.2.9: *Nearest neighbor graph*

The nearest neighbor graph for a set of objects P in a metric space is a directed graph with P being its vertex set and with a directed edge from p to q where $q, p \in P$, whenever q is a nearest neighbor of p.

Sometimes the graph is considered as undirected, but it is important to note that the property of the nearest neighbor is not a symmetric one.

2.3. Network structure

Here we will define the various notions around a network's structure. Ordinarily, the categorization of networks focuses on their degree distribution and on occasion on other elements.

Definition 2.3.1: *Path/Linear graph, Cycle graph, Star graph, Wheel graph, Gear graph, Bipartite Wheel graph, Complete graph*

- A path graph or linear graph is a graph whose vertices can be listed in the order $v_1, v_2, \dots, v_n$ such that all the edges are of the form (v_i, v_{i+1}).
- A cycle graph is a path graph with the addition of the edge (v_1, v_n).
- A star graph is a tree with one central node and all other nodes connected to it.
- A wheel graph is a cycle graph with the addition of one node connected to all others.
- A gear graph, or bipartite wheel graph, is a wheel graph with a node added between each pair of adjacent graph vertices of the outer cycle.
- A complete graph is a graph in which every pair of distinct vertices is connected by an edge.

Definition 2.3.2: *Assortativity coefficient, Assortative/Non-Assortative/Disassortative network*

The assortativity coefficient r is the Pearson correlation coefficient of the degree between pairs of linked nodes (in this calculation, the link of the nodal pair is not included). Positive values of r indicate a correlation between nodes of similar degree, while negative values indicate relationships between nodes of different degree. When $r = 1$, the network is said to have perfect assortative mixing patterns, when $r = 0$ the network is non-assortative, while at $r = -1$ the network is completely disassortative.

$$r = \frac{\sum_{jk} jk\left(e_{jk} - q_i q_j\right)}{\sum_k [k^2 q_k - (k q_k)^2]} \qquad (2.3.1)$$

Where j, k are the degrees of the adjacent vertices associated with an edge, e_{ij} is the joint probability distribution of the residual degrees at either of the ends of a randomly chosen edge, and q_k is the normalized degree distribution

$$q_k = \frac{(k+1)P(k+1)}{\sum_k kP(k)} \tag{2.3.2}$$

Where $P(k)$ is the probability a node has degree k.

Definition 2.3.3: *Clustering coefficient*

Clustering coefficient of a network is the number of closed nodal triplets in the network divided by the number of all nodal triplets (a triplet being a set of three nodes with at least two edges between them). Alternatively:

$$C = \frac{3 \cdot \{number\ of\ triangles\}}{\{number\ of\ triplets\}} \tag{2.3.3}$$

Definition 2.3.4: *Erdos-Renyi model*

The Erdos-Renyi *model* refers to one of two methods of constructing graphs. In the first case, originally proposed by Erdos and Renyi 1959, one graph from all possible graphs with a fixed number of nodes (N) and edges (E) is picked uniformly at random. While in the second case, proposed by Gilbert 1959, N nodes are placed, and every edge exists with a probability p, defined in practice as $p = \frac{E}{N(N-1)}$.

Definition 2.3.5: *Scale free/Scale invariant, Clustered scale free*

- A scale free or scale invariant network is a network whose degree distribution follows a power law, at least asymptotically. That is, the fraction $P(d_n)$ of nodes in the network having d_n connections to other nodes goes for large values of d_n as $P(d_n) \sim d_n^{-\gamma}$ where γ is a parameter whose value is usually, but not necessarily, in the range $2 < \gamma < 3$.

- A clustered scale free network is a scale free network modelled with increased clustering coefficient. This is achieved by reserving some edges for each node to be connected to neighbors of neighbors (Holme and Kim 2002).

Definition 2.3.6: *Small world*

A small-world network is a type network in which most nodes are not neighbors of one another, but the neighbors of any given node are likely to be neighbors of each other and most nodes can be reached from every other node by a small number of steps. Specifically, a small-world network is defined to be a network where the average path length L between two randomly chosen nodes grows proportionally to the logarithm of the number of nodes N in the network, that is: $L \propto \log(N)$.

Definition 2.3.7: *Hierarchical network*

A hierarchical network is a subclass of scale-free networks distinguished by the behavior of the clustering coefficients: decreasing as the degree increases, and remaining invariant as the order of the network. We remind the reader that in most scale free networks, the clustering coefficients decrease as the order increases.

Dodds et al. 2003 developed a model that takes as input a "pure hierarchy", that is, a tree with a single node as a root that is connected to b (branching ratio) other nodes, who, in turn, are connected to b more nodes. There are L levels (including the root node in the 1st level). Afterwards, each edge, between every pair of nodes not already connected with an edge, appears with a suitably designed probability.

Definition 2.3.8: *Bipartite*

A bipartite graph is a graph whose vertices can be divided into two disjoint and independent sets U and V such that every edge connects a vertex in U to one in V. Vertex sets U and V are usually called the parts of the graph. Equivalently, a bipartite graph is a graph that does not contain any odd-length cycles.

Definition 2.3.9: *L-Expansion (Costa 2004)*

The L-expansion of a given network $G(V, E)$ (directed or not) is a graph where connections from node i to node j are established whenever there exists a self-avoiding path (i.e. never passing by the same node twice) of length L connecting i to j in G.

Definition 2.3.10: *Q-Augmentation (Costa 2004)*

The Q-augmentation of a network $G(V, E)$ (directed or not) is the union of every L-expansion of $G, \forall\ L \leq Q$.

For example

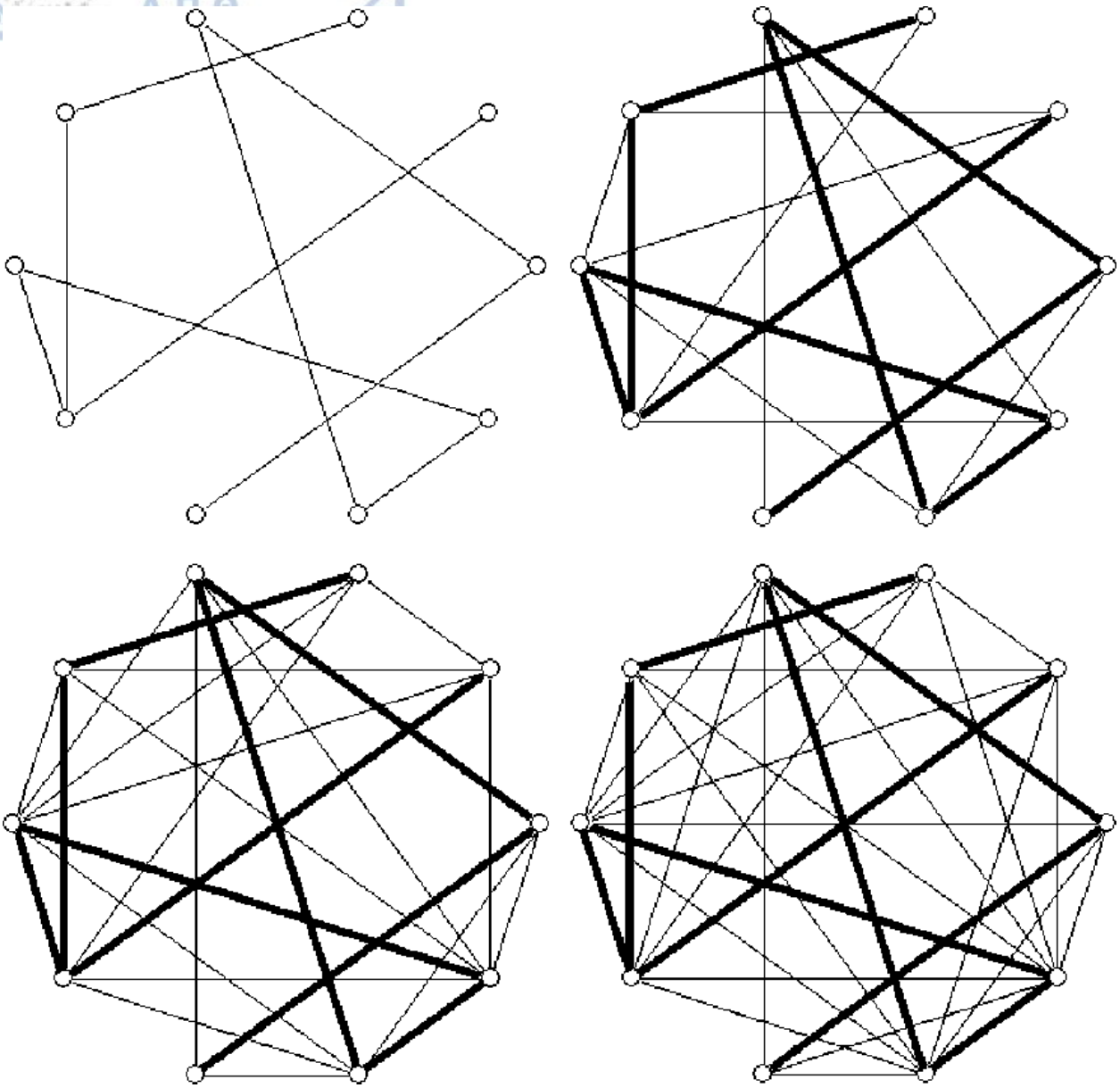

Figure 2.3.1: A simple graph and its respective 1, 2 and 3-augmentations. The added edges at each step are shown with thicker lines.

(Costa. 2004)

2.4. Network functionality

There are several real-world evolutions modeled by networks. For example, a public transport network could have values for the number of passengers that go from place to place, as well as the maximum number of passengers able to, or a food web might require that each species has access to some other species that it feeds upon.

Real networks usually involve structural and/or processability requirements. The structural limitations refer to conditions on the nodes and/or edges. For example, the nodes which if disconnected, certain functions will stop. For example, food webs (Allesina and Pascual 2009, Dunne et al. 2002) and plant-pollinator networks (Memmott et al. 2004, Kaiser-Bundury et al. 2010, Santamaria et al. 2014, Dominguez-Garcia and Munoz 2015 and Garcia-Algarra et al. 2017.)

Examples of processability conditions are: In flow networks, there is a limit to how much flow a node or an edge can handle. For example, there is finite space in a train and thus, a finite number of people can travel at any given time, another example is the electrical current that a power line can hold before it fails due to overloading. In some cases, like the train example, the edge continues functioning, but only up to its capacity. Such cases have been studied by Dodds et al. 2003, Wilkinson et al. 2011, Wagner 2015, Dunn and Wilkinson 2015. While in other cases, like the power line example, the node or edge stops functioning entirely, and the respective flow might be lost or redistributed in the network, causing other edges or nodes to fail. Such cases have been studied by Motter and Lai 2002, Moreno et al. 2003

Examples of networks involving both processability and structural limitations are: electricity or water distribution networks. There are certain nodes that any node must be connected to at least one of them, such as power distribution stations or water tanks, and there is always an upper limit on flow between nodes. Such a case has been studied by Albert et al. 2004.

There are also networks with the ability to rewire the edges whose start or end node has failed. For example, air traffic networks, where a plane can land on a different airport

than the one it started in. We call such networks self-rewiring networks. Such networks have been studied by Kaiser-Bundury et al. 2010, Wilkinson et al. 2011, Duijn et al. 2014 and Dunn and Wilkinson 2015.

3. Indices for robustness and vulnerability

3.1. Introduction

In this chapter, we will analyze the methods to measure the robustness or the fragility of a network. We shall divide these into two categories, with a subcategory each.

- Measures defined according to an attack. This means picking a strategy to remove nodes, and defining the mechanics of the network, and then measuring the impact the specific attack strategy has according to these mechanics.
    - Worst scenario measures. Derived by measuring the most effective attack possible on a network, according to the mechanics of the network and the variables we are interested in.

Quoting Li et al. 2005: *In an analysis of the vulnerability of networks to disruption, three important quantities [...] are (1) the number of elements that are not functioning, (2) the number of remaining connected subnetworks and (3) the order of a largest remaining group within which mutual communication can still occur.*

- Measures derived from the networks structure. These measures are dependent only on the network itself, and perhaps its mechanics, and are independent of any attack method. Such measures are usually obtained by making compromises on their accuracy, since the impact of an attack can vary greatly, but are expected to have their highest values for complete networks (networks that are the hardest to disconnect).
    - Spectral graph measures. This subcategory includes measures derived from the eigenvalues and eigenvectors of the adjacency or the Laplacian matrix, their weighted variants, or other matrices derived by them.

There can be cases of overlapping in these subcategories, as many network measures can be calculated both by calculations on the network and by spectral methods. Such cases are listed among the spectral measures, since they are usually easier to calculate this way.

3.2. Worst case scenario - attack dependent measures

Counterintuitively, the simplest, and historically the first, methods of measuring a network's robustness, are worst case scenario attack dependent measures. Such measures are unique for each network and procure a quantitative estimate for the damage a network can take while maintaining a specified condition. As such, they have been widely studied for a lot of basic graph structures. However, they frequently demand the calculation of every possible attack, and they cannot always be adjusted easily for changes in the network. Because they are computationally taxing they are not generally used in real world network studies. A point to be made, is that most of the following measures can be used to measure the course of the impact of a specific, repeated, attack on a network. However, this would not necessarily be a decreasing function of removed nodes.

Connectivity indices

Definition 3.2.1: *Vertex connectivity, Edge connectivity*

- The vertex connectivity of an undirected and unweighted graph is defined as the smallest cardinality of a set of vertices $S \subset V(G)$, such that $G - S$ is a disconnected graph. The vertex connectivity is usually denoted by κ_v.
- The edge connectivity of an undirected, unweighted graph is defined as the smallest cardinality of a set of edges $S \subset E(G)$, such that $G - S$ is a disconnected graph. The edge connectivity is usually denoted by κ_e.

As the removal of a node, entails the removal of all the edges connected to it, the following is self-evident.

Corollary 3.2.1

For any undirected, unweighted graph, graph G, where $C_d(v)$ is the degree of node v, it follows that

$$\kappa_v \leq \kappa_e \leq \min_{v \in V(G)} C_d(v) \qquad (3.2.1)$$

Definition 3.2.2: *Conditional connectivity (Harary 1983)*

Conditional connectivity of any graph G, with respect to some specific property P is defined as the smallest cardinality of a set of vertices $S \subset V(G)$, such that every component in $G - S$ has the property P. The conditional connectivity is usually denoted by κ_v^P.

The same definition can be applied for a conditional edge connectivity. It is obvious that both edge and vertex connectivity are specific instances of their respective conditional connectivity, with the property of being connected to the rest of the nodes in the graph.

The vertex and edge connectivity of the network give answer to the question of how many nodes or edges must be removed to create a disconnection. However, it does not address the issue of how severe that disconnection is. For example, a complete network K_N with an extra node connected with only one other, will have the exact same edge and vertex connectivity as any tree.

Isoperimetric number

Expanding on the idea of connectivity, the isoperimetric number is defined, based on the notion that smaller graphs will inevitably have small values for their connectivity, yet larger graphs with similar values are flawed.

Definition 3.2.3: *Isoperimetric number/Cheeger constant (Mohar 1989)*

Let $G(V, E)$ be a finite graph. If $X \subseteq V$ then ∂X denotes the edge-border, that is the set of edges of G that have one end in X and the other end in $V \backslash X$. The quantity

$$i(G) = \min_{X \subseteq V} \frac{|\partial X|}{|X|} \tag{3.2.2}$$

Where the minimum is taken over all non-empty $X \subseteq V$ satisfying $|X| \leq \frac{1}{2}|V|$, is called the isoperimetric number of G, also called the Cheeger constant of G.

Fault diameter

Fault diameter of a network is the greatest damage to its ability to communicate, when subjected to an intentional removal of nodes that is not enough to disconnect it.

Definition 3.2.4: *Fault diameter (Krishnamoorthy and Krishnamurthy 1987)*

Fault diameter of a graph G with vertex connectivity κ_v is defined as the largest obtainable diameter of that graph, after the removal of $\kappa_v - 1$ vertices. It is symbolized as f_G.

Definition 3.2.5: *Strongly resilient, Weakly resilient (Krishnamoorthy and Krishnamurthy 1987)*

- A graph, or a class of graphs is strongly resilient if there exists a constant t such that

$$f_G \leq \delta + t \tag{3.2.3}$$

- A graph, or a class of graphs is weakly resilient if there exists a constant t such that

$$f_G \leq t \cdot \delta \tag{3.2.4}$$

Toughness

Definition 3.2.6: *Toughness (Chvatal 1973)*

Toughness of an undirected graph G is the maximum real number t, such that deletion of $|S|$ nodes (can also be defined for edges) from G results in a connected graph, or one that has at most s/t components.

$$t(G) = \min_{\forall S \subseteq V} \left\{ \frac{|S|}{c(G-S)} \right\} \tag{3.2.5}$$

Where S denotes the set of deleted vertices and $c(G)$ the number of components of G.

Theorem 3.2.1

The toughness of a graph has the following properties

- Increases (not purely) with edge addition
- If G is not complete, then

$$t(G) \leq \frac{1}{2}\kappa_v(G) \tag{3.2.6}$$

- If G is not complete, then

$$t(G) \leq \frac{|V - MIS|}{|MIS|} \tag{3.2.7}$$

where MIS denotes the maximal independent set.

Proof: Chvatal 1973

The toughness of a graph has been studied for gear graphs by Kirlangic 2009 and for nearest neighbor graphs by Dunkum and Lanphier 2014.

Scattering Number

Definition 3.2.7: *Scattering number (Jung 1978)*

Scattering number of a finite graph $G(V, E)$ is defined as

$$sc(G) = \max_{\forall S \subseteq V}\{c(G - S) - |S| : \mathrm{c(G - S)} \neq 1\} \qquad (3.2.8)$$

Where $c(G)$ denotes the number of components in G.

Definition 3.2.8: *Edge scattering number (Aslan 2014)*

Scattering number of a finite graph $G(V, E)$ is defined as

$$sc_e(G) = \max_{\forall S \subseteq E}\{c(G - S) - |S| : \mathrm{c(G - S)} \neq 1\} \qquad (3.2.9)$$

Where $c(G)$ denotes the number of components in G.

"The scattering number is in a certain sense the 'additive dual' for the concept of toughness" (Jung 1978).

Its basic properties, along with some parallelisms with toughness have been studied by Kirlangic 2002. Some of its properties in Hamiltonian graphs have been observed by Zhang and Wang 2001. It has also been studied for split graphs by Li et al. 2008 and nearest neighbor graphs by Dunkum and Lanphier 2014.

Integrity

Definition 3.2.9: *Integrity (Barefoot et al. 1987, as presented by Goddard and Sward 1990)*

The integrity of a graph $G(V, E)$ is defined as such

$$I(G) \coloneqq \min_{S \subset V}\{|S| + m(G - S)\} \tag{3.2.10}$$

Where $m(G)$ is the order (number of vertices) of the largest component of G.

Definition 3.2.10: *Edge integrity (Bagga et al. 1992)*

The integrity of a graph $G(V, E)$ is defined as such

$$I_e(G) \coloneqq \min_{S \subset E}\{|S| + m(G - S)\} \tag{3.2.11}$$

Where $m(G)$ is the order (number of vertices) of the largest component of G.

Integrity is based on the idea that an attacker wants to disconnect a network as much as possible, with the least amount of attacks (Bagga et al. 1992). It has the advantage of not being oversensitive in local weakness (Goddard and Sward 1990), as for example having a relatively small number of isolated nodes connected in a single other node being an obvious liability to some of the aforementioned measures.

Theorem 3.2.2

- $$I(G) \leq VC(G) + 1 \tag{3.2.12}$$

  where $VC(G)$ is the vertex cover number of G.

- $$1 \leq I(G) \leq p \tag{3.2.13}$$

where p is the order of the network.

- $$I(G) \geq \min\big(d(G)\big) + 1 \qquad (3.2.14)$$

- If G is connected and not trivial, then

$$I(G) = 1 + \min_{v \in V(G)} I(G - v) \qquad (3.2.15)$$

Proof: Goddard and Sward 1990.

Definition 3.2.11: *Integrity family (Goddard 1994)*

The integrity family of measures is the class of measures of the form

$$\Psi(G) \coloneqq \min_{\mathrm{X} \subset \mathrm{V(G)}\ or\ E(G)} \{|X| + \psi(G - X)\} \qquad (3.2.16)$$

Where $\psi(G)$ is a network parameter.

For example, for ψ being the order of the largest component we get integrity and for ψ being the number of components we get toughness.

A more general class of measures was also proposed as

$$\Phi^f(G) \coloneqq \min_{\mathrm{X} \subset \mathrm{V(G)}} f\big(|X|, \psi(G - S)\big) \qquad (3.2.17)$$

Where f is a given function.

The integrity has further been studied for cubic graphs by Vince 2004, for nearest neighbor graphs by Dunkum and Lanphier 2014, for split graphs by Li et al. 2008, for Harary graphs by Li et al. 2009 and in general terms by Goddard and Sward 1990, Beineke et al. 1991, Bagga et al. 1992, Drange et al. 1996 and by Aslan and Bacak-Turan 2016. While the

edge-integrity has been studied by Bagga et al. 1992, Laskar et al. 1993, Bagga et al. 1994 and Moazzami 2011.

Theorem 3.2.3

- $$2 \leq I(G) \leq I_e(G) \leq p \tag{3.2.18}$$
- $$I_e(G) \geq \left[2\sqrt{p}\right]^+ - 1 \tag{3.2.19}$$

If G is connected

- If $\kappa_e \geq 2$, then

$$I_e(G) \geq \min\left\{\left[\sqrt{2p \cdot \kappa_v(G)}\right]^+, p\right\} \tag{3.2.20}$$

Where p is the order of the graph.

Proof: Bagga et al. 1994.

There are a lot of measures based around the idea of integrity, but most of them have not been thoroughly studied, since other, more informative measures have appeared. Nevertheless, they are mentioned here in order to complete the list.

Definition 3.2.12: *Hub-integrity (Mahde et al. 2010)*

The hub-integrity of a graph is defined as

$$I(G) \coloneqq \min_{X \subset V(G)} \{|X| + m(G - X)\} \tag{3.2.21}$$

Where X is taken such that $\forall x, y \in V(G)\ \exists\ X -$ path in G between x and y.

Definition 3.2.13: *Mean integrity (Bagga et al. 1992)*

The mean integrity of a graph is defined as

$$J(G) \coloneqq \min_{X \subset V(G)} \{|X| + \bar{m}(G - X)\} \qquad (3.2.22)$$

Where $\bar{m}$ the average component order $\bar{m}(G) \coloneqq \frac{1}{p}\sum_{v \in V(G)} p_v(G)$, where p_v denotes the number of vertices of v-th component and p the number of vertices in the graph.

Mean integrity has also been recently studied by Aslan and Bacak-Turan 2016.

Tenacity

Integrity, toughness and scattering number, all take into account the number of nodes deleted in an optimal attack on the network. In addition, integrity considers the order of the largest component after the attack, where scattering number and toughness consider the number of the components remaining. Combining these ideas, we get the next two measures.

Definition 3.2.14: *Tenacity (Cozzens et al. 1995, as presented by Li et al. 2008)*

For a noncomplete graph G, tenacity is defined as

$$T(G) = \min_{S \subset V(G)} \{\frac{|S| + m(G - S)}{c(G - S)}\} \qquad (3.2.23)$$

Such that $\omega(G - S) \geq 1$. Where $m(G)$ denotes the order (number of vertices) of the largest component in G, and $c(G)$ the number of components in G.

Definition 3.2.15: *Mix-tenacity (Moazzami and Salehian 2008)*

For a noncomplete graph G, mix-tenacity is defined as

$$T_m(G) = \min_{S \subset E(G)} \{\frac{|S| + m_e(G-S)}{c(G-S)}\} \quad (3.2.24)$$

Such that $\omega(G-S) \geq 1$. Where $m(G)$ denotes the order (number of vertices) of the largest component in G, and $c(G)$ the number of components in G.

Definition 3.2.16: *Edge tenacity*

For a noncomplete graph G, *edge tenacity* is defined as

$$T_e(G) = \min_{S \subset E(G)} \{\frac{|S| + m(G-S)}{c(G-S)}\} \quad (3.2.25)$$

Such that $\omega(G-S) \geq 1$. Where $m(G)$ denotes the size (number of edges) of the largest component in G, and $c(G)$ the number of components in G.

It has been shown (Moazzami 1999) that in some cases of graphs, tenacity is more able to measure the differences in the vulnerability of graphs than toughness or integrity, *"between graphs that intuitively should have different levels of vulnerability"*.

The various instances of tenacity have further been studied by Li et al. 2008 for split graphs, by Moazzami 2010 for n-connected graphs with minimal edges and by Dunkum and Lanphier 2014 for nearest neighbor graphs.

Rupture Degree

Following the relationship of scattering number and toughness, rupture degree, the additive dual of toughness, is introduced.

Definition 3.2.17: *Rupture degree (Li et al. 2005)*

The rupture degree of an incomplete connected graph G is defined as

$$r(G) = \max_{S \subset V(G)} \{c(G-S) - |S| - m(G-S)\} \quad (3.2.26)$$

Such that $\omega(G-S) \geq 1$. Where $m(G)$ denotes the order (number of vertices) of the largest component in G, and $c(G)$ the number of components in G.

The rupture degree has further been studied by Dunkum and Lanphier for nearest neighbor graphs, by Li et al. 2008 for split graphs and by Kirlangic 2009 for gear graphs. Finally, by Li and Zhang 2010 the graphs with maximal rupture degree for a given number of edges and vertices are calculated, as well as the graphs with maximal number of edges for a given number of vertices and rupture degree, but the problem of finding the graph with minimal number of edges in the latter case, remains open due to its complexity.

Definition 3.2.18: *Mean rupture degree (Aslan and Bacak-Turan 2016)*

The mean rupture degree of an incomplete connected graph G is defined by

$$\bar{r}(G) = \max_{S \subset V(G)} \{c(G-S) - |S| - \bar{m}(G-S)\} \quad (3.2.27)$$

Such that $\omega(G-S) \geq 1$. Where $\bar{m}$ is the average component order $\left(\bar{m}(G) \coloneqq \frac{1}{p}\sum_{v \in V(G)} p_v(G)\right)$, where p_v denotes the number of vertices of the v-th component and p the number of vertices in G and $c(G)$ the number of components.

Elasticity of MGG-Robustness

Let f_{ij} denote the interaction, or flow, between each pair of nodes i and j, let also Ω indicate the total node interaction in the network, expressed as

$$\Omega = \sum_{i \in V(G)} \sum_{j \in V(G)} f_{ij} \tag{3.2.28}$$

and let Ω_X indicate the total node interaction inhibited by removing $X \subset E(G)$ from G, expressed as

$$\Omega_X = \sum_{i \in V(G)} \sum_{j \in V(G)} f_{ij} Z_{ij}^X \tag{3.2.29}$$

Where $Z_{ij}^X = 0$, if a path from i to j exists in $G - X$, and 1 otherwise.

Definition 3.2.19: *MGG-robustness (Matisziw et al. 2012)*

The MGG-robustness of a network, under an attack of magnitude k, is defined as the lowest value for the uninhibited total nodal interactions, expressed as

$$\Gamma^{\mathrm{k}} = \Omega - \left(\max_{X \subset E,\ |X|=k} (\Omega_X) \right) \tag{3.2.30}$$

Originally called simply robustness, we call it MGG-robustness for clarification purposes, after the initials of the authors of the article.

Definition 3.2.20: *MGG-elasticity/Elasticity of MGG-robustness (Matisziw et al. 2012)*

The elasticity of MGG-robustness or MGG-elasticity is defined as the percentage of change in the robustness of the network over the percentage of change in the number of deleted edges.

$$E_{\Gamma^k} = \left| \frac{\%\Delta \Gamma^k}{\%\Delta p} \right| \tag{3.2.31}$$

The MGG-elasticity was originally defined in a dynamic environment for each epoch. This has been omitted since it only served the purposes of comparing the network at different timeframes (essentially comparing different networks) and does not affect the definition.

"A large value in MGG-elasticity indicates greater potential for a mechanism to efficiently degrade network performance. Conversely, robustness elasticity less than 1.0 indicates decreasing returns to scale, where changes in network robustness are less sensitive to changes in the magnitude of arc deletion" (Matisziw et al. 2012).

The MGG-elasticity was demonstrated as a measure on the Internet2 backbone network.

3.3. General attack dependent measures

In this section, the various measures quantify the impact of a specific attack strategy on the network. These types of measures have the advantage of being able to compare the resilience of two different networks to a specific attack, as well as the impact of two different attacks on a specific network.

The worst-scenario measures that take into account the values that Li et al. 2005 proposed (integrity, toughness, rupture degree, scattering number, tenacity and their variants), can be adapted to compare any attack pattern, or to find the optimal points in them. For example, one can compute and compare the impact of a few attack strategies by calculating the values of the tenacity formula,

$$\frac{|S| + m(G - S)}{\omega(G - S)} \tag{3.3.1}$$

without minimizing it over the possible attack, but rather calculating it for a specific attack.

Critical Threshold

The study of the order of the change in the network's components leads to the idea of considering the specific instance when an attack disintegrates the network into a lot of small, disconnected components.

Definition 3.3.1: *Critical threshold (Cohen et al. 2000)*

> The critical threshold of a graph subjected to an attack is the critical value p_c of the probability of removal existence of each node, below which the graph contains a connected cluster (its giant component) that spans the entire graph (its order is proportional to that of the entire graph); while beyond it, there is no such component.

Other definitions require the giant component to be of the same order of magnitude as the whole graph above criticality.

In simple terms this means that if the nodes that are removed are below this percentage, then the network resembles one large network, possibly with a few isolated clusters. If more nodes than this percentage are removed, then the network is consisted of many small clusters.

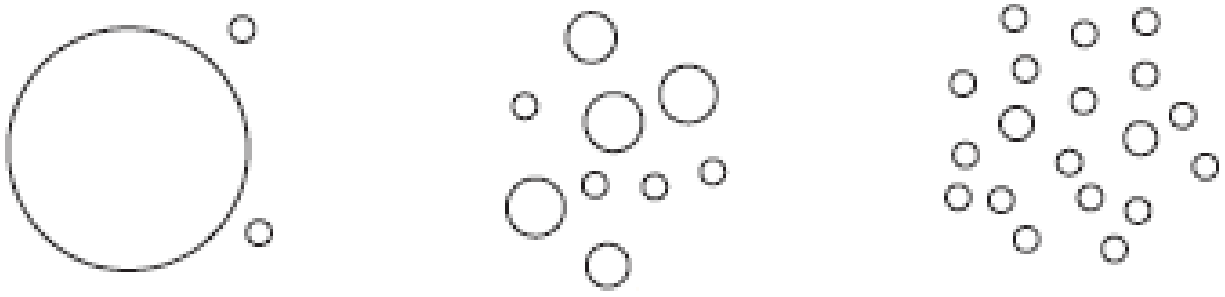

Figure 3.3.1: (from left to right) Above the critical threshold. At critical threshold. Below the critical threshold. (Albert et al. 2000)

This idea was first shown by Albert et al. 2000 for scale free and Erdos-Renyi networks, for random and degree-based attacks. It was further studied and used by Albert et al. 2000, Paul et al. 2004, Cohen et al. 2000, Valente et al. 2004, Tanizawa et al. 2005, Wu et al. 2007.

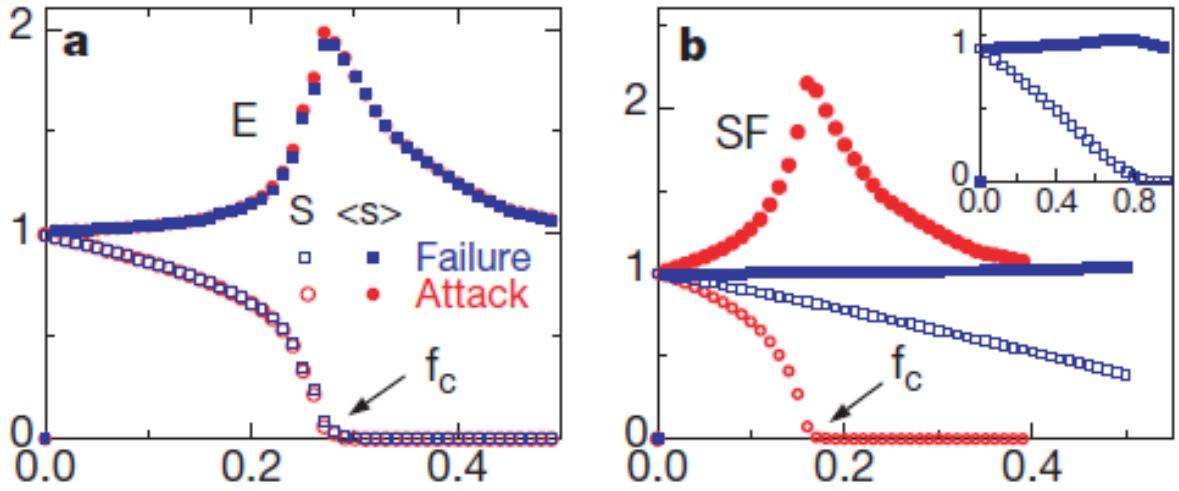

Figure 3.3.2: Network fragmentation under random failures and attacks on the nodes with the highest degree. Left panel is an Erdos-Renyi network, right panel is a Scale-Free. The relative order of the largest connected component S (open symbols) and the average order of the isolated components ⟨s⟩ (filled symbols) as a function of the fraction of removed nodes f. f_c denotes the instance where the network has no single distinct larger component.

(Albert et al. 2000)

As the definition is dependent on a variable order of a network model, which real networks may lack, we propose the following measure.

a-Fragmentation threshold

Definition 3.3.2: *a-Fragmentation threshold*

a-Fragmentation threshold, $0 \leq a \leq 1$, of a network $G(V, E)$ is the minimal percentage of nodes or links to be removed in order to split the giant component into components

with maximal order not greater than aN_{giant}, where N_{giant} is the order of the giant component. a is the percentage of the reduction of the giant component.

For example, $a = 0.5$ means that the resulting giant component contains at most half of the nodes of the original giant component, $a = 0$ means that all the nodes or links of the network have been removed, $a = 1$ means that no removals have occured.

This percentage can be calculated exactly in case of a deterministic method of attack, or approximated as an expected value in case of probabilistic attacks.

The amount of damage usually is a percentage of the disconnected, or cascaded nodes (Dunne et al. 2002, Buhl et al. 2004, Santamaria et al. 2014) or the time of appearance of many small clusters (Xu and Chen 2008).

Effect of the attack on the network's functionality

There is a variety of measures that can be important for a network, but not directly related to robustness. For example, a low value of the average path length can indicate a well-connected and efficient network. However, the average path length of a finite star graph is strictly less than 2, showing us that no message has to travel a long distance to reach its destination, but a single targeted node removal can shatter the network. It is also important to note that the star graph is very robust to edge removals and random node removals. Similarly, the wheel graph's average path length is also strictly lower than 2, but after removing the central node it rises to that of a cycle graph.

The previous examples indicate two things. Firstly, different methods of attack can have very different effects on a network and secondly, it can be important to consider the changes a network measure undergoes when the network is subject to a specific attack. This can be done by plotting the said measure over the number, of nodes or edges removed in the

network. The information of this plot can also be further compressed in a single value, by studying the area under this curve (AUC), but this is only useful if the measure studied is decreasing, not necessarily purely, as nodes or edges are removed. While also taking care to fulfill one of the following:

- Chose an attack that keeps the network connected.
- Chose a measure that can be defined for disconnected networks.
- At every disconnection chose only one component of the network to be studied.

Finally, when studying the area under the curve of a measure, it is more functional if both axes of the graph are normalized, so it can be used comparatively with greater ease.

Definition 3.3.3: *Area under curve*

Let k be a network measure defined for a network $G(V, E)$ according to its properties (directions, weights, etc.). We define the Area Under the Curve of the metric k, when the network is under n distinct node or edge attacks as such

$$AUC_k(G) = \frac{2\sum_{k=0}^{n} k_i - k_0 - k_n}{2\,n^2} \qquad (3.3.2)$$

Where $n \leq |V|$ in the case of node attacks, or $n \leq |E|$ in the case of edge attacks.

This is essentially the normalized average of the measure over the attack.

Change in the diameter or the average distance - Distance vulnerability

It has been noted (Albert et al. 2000, Ellens 2011) that the diameter of a network, and its ability to remain unchanged during an attack, is an important indicator for the effect an attack has on a network. Similarly, the average distance between nodes in a connected network shows us how well connected it is, and is less affected, as a measure, by long tails that might exist in the network. These measures are useful for comparing various sequences

of attacks, but only for small proportions of large networks, and only in cases where the network does not split in two large disconnected components. They do, however, give rise to various other measures based on the idea that disconnection is not the only danger a network faces when subjected to an attack, but also the increased difficulty on communications. Furthermore, the existence of frequently used long paths is a sign of vulnerability, because they are harder to protect against intentional attacks, and are more probable, to suffer random failures as they contain more nodes and edges. Therefore, a measure has been proposed (Ellens 2011) for flow networks whose flow is transferred by the shortest available path.

Definition 3.3.4: *Distance vulnerability (Ellens 2011)*

The distance vulnerability for a network G that has a traffic matrix T with entries t_{ij} and d_{ij} denoting the distance between nodes i and j, is defined as the average distance weighted by traffic.

$$\bar{d}^T = \frac{1}{\sum_{i=1}^{n}\sum_{j=i+1}^{n} t_{ij}} \sum_{i=1}^{n}\sum_{j=i+1}^{n} t_{ij} d_{ij} \qquad (3.3.3)$$

The problem of such a measure is that it does not take into account alternative paths, or different ways of distributing the traffic of the network.

Change in the order of the giant component

The fundamental concept of networks is interconnectedness. However, simply disconnecting one, or even a few nodes, cannot be considered as significant damage for a network with a few thousand of them. So, it follows, that the decline the largest component of the network undergoes, is an important index in the attempt to understand the impact an attack has on it, either by studying the order of the giant component relative to its original order, S'/S (Albert et al. 2000, Sole and Montoya 2001, Holme et al. 2002, Buhl et al. 2004,

Costa 2004, Estrada 2006, Xu and Chen 2008, Zhang et al. 2011, Iyer et al. 2013) or by studying the proportion of nodes still connected in the largest component, S'/N' (Deng and Wu 2015, Deng and Wu 2016), while another approach was to measure the percentage of nodes disconnected from the giant component (Albert et al. 2004). For the same network, the efficiency of different deconstruction strategies can be compared by examining the area under the curve of the relative order of the largest component and the percentage of removed nodes (Kasthrinathna and Mahendra 2013, Wagner 2015. This idea was also implemented by Schneider et al. 2010, who measured the average giant component order at every instance of the attack.

The order, normalized with respect to the order of the initial giant component, could also be used to compare the impact of a specific deconstruction strategy on different networks.

Another view of considering the interconnectedness, is considering the availability on alternative paths. Newman and Ghosal 2007 studied the change in the order of the largest biconnected component both as it appears in various network models, and as it declines when subjected to an attack. It was shown that in most network creation models (Newman and Ghosal 2007), the probability of a node to belong in a small bicomponent goes to zero as the network increases in order. It is also found that the same phenomenon arises in some real-world networks as well.

Dodds et al. 2003 studied the propagation of failures due to congestion and suggested another point of view of the change in the order of the giant component of a network under attack, attempting to bridge the simplicity of studying the order of the largest component and the importance of the magnitude of the attack.

Definition 3.3.5: *Connectivity robustness (Dodds et al. 2003)*

The connectivity robustness of a network is defined as

$$C_r = \frac{S}{N - N_a} \tag{3.3.4}$$

Where S is the order of the giant component after the removal of N_a $nodes$.

The opposite approach has been taken by Dunne et al. 2002 and by Memmot et al. 2004 for bipartite graphs, where they measured the cumulative indirect removals (caused by the mechanics of the network) over the direct removals (caused by the attacker). After taking into consideration the percentage of vertices removed to achieve these secondary extinctions, this measure contains the same information as the percentage of vertices still in the network.

Definition 3.3.6: *Extinction area (Allesina and Pascual 2009), Bipartite extinction area (Kaiser-Bunbury et al. 2010), Weighted extinction area (Kaiser-Bunbury et al 2010)*

- In networks with mechanics that cause nodes to fail as a result to an attack, extinction area is the area under the curve of the normalized cumulative number of indirectly removed nodes (nodes that failed due to the mechanics of the network, also called secondary extinctions) over the normalized number of directly removed nodes (nodes removed during the attack, also called primary extinctions). Extinction area is equal to 1 if all nodes fail after the first removal and equal to 0.5 if no secondary extinctions occur.
- The bipartite extinction area is defined for attacks only on one group of nodes, where only the indirect removals of the other group are calculated.
- If a value is attached to each node, for example, if the nodes represent species with different populations, then then instead of using the normalized number of nodes, one could use the normalized population. This is the weighted extinction area or the weighted bipartite extinction area, and the same limits apply for its values.

The extinction area was defined and used by Allesina and Pascual 2009.

The bipartite extinction area was defined and used by Kaiser-Bunbury et al. 2010, it was also used by Dominguez-Garcia and Munoz 2014.

The weighted variant was defined by Kaiser-Bunbury et al. 2010

SSSK-Elasticity

For the specific case of flow networks, the SSSK-elasticity of a network has been proposed

Definition 3.3.7: *SSSK-Elasticity (Sydney et al. 2008)*

The *SSSK-elasticity* of an undirected, unweighted network in respect to an attack is defined as the area under the curve (AUC) of

$$Tp(G) = \frac{1}{f_{max}} \sum_{jk} T_{jk} \tag{3.3.5}$$

over the percentage of the remaining nodes in the network, denoted as $El(G)$, where $T_{jk} = 0$ for $j = k$ and $T_{jk} = 1$ for $j \neq k$ and f_{max} denoting the maximum number of flows through a bottlenecked link.

Originally called elasticity, we refer to it as SSSK-elasticity, for clarification purposes, after the initials of the authors of the article.

SSSK-elasticity can be used to compare various methods of distributing traffic in a network as well as comparing the ability of networks to withstand attacks. It has been shown, on various networks (Sydney et al. 2008) subjected to highest degree attacks, to be correlated positively with assortativity.

Change in the efficiency of the network

Following closely the idea that the distance information must travel in the network is important, a measure is defined (Latora and Marchiori 2001 called it *global efficiency,* the same measure was called *average inverse geodesic length by* Holme et al. 2002) that encompasses the distance between every set of nodes in a single value.

Definition 3.3.8: *Efficiency/Global efficiency*

The efficiency or global efficiency of network is defined as the average inverse geodesic. In other words

$$E(G) = \frac{1}{N(N-1)} \sum_{i \neq j \in V} \frac{1}{d(i,j)} \tag{3.3.6}$$

Where $d(i,j)$ is the distance between the nodes i and j, and N is the number of vertices of G.

The definition may apply for weighted networks, after exchanging each weight defined in a sense of similarity with another, defined in a sense of distance.

Bocaletti et al. 2006 considered efficiency as an alternative, measure to the average path length.

Definition 3.3.9: *Local efficiency*

The local efficiency of a network is defined as

$$E_{loc}(G) = \frac{1}{N} \sum_{i \in V} E(G_i) \tag{3.3.7}$$

Where G_i is the subnetwork of G containing all the neighbours of node i but not i itself.

Bocaletti et al. 2006 considered local efficiency as an alternative, measure to the clustering coefficient.

The efficiency is not a measure of robustness. On the contrary it has been shown to be negatively correlated to it (Figure 3.3.2, Peng et al. 2016). However, studying the way the efficiency changes during an attack can help us understand the effect it has on a network. The disadvantages of the distance vulnerability carry on, meaning that the efficiency does not considered alternate paths unless we study its AUC during an attack.

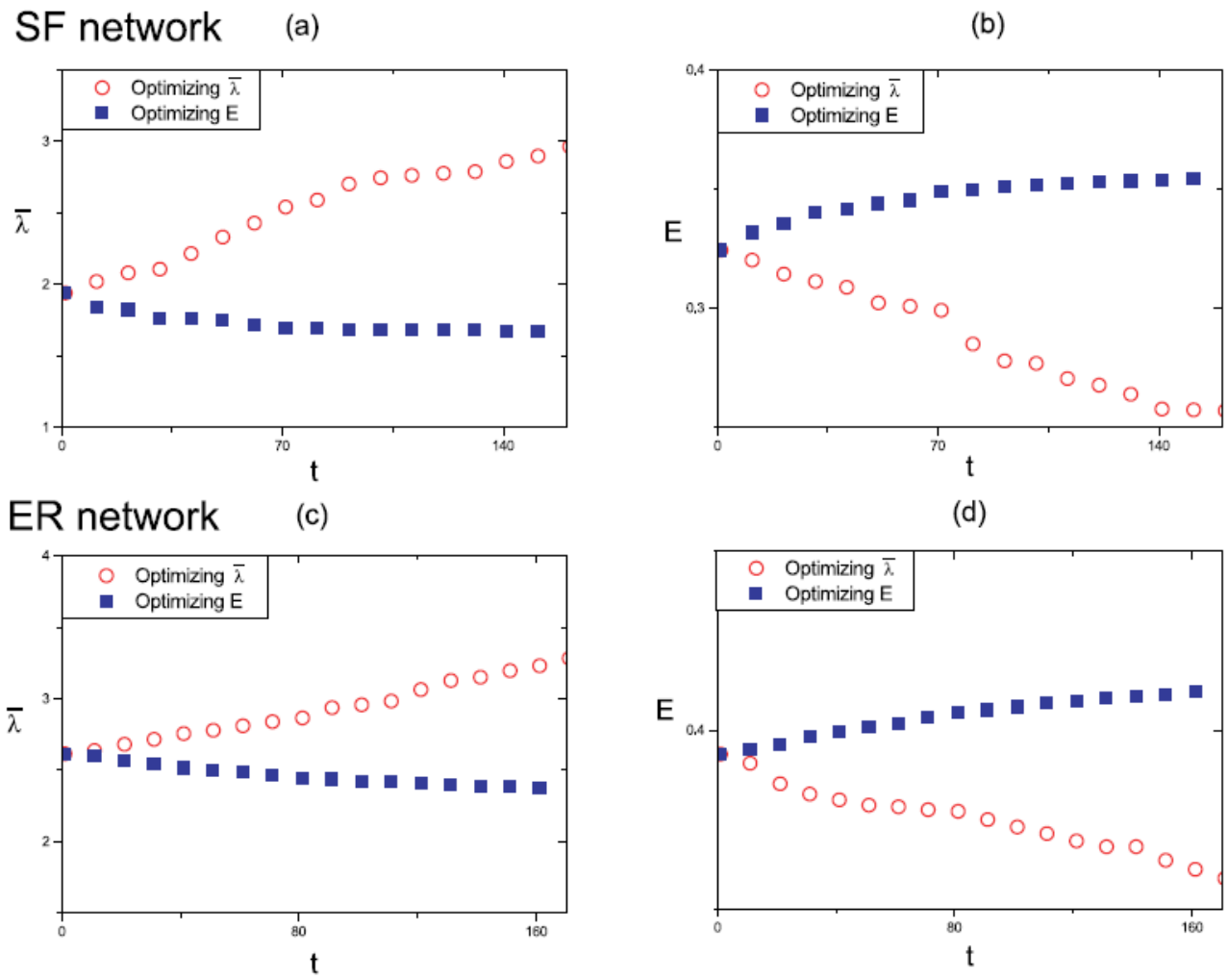

Figure 3.3.3: The change of the natural connectivity $\bar{\lambda}$ (left), and the efficiency E (right), versus iteration of a degree-preserving rewiring algorithm optimized to maximize either $\bar{\lambda}$ or E, on scale free networks (up) and Erdos-Renyi networks (down).

(Peng et al. 2016)

An exceptional advantage of network efficiency as a measure is that instead of the geodesic length, it can accept as input the distance travelled, and thus be used to study different methods of distributing the network's traffic, other than the shortest possible path (Pu et al. 2012).

Efficiency has been used by Holme et al. 2002 to compare the impact of various attacks on various networks, by Krebs 2002 to detect important actors on the internet and on terrorist networks, by Crucciti et al. 2003 to compare actors in criminal and terrorist networks, by Buhl et al. 2004 to study ant galleries, by Zhang et al. 2011 to compare attacks on the Shanghai subway network, by Yin et al. 2016 to compare attacks on the Beijing subway network and by Zhang et al. 2018 to study the subway networks of Beijing, Shanghai and Guanzhou when subjected to attacks.

Local efficiency has not been the subject of much research. It has been studied by Crucciti et al. 2003 and by Yin et al. 2016.

The efficiency is not necessarily decreasing as nodes are removed. If an isolated node is removed, then the efficiency of the network increases. However, the removal of a node is expected to decrease the efficiency. We propose the approach of treating each removed node, as having all his edges removed, but still existing completely isolated in the network, thus, normalizing the efficiency after a number of attacks by the number of nodes in the original network. This value is strictly decreasing if a non-isolated node is removed even if he belongs to an isolated cluster and remains constant on the removals of isolated nodes.

Theorem 3.3.1

The expected change for the global efficiency of a network if a node is removed is a decrease.

Proof:

Let $A = \sum_{i \neq j} \frac{1}{d(i,j)}$ be the sum of the inverse geodesic lengths of all the network, let also G be a network with $N > 3$ nodes and G' the same network with an arbitrary node k removed. Then

$$E_{global}(G) = \frac{A}{N(N-1)} \; and \; E_{global}(G') = \frac{A - Y}{(N-1)(N-2)}$$

Where Y is the sum of the inverse geodesics to and from the node removed plus a value for the change of the length of every geodesic k was a part of.

Let X be the sum of the inverse geodesics to and from the node removed. Then the expected value of X is

$$E(X) = E(X_{from} + X_{to}) =$$

$$= E\big(X_{from}\big) + E(X_{to}) =$$

$$= E\left(\sum_i \frac{1}{d(k,i)}\right) + E\left(\sum_j \frac{1}{d(j,k)}\right) =$$

$$= \frac{A}{N} + \frac{A}{N} = 2\frac{A}{N}$$

So

$$E(X) = 2\frac{A}{N} \Leftrightarrow$$

$$\Leftrightarrow E(X)N = 2A \Leftrightarrow$$

$$\Leftrightarrow E(X)N - AN = 2A - AN \Leftrightarrow$$

$$\Leftrightarrow AN - E(X)N = AN - 2A \Leftrightarrow$$

$$\Leftrightarrow \big(A - E(X)\big)N = A(N-2) \Leftrightarrow$$

$$\Leftrightarrow \frac{A - E(X)}{N-2} = \frac{A}{N} \Leftrightarrow$$

$$\Leftrightarrow \frac{A - E(X)}{(N-1)(N-2)} = \frac{A}{N(N-1)}$$

However, $X = Y$ *iff* the removal of node k does not affect the length of any geodesic $d(i,j)$ with $i, j \neq k$, that does not pass through k. This, $E(X) < E(Y)$. Therefore,

$$\frac{A - E(Y)}{(N-1)(N-2)} < \frac{A - E(X)}{(N-1)(N-2)} = \frac{A}{N(N-1)} \Leftrightarrow$$

$$\Leftrightarrow E\left(E_{global}(G')\right) < E_{global}(G)$$

■

The change in the efficiency of a network when under attack can be seen in the following example. Note that any form of successful attack is expected to be better (and thus create greater drops in the efficiency) than the presented random attacks.

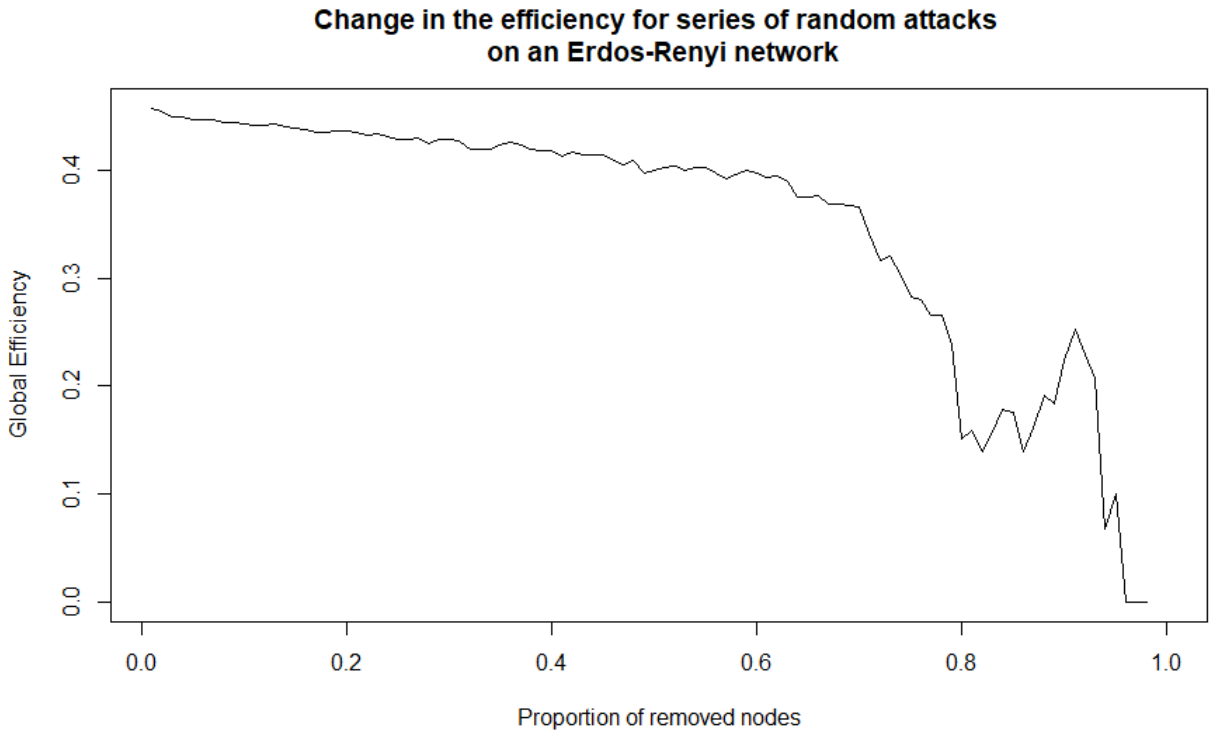

Figure 3.3.4: The drop in the global efficiency in an Erdos-Renyi network with 100 nodes and 400 edges, when subjected to a sequence of random node removals.

In order to study the efficiency as a decreasing value over the attack, we propose to consider an attack that removes a node as an attack that removes only the adjacent edges. This way the value of the efficiency constantly decreases (unless an isolated node is removed, when it remains constant) as seen in the following example.

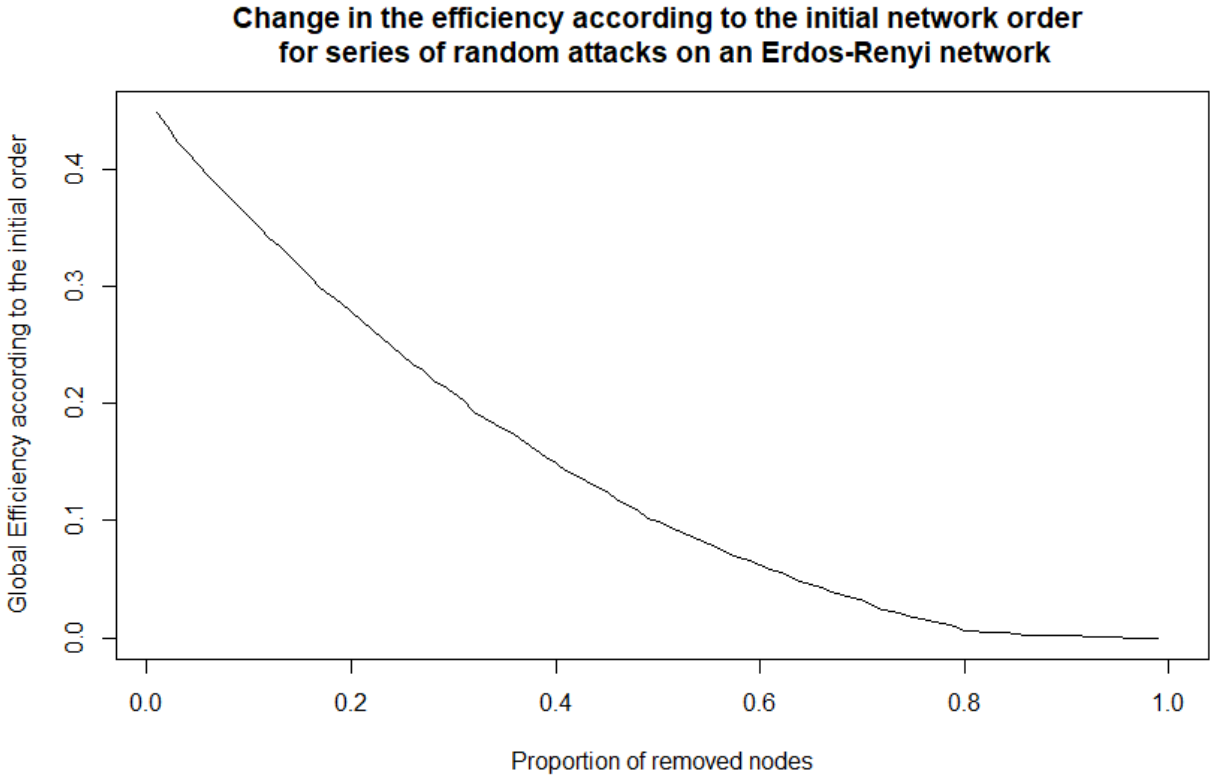

Figure 3.3.5:The drop in the global efficiency according to the initial network order on the same network as in figure 3.3.4 when subjected to the same sequence of random node removals.

JS-robustness

"Quantification of network robustness could be thought as the distance that a given topology is apart from itself after a failure" (Schieber et al. 2015).

Definition 3.3.10: *Shannon entropy, Jensen-Shannon divergence*

- The Shannon entropy of a probability distribution is defined as

$$H(P) = -\sum_i p_i \log(p_i) \tag{3.3.8}$$

- The Jensen-Shannon divergence between two probability distributions is defined as

$$J^H(P,Q) = H\left(\frac{P+Q}{2}\right) - \frac{H(P)+H(Q)}{2} \tag{3.3.9}$$

The Jensen-Shannon divergence is the square of a distance between probability distributions (Lin 1991).

The Jensen-Shannon divergence makes it possible to use the Jensen-Shannon divergence to compare the difference in the probability distribution of any measure in the nodes or edges of the network, for any given sequence of n failures.

Definition 3.3.11: *JS-robustness (Schieber et al. 2015)*

The info-theoretic rob*ustness* of a network is defined as

$$R_P\left(G\,\middle|(G_t)_{t\in[1,2,\dots,n]}\right) = \prod_{t=1}^{n}\left[1 - J^H\left(P(G_t), P(G_{t-1})\right)\right] = \prod_{t=1}^{n} R_p(G_{t-1}|G_t) \tag{3.3.10}$$

Where P is a probability distribution of a network measure, G is the original network and G_t is the network after t failures.

Originally called simply "robustness", we will call it JS-robustness for clarification purposes.

It is worthy of note that this measure can be used along any probability distribution, whether it is from a nodal measure such as degree, or an edge measure such as edge-betweenness. It can even be used with other types of distributions, as is demonstrated by Schieber et al. 2015, where the distance distribution is also used along with the degree.

Resilience Factor

We can consider the capability of a network to remain connected when vertices or edges fail in a combinatorial way, by looking at the number of possible connected networks as a result of a number of attacks over the number of all possible resulted networks.

Definition 3.3.12: *Resilience factor (Salles and Marino 2011)*

The resilience factor of a network G is defined as the average of all the fractions of connected subgraphs of G, where 1 up to $N-1$ vertices have been removed, over the number of all possible subgraphs after removal. Specifically

$$R_F = \frac{\sum_{i=1}^{N-2} k(i)}{(N-2)} \in [0,1] \qquad (3.3.11)$$

Where $k(i) = \frac{|connected\ subgraphs\ after\ i\ attacks|}{|all\ subgraphs\ after\ i\ attacks|}$, the denominator of which is equal to $\binom{n}{i} = \frac{N!}{i!(N-i)!}$.

The resilience factor can be used as a measure independent of the attack but has also been demonstrated (Salles and Marino 2011) to be a workable measure to detect changes on a network over an attack and has been compared alongside the average shortest path length and the diameter in such a case. However, this comparison has been implemented only in small networks and only for attacks that don't disconnect them, and thus, some more thorough examination is required to verify its usefulness, perhaps alongside other, more sophisticated measures.

3.4. Robustness measures dependent only on the network

Density

The density of the network (called connectance in the article) has been shown to be highly correlated with the number of nodes needed to be removed in a food web to achieve a 50% nodal cascade by Dunne et al. 2002.

Bounding network measures

A very simple and intuitive, yet sometimes effective way, to judge the robustness of a network, is to check the extremes of some of its key measures. We have already seen the minimal degree bounding various measures of robustness. Another example can be the nodes with high betweenness centrality being crucial to distributing information. In fact, it is around such ideas that the deconstruction strategies are based on. The maximal congestion centrality has been used as a measure of robustness by Dodds et al. 2003.

Definition 3.4.1*: Congestion centrality*

The congestion centrality of a node i is defined as the probability that any message sent between a pair of nodes (a, b) will pass through node i.

The congestion centrality depends on the method of distributing traffic in a network, and can be used to rank such methods for a specific network.

Assortativity coefficient

Newman 2002 showed that the assortativity of a network is negatively correlated to its robustness against random attacks, but positively correlated to its robustness against

targeted attacks. On the contrary, Vazquez and Moreno 2003 showed that power law networks are more robust when they are assortative even to random failures. Sydney et al. 2008 showed a positive correlation between assortativity and sssk-elasticity. Finally, Iyer et al. 2013 found assortative networks to be more vulnerable to degree attacks.

Reliability polynomial

In order to create more robust to random failure communications, the following measure has been proposed.

Definition 3.4.2: *Reliability polynomial (Moore and Shannon 1956, as presented by Ellens 2011)*

The reliability polynomial $Rel(G)$ of a graph G is equal to the probability that the graph is connected when each edge is (independently of the others) present with probability $p = 1 - q$, in other words

$$Rel(G) = \sum_{i=0}^{m} F_i (1-p)^i p^{m-i} \qquad (3.4.1)$$

Where F_i denotes the number of sets of i edges whose removal leaves G connected.

Theorem 3.4.1 *(Kelmans et al. 1981, as presented by Ellens 2011)*

For a graph with given size and order, there is no guarantee that a uniformly (for all p) optimal graph exists.

Proof: Kelmans 1981

The reliability polynomial has been used (Page and Perry 1994) to compare the criticality of two edges, by considering the inequalities $Rel(G - e_1) \geq Rel(G - e_2)$ and $Rel(G * e_1) \leq Rel(G * e_2)$, where $Rel(G * e)$ is the reliability polynomial given edge e has not failed and $Rel(G - e)$ is the reliability polynomial given edge e has failed. If the above inequations stand for all p, then it is safe to assume that edge e_2 is more critical than edge e_1. Such inequalities do not always stand, and in fact, reliability polynomials can cross each other more than once (Colbourn et al. 1993), or they might have the same direction (Page and Perry 1994).

Theorem 3.4.2 *(Moore and Shannon 1956)*

The relation between the reliability polynomial $Rel(G)$ of a graph and the edge connectivity $\kappa_e(G)$ satisfies the following two properties

1. If $\kappa_e(G_1) < \kappa_e(G_2)$ then, for p close enough to one $Rel(G_1) < Rel(G_2)$.
2. Let $ds(G)$ be the number of subsets of $\kappa_e(G)$ edges whose removal disconnects G. If $\kappa_e(G_1) = \kappa_e(G_2)$ and $ds(G_1) < ds(G_2)$ then for p close to one $Rel(G_1) < Rel(G_2)$.

Proof: Moore and Shannon 1956. Another proof given by Ellens 2011.

Graph diversity

For any pair of vertices in a graph, all different paths can be calculated and quantify the extent to which they diverge from the shortest path and each other.

Definition 3.4.3: *Path diversity Effective path diversity, Total graph diversity, Compensated total graph diversity (Rohrer and Sterbenz 2011)*

- The path diversity between paths a and b is defined as

$$D(P_{a,b}) = 1 - \frac{|P_a \cap P_b|}{\min(|P_a|, |P_b|)} \quad (3.4.2)$$

Where $|P_\gamma| = 2l + 1$, where l is the length of the path γ.

- The effective path diversity between nodes i and j is defined as

$$EPD_{ij} = 1 - e^{\lambda \sum_{i=1}^{k} D_{min}(P_i)} \quad (3.4.3)$$

Where λ is an experimentally determined constant that scales the impact of the diversity between paths a and b based on its utility.

- The total graph diversity is defined as the average of all effective path diversities.
- The compensated total graph diversity, (TGD) is defined in order to compensate for various path lengths as

$$cTGD = e^{TGD-1} \times h^{-a} \quad (3.4.4)$$

Where h is the average hop-count and a is an experimentally tuned parameter. (In Rohrer and Sterbenz 2011 it is suggested that $a = 1.125$ gives the best correlation to the simulation results).

Rohrer and Sterbenz 2011 ordered various networks by their TGD and cTGD along with other simpler measures.

3.5. Spectral Measures of Robustness

A network can be studied either directly, as we have seen until now, or through the matrices associated with it. Specifically, its adjacency matrix and its Laplacian.

Definition 3.5.1: *Laplacian matrix, Weighted Laplacian matrix*

- The Laplacian L is the difference $\Delta - A$ of the degree matrix Δ ($\Delta_{ii} = d(v_i), \Delta_{ij} = 0, i \neq j$) and the adjacency matrix A, i.e.

$$L_{ij} = \begin{cases} d(v_i), & i = j \\ -1, & if\ (i,j) \in E(G) \\ 0, & otherwise \end{cases} \tag{3.5.1}$$

- The weighted Laplacian L^W, for a weighted graph G with non-negative weights w_{ij}, is defined similarly as

$$L_{ij}^W = \begin{cases} s(v_i) = \sum_j w_{ij}\,, & i = j \\ -w_{ij}, & if\ (i,j) \in E(G) \\ 0, & otherwise \end{cases} \tag{3.5.2}$$

From the eigenvalues of the Laplacian of a graph, it is possible to determine the number of connected components it has.

Theorem 3.5.1

For the *Laplacian* of a graph $G(V,E)$, the multiplicity of the eigenvalue zero corresponds to the number of connected components of G.

Proof: Ellens 2011.

Theorem 3.5.2: *Weyl's Theorem (as presented by Ellens 2011)*

Let a weighted graph G be given and let G' be obtained by increasing the weight of an edge, the Laplacian eigenvalues of the new graph satisfy

$$\lambda_i^W(G') \geq \lambda_i^W(G) \tag{3.5.3}$$

This means that the algebraic connectivity increases (not purely) with edge addition.

Proof: Ellens 2011.

Algebraic Connectivity

Definition 3.5.2: *Algebraic connectivity (Fiedler 1972)*

The second smallest eigenvalue λ_2 of the Laplacian is called the algebraic connectivity.

Corollary 3.5.1: *(to theorem 8)*

The algebraic connectivity is equal to zero *iff* the graph is unconnected.

Lemma 3.5.1 *(Fiedler 1973)*

Removing k vertices, reduces the algebraic connectivity by at most k. More formally, let G be a graph and G_k a graph obtained by deleting k vertices from G, then

$$\lambda_2(G_k) \geq \lambda_2(G) - k \tag{3.5.4}$$

Proof: Fiedler 1973.

Lemma 3.5.2 *(Fiedler 1973)*

The algebraic connectivity is bound by its vertex and edge connectivity.

$$\lambda_2 \leq \kappa_v \leq \kappa_e \leq \min\big(d(G)\big) \tag{3.5.5}$$

Proof: Fiedler 1973.

Theorem 3.5.3 *(Fiedler 1973)*

The algebraic connectivity can be bound using the independent sets of $G(V, E)$.

$$\lambda_2 \leq |V| - |MIS| \tag{3.5.6}$$

Where MIS is a maximal independent set of G.

Proof: Fiedler 1973.

Definition 3.5.3: *Fiedler vector*

The Fiedler vector of a graph G is the eigenvector corresponding to the second smallest eigenvalue (i.e the algebraic connectivity).

Theorem 3.5.4 *(Maas 1987 as presented by Wang and Van Mighem 2008)*

Let G be a graph and G_e be the same graph after the addition of an edge e between nodes i and j. The upper and lower bounds of the algebraic connectivity $\lambda_2(G_e)$ are

$$\min\left\{\lambda_2(G) + \frac{\varepsilon a^2}{\varepsilon + (2 - a^2)}, \lambda_2(G) - \varepsilon\right\} \leq \lambda_2(G_e) \leq \min\{a^2 + \lambda_2(G), \lambda_2(G)\} \tag{3.5.7}$$

Where $a = |u_i - u_j|$, u_i being the i-th element of the Fiedler vector.

In the lower bound the first term increases with increasing ε, whereas the second one decreases.

The highest lower bound can be achieved by a choice of ε that makes both terms equal:

$$\varepsilon = \frac{\beta - 2}{2} + \left(\frac{(\beta - 2)^2}{4} + \beta(2 - \alpha)^2\right)^{\frac{1}{2}} \geq 0 \tag{3.5.8}$$

Where $\beta = \lambda_3(G) - \lambda_2(G) \geq 0$.

The higher a is, the higher is ε, and the higher the highest lower bound is. Higher a also contributes possible to a higher upper bound. Hence $\lambda_2(G_e)$ tends to be large if a is large.

Proof: Maas 1987.

The lower bound for vertex and edge connectivity that the algebraic connectivity presents, has been shown on various graph models, especially by Watts-Strogatz small world graphs (Figures 3.5.1 & 3.5.2), to be rather loose and their difference increases as the order of the network increases (Figure 3.5.3, Jamakovic and Uhlig 2007).

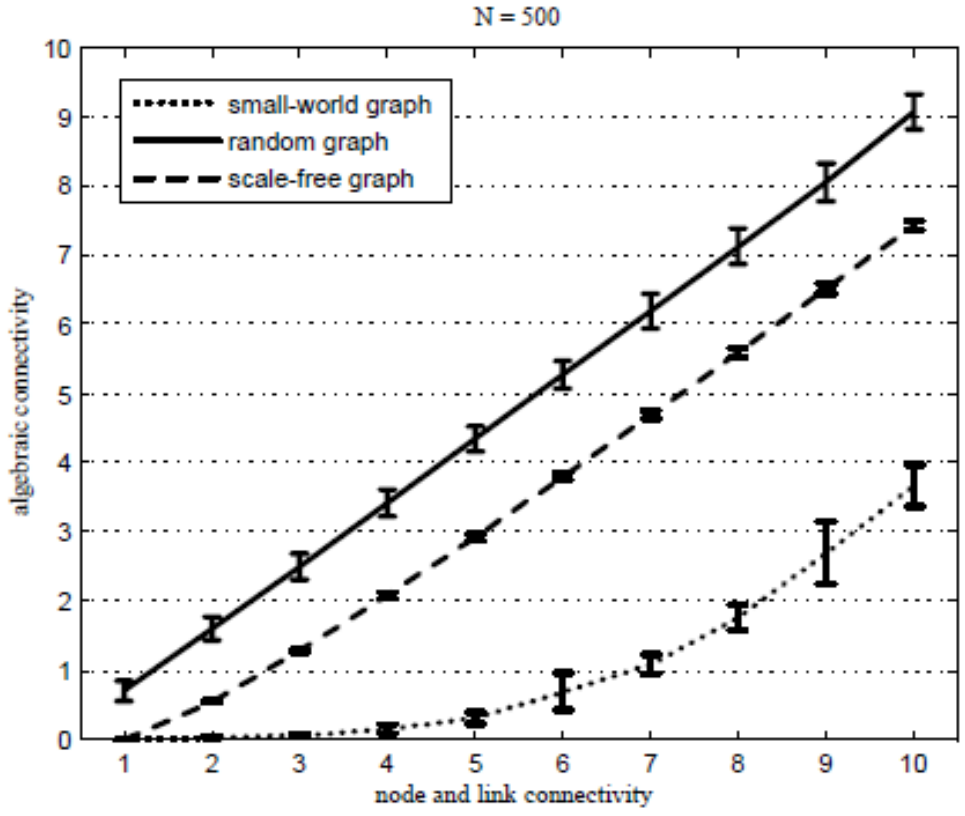

Figure 3.5.1: The mean as well as the standard deviation (error bars) of the algebraic connectivity as a function of the node and the link connectivity in the random graph of Erdos-Renyi, the small-world of Watts-Strogatz and scale-free graph of Barabási-Albert. All graphs have $N = 500$ nodes.

(Jamakovic and Uhlig 2007)

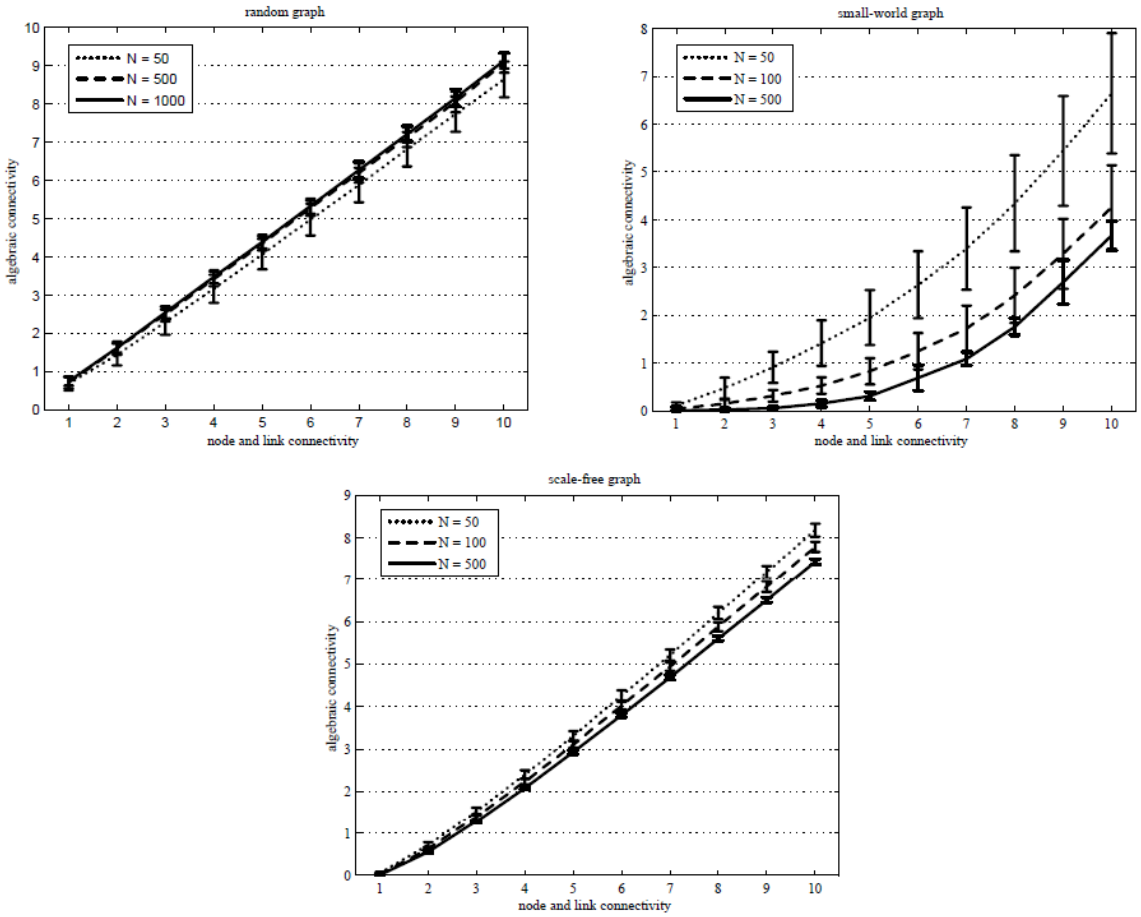

Figure 3.5.2: The mean as well as the standard deviation (error bars) of the algebraic connectivity as a function of the node and the link connectivity in the Erdos-Renyi random graphs, Watts-Strogatz small world graphs and Barabasi-Albert scale-free graphs with $N = 50, 100$ *and* 500 *nodes.*

(Jamakovic and Uhlig 2007)

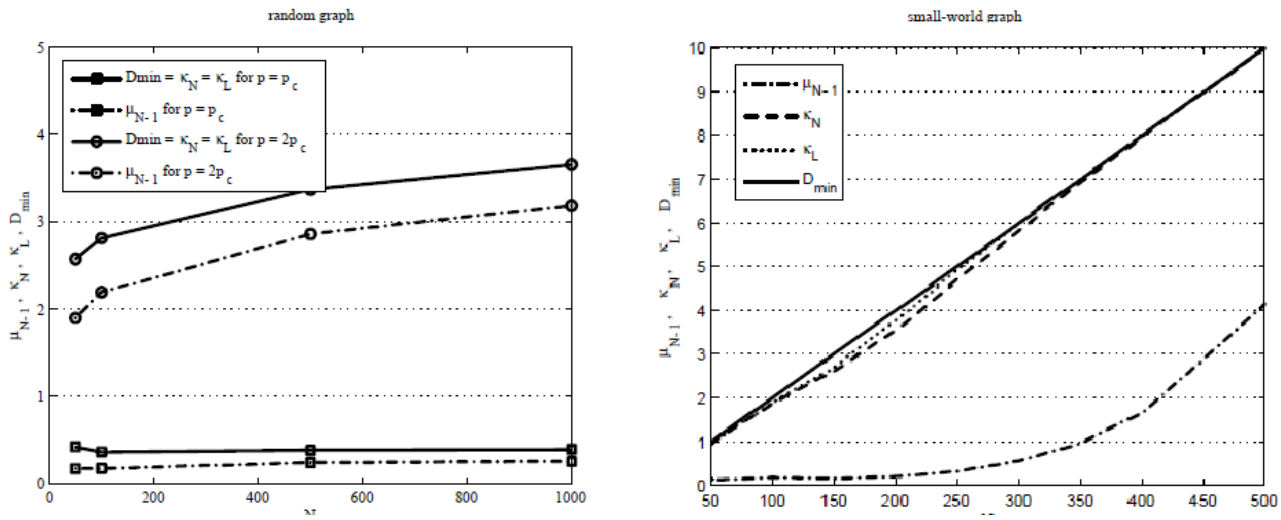

Figure 3.5.3: The mean over 10^3 *graphs of the node connectivity* κ_N*, the link (edge) connectivity* κ_L*, and the algebraic connectivity* μ_{N-1} *as a function of the number of nodes. Left for The Erdos-Renyi random graph with a given link density* $q = \frac{L}{L_{max}} = p$*.* p_c *is the critical value where the giant component appears. Right for the Watts-Strogatz small world model, with a given link density* $q = 0.04$*.*

(Jamakovic and Uhlig 2007)

Nevertheless, the algebraic connectivity serves as a stepping stone and a baseline to compare other measures. Ghosh and Boyd 2006 showed that optimizing the algebraic connectivity requires relatively few edge additions, and a greedy algorithm is presented for such optimization. Wang and Van Mighem 2008 discussed two edge addition strategies, simpler to compute, but close to optimal values, for increasing the algebraic connectivity of a graph. Specifically, they study the increase in the algebraic connectivity in two cases. After connecting the lowest degree nodes, and after connecting the nodes with the highest value for $a = |u_i - u_j|$ (u_i being the i-th element of the Fiedler vector), in Erdos-Renyi, Scale-Free and k-ary tree networks. The effectiveness of this increase is proved algebraically, but it is not tested with the network under attack.

Theorem 3.5.5

Adding an edge does not necessarily affect the first $n - 2$ Laplacian eigenvalues.

Proof: Ellens 2011 via example of a star graph.

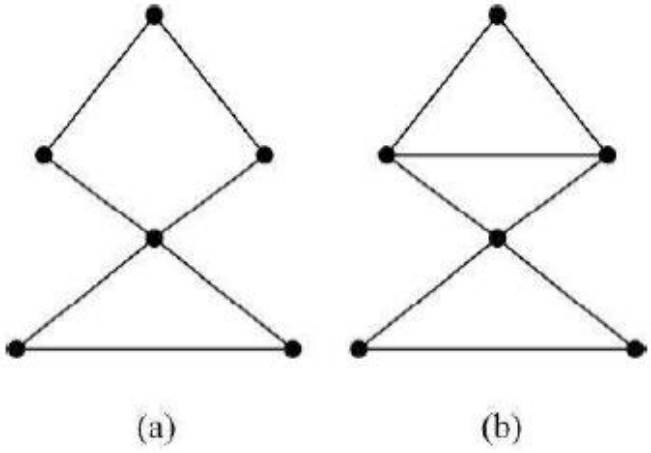

Figure 3.5.4: Two graphs with identical algebraic connectivity equal to 0.7369. (Wu et al. 2008)

Under this light, an interesting topic of research would be to study how the rest of the eigenvalues of the Laplacian project the robustness of the network.

In Wu 2005, the properties of the algebraic connectivity on directed, weighted and singed graphs are explored, showing that most properties still stand for directed graphs, and some of them stand for weighted and even singed ones. This is the only attempt so far to obtain a robustness measure in a signed graph, without altering the weights to positive values.

Theorem 3.5.6: *(Wu 2005)*

Let G be a directed graph

- If G has two vertices with 0 out-degree, then $a(G) \leq 0$.
- If G is unconnected $a(G) \leq 0$.
- If G has non-negative weights, $a(G) \geq 0$, if it is also connected $a(G) > 0$.

- If G has non-negative weights $a(G) \leq \kappa_v(G) \leq \kappa_e(G)$.

Proof: Wu 2005.

However, while adding undirected edges cannot decrease the algebraic connectivity of a graph, adding directed edges might, as shown in the following example by Wu 2005.

Figure 3.5.5: The empty graph has an algebraic connectivity of 0, However, the graph in this image has an algebraic connectivity of -0.0774.
Wu 2005.

Theorem 3.5.7: *Relating isoperimetric number and algebraic connectivity (Wu 2005)*

$$i(G) \geq \frac{a(G) - \frac{|V|}{2}}{\frac{|V|}{2}} \tag{3.5.9}$$

Proof: Wu 2005.

Theorem 3.5.8: *(Wu 2005)*

Let H be a set of vertices with zero in-degree in a graph G with non-negative weights. Then $a(G - H) \geq a(G)$.

Proof: Wu 2005.

Based on the principle that the number of alternate paths between each pair of nodes in a network plays a critical role in its robustness, but also on the notion that identifying the number of all alternative paths of different lengths between each node, the natural connectivity is proposed as a measure based on closed walks. To avoid repeated back-and-forths on an edge, the cycles are weighted by the factorial of their length. Let $N = |V|$ and $S = \sum_{k=0}^{\infty} \frac{n_k}{k!}$, where n_k is the number of closed walks in the network with length k. From matrix theory, we know that for symmetric matrices

$$n_k = trace(A^k) = \sum_{i=1}^{N} \lambda_i^k \tag{3.5.10}$$

Therefore,

$$S = \sum_{k=0}^{\infty} \frac{n_k}{k!} = \sum_{k=0}^{\infty} \sum_{i=1}^{N} \frac{\lambda_i^k}{k!} = \sum_{i=1}^{N} \sum_{k=0}^{\infty} \frac{\lambda_i^k}{k!} = \sum_{i=1}^{N} e^{\lambda_i} \tag{3.5.11}$$

Where S is equal to the sum of the subgraph centralities (see Estrada 2006).

Finally, noting that S increases with N it is scaled, and the natural connectivity is derived.

Definition 3.5.4: *Natural connectivity/Natural eigenvalue (Wu et al. 2008)*

The natural connectivity or natural eigenvalue of an undirected network is defined as the scaled sum of all closed walks in the network.

$$\bar{\lambda} = \ln\left(\frac{S}{N}\right) = \ln\left(\frac{\sum_{i=1}^{N} e^{\lambda_i}}{N}\right) \tag{3.5.12}$$

We note that there is another approach to defining the natural connectivity applicable to asymmetric matrices (and thus to directed graphs). Let a_{ij}^k denote the elements of A^k

$$n_k = trace(A^k) = \sum_{i=1}^{N} a_{ii}^k \tag{3.5.13}$$

$$S = \sum_{k=0}^{\infty} \frac{n_k}{k!} = \sum_{k=0}^{\infty} \left(\frac{1}{k!} \sum_{i=1}^{N} a_{ii}^k \right) = \sum_{k=0}^{\infty} \sum_{i=1}^{N} \frac{a_{ii}^k}{k!} = \sum_{i=1}^{N} \sum_{k=0}^{\infty} \frac{a_{ii}^k}{k!} = trace(e^A) \tag{3.5.14}$$

Theorem 3.5.9 *(Wu et al. 2008)*

- $\bar{\lambda}$ increases purely with edge addition
- $$\bar{\lambda} \leq N - \ln N \tag{3.5.15}$$

Proof: Wu et al. 2008.

The natural connectivity has been shown to produce better judgement than other measures, namely, edge connectivity, algebraic connectivity and critical threshold, on four different edge attack strategies (random, rich-rich, rich-poor, poor-poor (see chapter 4) on a scale free network (Wu et al. 2008, Wu et al. 2010, Wu et al. 2011) and on the Chinese internet (Wu et al. 2011). It has also been used to show that the efficiency as a single value is not a measure of robustness (Peng et al. 2016).

Number of spanning trees

Definition 3.5.5: *Spanning tree*

> Spanning tree of a connected undirected graph, is a tree subgraph that contains (spans through) all vertices. In other words, for an undirected graph G, a *spanning tree* is a connected subgraph that contains N vertices and $N - 1$ edges.

This definition is not exclusive. A graph can have multiple different spanning trees indicating alternate pathways in the network, and thus their number can be considered as a measure of robustness.

Although the number of spanning trees in a graph is a structural element, we consider it in this section as it is possible to determine it by the eigenvalues of the Laplacian.

Theorem 3.5.10 *(Van Mieghem 2010, as presented by Ellens 2011)*

The number of spanning trees in a simple graph G with Laplacian L and Laplacian eigenvalues $\lambda_i, i = 1,2, \dots, n$ is

$$\xi(G) = \frac{1}{n}\prod_{i=2}^{n}\lambda_i \tag{3.5.16}$$

Proof: Van Mieghem 2010.

Theorem 3.5.11 *(Ellens 2011)*

The reliability polynomial of a simple graph G, satisfies

$$Rel(G) = \xi(G)p^{n-1} + o(p^{n-1}), \;\; p \to 0 \tag{3.5.17}$$

Proof: Ellens 2011.

According to Ellens 2011, this indicates that the number of spanning trees is not a good measure of robustness to failure for real world networks, since, in such networks, random failures are generally scarce. In these cases, it is reasonable to assume that p will be close to one rather than zero.

An algorithm has been proposed (Tsen et al. 1994) that identifies the most important edge to be deleted according to the number of spanning trees with complexity $O(N^{2.376})$. The reversal of this algorithm can indicate edges to be added to strengthen the network. Another algorithm for optimizing the number of spanning trees by edge addition has been proposed by Baras and Hovareshti 2009.

Theorem 3.5.12 *(Baras and Hovareshti 2009)*

In a simple graph, the optimal edge to be added in order to maximize the number of spanning trees is between two nodes with maximal effective resistance distance (see below).

Proof: Baras and Hovareshti 2009.

Effective Resistance

Definition 3.5.6: *Effective resistance/Resistance distance*

By considering an undirected graph G as an electrical circuit, we can treat each edge (i, j) as a resistor of $r_{ij} = 1\ Ohm$ for unweighted graphs, or $r_{ij} = w_{ij}^{-1}$ for weighted ones, effectively treating the weight of an edge as its conductance. The resistance between a pair of nodes in the graph is called resistance distance, or effective resistance.

Theorem 3.5.13 *(Klein and Randic 1993)*

The effective resistance between two nodes of a simple graph i, j is a metric.

Proof: Klein and Randic 1993.

Theorem 3.5.14 *(Klein and Randic 1993)*

The effective resistance between two nodes of a simple graph i, j is non-increasing with edge addition.

Proof: Klein and Randic 1993.

This alone suggests that it can be used instead of the geodesic length in the various measures that depend on it. In addition, as we have already seen, minimizing the effective resistance is the best way to increase the number of spanning trees.

The effective resistance also agrees with some very basic intuitive ideas about the robustness of communications between two nodes. Specifically:

- Long paths offer greater resistance (n resistors in series give $R_{ij} = r_1 + r_2 + \cdots + r_n$)
- Alternate paths decrease the resistance (n resistors in parallel give $R_{ij} = \frac{1}{r_1^{-1} + r_2^{-1} + \cdots + r_n^{-1}}$)
- Increasing the number of alternate paths has diminishing returns (similarly with b, e.g. for one, two and three 1-ohm resistors we get effective resistance of $R_1 = 1, R_2 =$

$\frac{1}{2}, R_3 = 1/3$. Essentially, to halve the effective resistance we must double the conductivity.)

This way we can study both basic ideas of network robustness. Namely, the idea that shorter paths are more reliable since they are less prone to failure, and the idea that having alternative paths can keep a network connected even after certain parts of it have failed.

Although the effective resistance does not initially appear to be a spectral measure of the graph, it has been shown that it can be calculated from the Laplacian.

Theorem 3.5.15 *(Klein and Randic 1993)*

For a [weighted] undirected graph $G(V, E)$ with edge weights w_{ij}, the effective resistance between vertices i and j is

$$R_{ij} = \left(e_i - e_j\right)^T L^+ \left(e_i - e_j\right) = L_{ii}^+ - 2L_{ij}^+ + L_{jj}^+ \qquad (3.5.18)$$

Where L_{ij}^+ is the (i, j) element in the Laplacian pseudoinverse matrix and e_i is the vector with 1 on the i-th element and 0 on all others.

Proof: Klein and Randic 1993.

Definition 3.5.7: *Effective graph resistance/Total effective resistance*

The effective graph resistance or total effective resistance of an undirected graph G with N vertices is the sum of the effective resistances over all pairs of vertices.

$$R(G) = \sum_{i=1}^{N} \sum_{j=i+1}^{N} R_{ij} \qquad (3.5.19)$$

The total effective resistance is also called Kirchhoff index.

Theorem 3.5.16 *(Klein and Randic 1993)*

The total effective resistance of a connected weighted undirected graph G, satisfies

$$R(G) = N \sum_{i=2}^{N} \frac{1}{\lambda_i^W} \tag{3.5.20}$$

Proof: Klein and Randic 1993.

Corollary 3.5.2 *(Ellens 2011)*

The total effective resistance of an undirected, unweighted graph can be bounded by functions of λ_2^W in the following way

$$\frac{N}{\lambda_2^W} \leq R(G) \leq \frac{N(N-1)}{\lambda_2^W} \tag{3.5.21}$$

Theorem 3.5.17 *(Ellens 2011)*

- The total effective resistance of an undirected graph strictly decreases when edges are added, or weights are increased.
- Let $D(i,j)$ be the ordinary (shortest path) distance between nodes i and j. Then $R_{ij} \leq D(i,j)$, with the equality holding iff there is only one path between i and j.

Proof: Ellens et al. 2011

The effective resistant can also be defined via a random walk on an undirected graph with transition probabilities $p_{ij} = \frac{w_{ij}}{s_i}$, where s_i stands for the strength of node i.

Theorem 3.5.18 *(Chandra et al. 1989, as presented by Ellens 2011)*

Let G be a weighted, undirected graph. First, we define an electrical circuit by setting $r_{ij} = \frac{1}{w_{ij}}$. Second, we define a random walk on G with transition probabilities $p_{ij} = \frac{w_{ij}}{s_i}$ and the expected travel time between nodes i and j denoted by $\boldsymbol{E}(T_{ij})$. It holds that

$$R_{ij} = \frac{1}{\sum_{i=1}^{N} s_i} \left(\boldsymbol{E}(T_{ij}) + \boldsymbol{E}(T_{ji}) \right) \tag{3.5.22}$$

For all $i, j \in V(G)$.

Proof: Ellens 2011.

Corollary 3.5.3

$$R(G) = \frac{1}{\sum_{i=1}^{N} s_i} \sum_{i=1}^{N} \sum_{j=1}^{N} \boldsymbol{E}(T_{ij}) \tag{3.5.23}$$

Following the approach of a random walk, and by studying the expected number of visits on every node, *network criticality* was defined by Tizghadam and Leon-Garcia 2008 in a way that it equals twice the total effective resistance, and thus we did not consider it as a separate measure.

Definition 3.5.8: *Normalized total effective resistance (Ellens 2011)*

The normalized effective resistance of an undirected graph is defined as

$$R^{norm}(G) = \frac{N-1}{R(G)} = \frac{N-1}{N \sum_{i=2}^{N} \frac{1}{\lambda_i}} \in [0,1] \tag{3.5.24}$$

The benefits of such a measure are twofold. First, it's a measure of robustness instead of vulnerability, and second, its values lie in [0,1] thus making it easier to handle and to be used comparatively.

Corollary 3.5.4

$$\frac{1}{N}\lambda_2^W \leq R^{norm}(G) \leq \frac{N-1}{N}\lambda_2^W \tag{3.5.25}$$

With this modification, the algebraic connectivity can be used to approximate the robustness of a network better than by itself alone.

Theorem 3.5.19 *(Ellens 2011)*

Let G be an undirected and unweighted graph with a given maximum degree and number of nodes. Then its normalized effective resistance is bound by that achieved

within the class of clique chain graphs $G(n_1 = 1, n_2, \dots, n_{d_{max}}, n_{d_{max}+1} = 1)$ with $\sum_{i=1}^{d_{max}+1} n_i = N$.

Proof: Ellens 2011.

Corollary 3.5.5

If the number of isolated nodes of a graph is known, it can be used to further bound its normalized effective resistance by lowering the maximal one possible in its corresponding clique chain family.

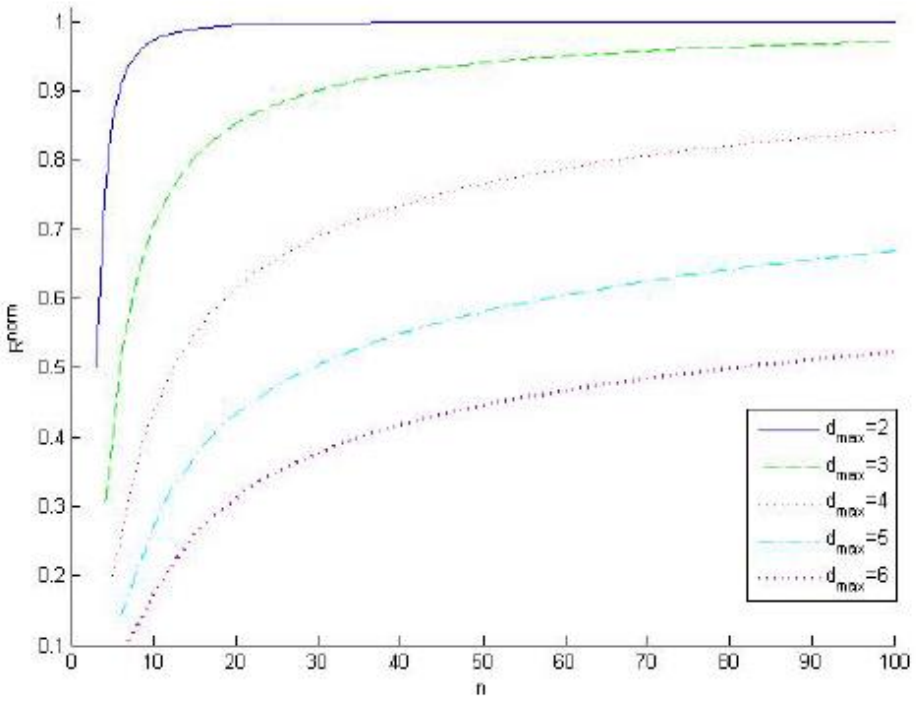

Figure 3.5.6 Bounds for the R^{norm} of the clique chains over their number of nodes, by their maximum degree.

(Ellens 2011)

Young et al. 2016a and Young et al. 2016b generalized the effective resistance for directed graphs. This attempt aims for uses in control theory and does not care for algorithms to calculate the total effective resistance, or any comparisons with other robustness measures, but it opens a large field for further research.

Deviation from the ideal good expansion character

Definition 3.5.9: *Good expander*

A network is a good expander or is said to have good expansion properties if every subset of nodes S (up to 50% of the nodes) has a neighborhood that is larger than some "expansion factor" Φ multiplied by the number of nodes in S. In other words, a network is a good expander if it has a sufficiently large vertex isoperimetric number.

Note: In general, a graph is considered to be a good expander if it has a sufficiently large vertex isoperimetric number and relatively low degree centralities.

Estrada 2006 noted that a necessary condition for a network to be a good expander is that the gap between the first and second eigenvalues of the adjacency matrix $\Delta\lambda = \lambda_1 - \lambda_2$ is sufficiently large. The problem of determining how large is "sufficiently large" is solved by assessing the degree of correlation between the largest eigenvector of the adjacency matrix (the eigencentralities) and the weighted sum of all odd-length walks that start and end at each node.

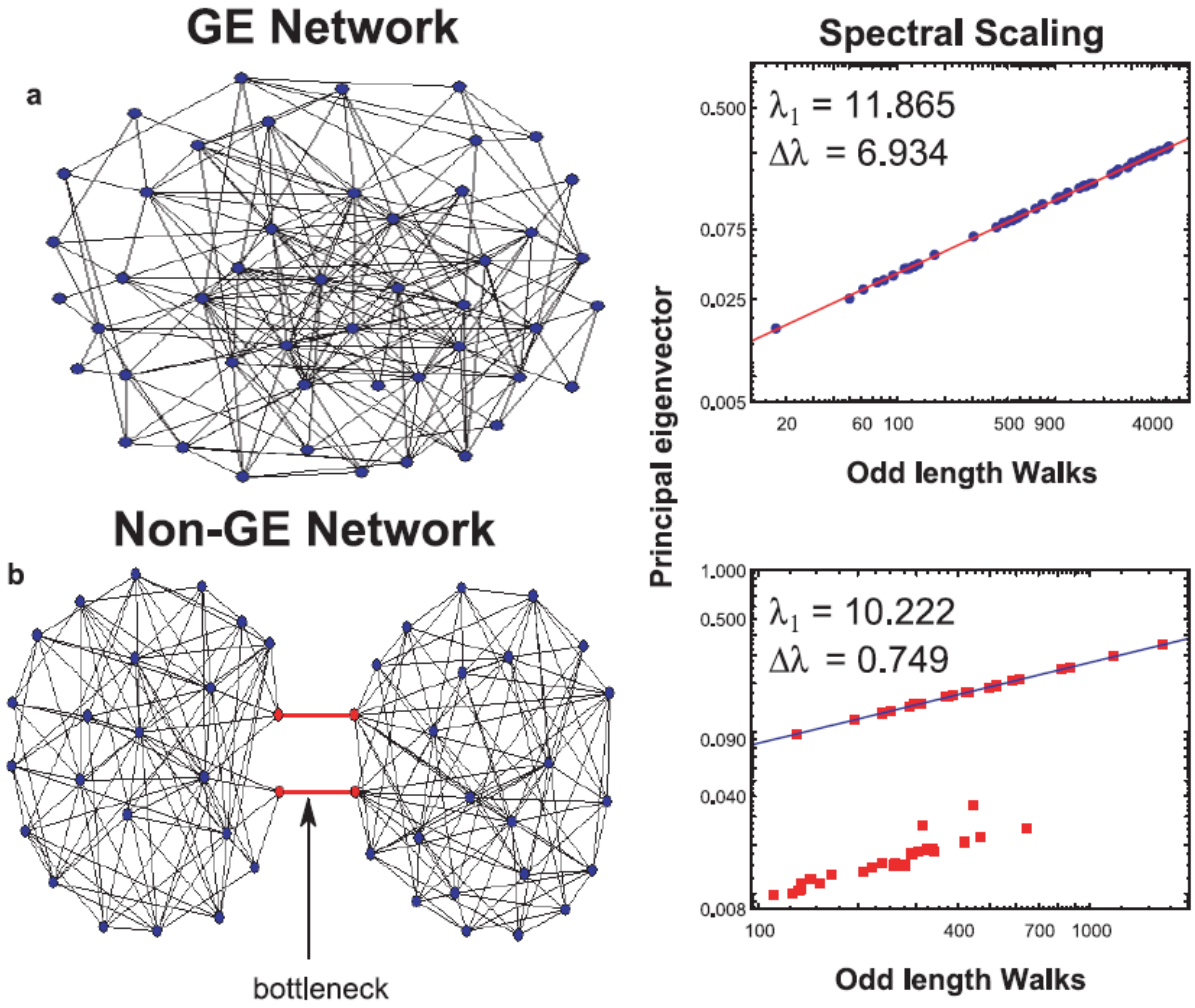

Figure 3.5.7: Illustration of the differences between entworks with and without good expansion properties. (Estrada 2006)

Definition 3.5.10: *Subgraph centrality, Odd subgraph centrality, Even subgaph centrality (Estrada and Rodriguez-Velazquez 2005)*

- Subgraph centrality is defined as the weighted by their length number of closed circles leading to a node

$$C_S(i) = \sum_{k=0}^{\infty} \frac{n_k(i)}{k!} \qquad (3.5.26)$$

Where $n_k(i)$ is the number of closed walks of length k beginning and ending at vertex i.

- We can similarly define the odd-subgraph centrality and even-subgraph centrality as the number of closed walks of odd or even length.

$$C_{S_{odd}} = \sum_{k=0}^{\infty} \frac{n_{2k+1}(i)}{(2k+1)!} \qquad C_{S_{even}} = \sum_{k=0}^{\infty} \frac{n_{2k}(i)}{(2k)!} \tag{3.5.27}$$

Theorem 3.5.20 *(Estrada and Rodriguez-Velazquez 2005)*

Let G be a simple undirected graph of order N. Let $v_1, v_2, \dots, v_N$ be an orthonormal basis of R^N composed by eigenvectors of the adjacency matrix A, associated to the eigenvalues $\lambda_1, \lambda_2, \dots \lambda_N$. Let v_j^i denote the i-th component of v_j. For all $i \in V$ the subgraph centrality may be expressed as follows:

$$C_S(i) = \sum_{j=1}^{N} v_j^i e^{\lambda_j} \tag{3.5.28}$$

Proof: Estrada and Rodrigues-Velazquez 2005.

Theorem 3.5.21 *(Rodriquez et al. 2007)*

Let G be a simple undirected graph of order N. Let $v_1, v_2, \dots, v_N$ be an orthonormal basis of R^N composed by eigenvectors of the adjacency matrix A, associated to the eigenvalues $\lambda_1, \lambda_2, \dots \lambda_N$. Let v_j^i denote the i-th component of v_j. For all $i \in V$ the odd and even subgraph centrality may be expressed as follows:

$$C_{S_{odd}} = \sum_{j=1}^{N} \left(v_j^i\right)^2 \sinh(\lambda_j) \qquad C_{S_{even}} = \sum_{j=1}^{N} \left(v_j^i\right)^2 \cosh(\lambda_j) \tag{3.5.29}$$

Proof: Rodriquez et al. 2007.

Definition 3.5.11: *Deviation from the good expansion character (Estrada 2006)*

Deviation from the good expansion character of a simple undirected graph is defined in such a way that perfect good expander networks will have $\xi(G) = 0$, specifically

$$\xi(G) = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\{\log[v_1(i)] - [\log(\alpha) + \eta \log[SC_{odd}(i)]]\}^2} \quad (3.5.30)$$

Where $\alpha \approx [\sinh(\lambda_1)]^{-0.5}, \eta \approx 0.5$ and v_1 is the eigenvector for the largest eigenvalue λ_1.

The above values are derived from the properties of good expander graphs by Estrada 2006.

The deviation from the good expansion character is shown to be a measure of vulnerability of a network by Estrada 2006 and Estrada 2007.

3.6. Combined Robustness

The idea of measuring a networks robustness in general has been studied by Van Mieghem et al. 2010, where a linear model for defining robustness is proposed as such

Definition 3.6.1: *Combined robustness (Van Mieghem et al. 2010)*

The combined robustness of a network is defined as

$$R = \sum_{k=1}^{m} s_k R_k \quad (3.6.1)$$

Where R_k are measures of robustness, or other network topological measures, preferably normalized, and s_k is a variable defining the importance of the corresponding measure of each measure of robustness or topological measure.

It is also noted, that such measures are frequently correlated and should be chosen in such a way that they are as independent as possible, but such a choice depends on the network being studied.

"The dependence between metrics in a graph seems a hard, inherent challenge of the robustness problem" (Van Mieghem et al. 2010).

It would, therefore, be an interesting point of research to study the correlation of the various measures according to the topology of a graph. Such correlations have, only partly, been explored by Kasthurinathna et al. 2013.

4. Network deconstruction

4.1. Introduction

In this chapter we will define the various methods of attack that have been used. We will present selected methods of attack with respect to robustness according to:

- The networks that have been studied, providing:
  - Structural information
  - Information on direction and/or weights
  - Network functionalities
- The origin of the network
- The measure used to assess the impact of each attack
- The relative reference

In this chapter some of the strategies of removing nodes from a network, are defined and presented according to the networks they have been tested on, the indices of robustness that have been used and finally the reference on where each study can be found.

Definition 4.1.1: *Instantaneous/Salvo attacks, Repeated attacks, Cascading attacks*

- Instantaneous or salvo attack on a network is multiple modifications performed at the same time in the network with no difference taking place in the network in-between them.
- Repeated attack on a network is any sequence of modifications where the mechanisms of the network, such as its limitations any self-rewiring capabilities, occur in-between the attacks.
- Cascading attack on a network is a special case of repeated attacks, or repeated salvos that start on a node or link and radiate from it. For example, the removal of a node

causes its neighbors to be removed as well in the next instance, either one by one or all together.

It is important to note that in the case of no limitations, self-rewiring capabilities or other such mechanics existing in a network, and in the case of a single node or link attack, there is no difference in the effects of repeated or instantaneous removals.

4.2. Random removals

We have already defined the notion of random removals of nodes or links in a network (definition 1.2.4). They have been widely studied, since almost any network is designed with the possibility of parts of it failing randomly, because such mishaps are present in any undertaking, e.g. transport networks and harsh weather, road networks and traffic accidents, power grids and distribution station malfunctions etc. In general, random removals are studied as a method of attack, either to check the ability of a network to cope with accidents, or as a baseline for other methods of attack on a network.

Table 4.2.1: Repeated, random node removals ordered by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Single value degree distribution (all degrees equal to 3) Undirected and unweighted Processability limitations	Artificial	Change in the order of the giant component	Motter and Lai 2002 Removal of a single node
2-Peak and 3-Peak degree distribution Undirected and unweighted	Artificial	Critical Threshold	Valente et al. 2004
Exponential degree distribution	Artificial	Change in the order of the giant component	Newman and Ghosal 2007

Undirected and unweighted		Change in the order of the giant bicomponent	Newman and Ghosal 2007
Exponential degree distribution Undirected and unweighted Processability limitations	North American Power Grid	Percentage of nodes disconnected from the giant component	Albert et al. 2004
	US power grid	Change in the order of the giant component	Motter and Lai 2002 Removal of a single node
Erdos-Renyi Undirected and unweighted	Artificial	Critical Threshold	Albert et al. 2000
		Change in the diameter	Albert et al. 2000
		Change in the efficiency	Crucitti et al. 2003
		Change in the local efficiency	Crucitti et al. 2003
		Change in the order of the giant component	Albert et al. 2000
Erdos Renyi Undirected and unweighted Processability limitations	Artificial	Change in the order of the giant component	Motter and Lai 2002 Removal of a single node
Erdos Renyi Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
2-dimentional exponential degree distribution Undirected and unweighted	Ant Galleries, their minimal spanning trees and their triangulated graphs	Fragmentation Threshold ($a = 0.5$)	Buhl et al. 2004
		Change in the efficiency	Buhl et al. 2004
		Change in the order of the giant component	Buhl et al. 2004
2-dimentional exponential degree distribution Undirected and weighted	Ant Galleries, their minimal spanning trees and their triangulated graphs	Fragmentation Threshold ($a = 0.5$)	Buhl et al. 2004
		Change in the efficiency	Buhl et al. 2004
		Change in the order of the giant component	Buhl et al. 2004
Poisson degree distribution Undirected and unweighted	Artificial	Change in the order of the giant component	Newman and Ghosal 2007
		Change in the order of the giant bicomponent	Newman and Ghosal 2007
Scale Free Undirected and unweighted	Artificial	Critical Threshold	Albert et al. 2000 Cohen et al. 2000
		Change in the order of the giant component	Albert et al. 2000 Cohen et al. 2000
		Change in the diameter	Albert et al. 2000
		Change in the efficiency	Crucitti et al. 2003

		Change in the local efficiency	Crucitti et al. 2003
Scale Free Undirected and unweighted Processability limitations	Artificial	Change in the order of the giant component	Motter and Lai 2002 Removal of a single node
Scale Free Undirected and unweighted Self-rewiring	Cannabis distribution network	Change in the efficiency	Duijn et al. 2014
		Change in the network density	Duijn et al. 2014
Scale Free Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Small World Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Unspecified structure Undirected and unweighted	Shanghai Subway	Change in the efficiency	Zhang et al. 2011
		Functionality loss	Zhang et al. 2011
		Change in the order of the giant component	Zhang et al. 2011
	Various real-world networks	Change in the order of the giant bicomponent	Newman and Ghosal 2007
	Western Power Grid of the US	JS-robustness and its change over the attack	Schieber et al. 2015
	Various artificial and real	Compensated total graph diversity	Rohrer and Sterbenz 2011
Unspecified structure Undirected and weighted	Beijing Subway	Change in the efficiency	Yin et al. 2016
		Change in the local efficiency	Yin et al. 2016
	Dolphin Social Network	JS-robustness and its change over the attack	Schieber et al. 2015
Unspecified structure Directed and unweighted Structural limitations	Food Webs	Fragmentation Threshold ($a = 0.5$)	Dunne et al. 2002
		Secondary extinctions and extinction area	Dunne et al. 2002
		Network density	Dunne et al. 2002
Unspecified structure Directed and weighted	Food Webs	Change in the order of the giant component	Sole and Montoya 2001

Bipartite with unspecified structure Undirected and unweighted Structural limitations	Pollinator networks	Fragmentation Threshold ($a \simeq 0.5$)	Santamaria et al. 2014
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Pollinator networks	% of nodes of the opposite type remaining	Kaiser-Bundury et al. 2010

Table 4.2.2: Repeated random edge removals ordered by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Erdos Renyi Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Scale Free Undirected and unweighted	Artificial	Natural Connectivity	Wu et al. 2010 Wu et al. 2011
		Change in the order of the giant component	Wu et al. 2010 Wu et al. 2011
	Chinese Internet	Natural Connectivity	Wu et al. 2011
		Change in the order of the giant component	Wu et al. 2011
Scale Free Undirected and unweighted Processability limitations	Artificial	Change in the probability the largest component has an order of the network's order of magnitude	Moreno et al. 2003 Removal of a single edge
Scale Free Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Small World Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Unspecified structure Undirected and unweighted	Western Power Grid of the US	JS-robustness and its change over the attack	Schieber et al. 2015
	Various artificial and real	Compensated total graph diversity	Rohrer and Sterbenz 2011
Unspecified structure	Beijing Subway	Change in the efficiency	Yin et al. 2016
		Change in the local efficiency	Yin et al. 2016

Undirected and weighted	Dolphin Social Network	JS-robustness and its change over the attack	Schieber et al. 2015

Table 4.2.3: Repeated random mixed node and link removals ordered by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Unspecified structure Undirected and unweighted	Various artificial and real	Compensated total graph diversity	Rohrer and Sterbenz 2011

4.3. Ranking

By ranking all the nodes or links in a network according to some measure that indicates their importance to the structure, we can get a sequence of nodes or links to remove that should be more effective according to the assumption that more important nodes play a larger role on the structure of a network.

A very important factor in such attack methods is whether these measures are calculated only on the initial network, or on every instance of a repeated removal attack (i.e. after the removal of each node). Calculating each instance separately can be more informative, as each removal targets the most important node or link, according to the selected ranking. However, it can also be very taxing to re-estimate such values in larger networks for every instance, and sometimes inefficient, as for example, the degree distribution of the network does not change significantly after each attack, although it has been shown to be more efficient (Holme et al. 2002, Chen et al. 2008, Morone and Makse 2015, Morone et al. 2016).

Definition 4.3.1: *Initial measure attacks, adaptive attacks*

An attack that targets edges or nodes according to a network measure such as the degree, can be formed either according to the initial values of the network, or to the

ones adapted to the new network created after the previous attacks. We call attacks that target nodes or edges according to the original ranking initial measure attacks (for example initial degree attacks) and attacks that target the recalculated values adaptive measure attacks.

Targeted attacks will be referred to in a format following the order, measure and target of the attack. For example, high degree node attacks target the nodes of the network in according to their degree in decreasing order. Furthermore, when a removal strategy is based on the measure re-estimated at every step, this will be noted.

For clarification purposes, let it be noted that when referring to a measure (e.g. the degree) that has a corresponding centrality, the centrality is considered to be the measure divided by the largest possible value for a node or edge in a graph of the same order (as defined in Freeman 1979)

Degree based node attacks

The most common, and perhaps most intuitive method studied in this regard is the degree centrality of each node.

Definition 4.3.2: *Degree, In-degree, Out-degree*

The degree of a node is defined as the number of neighbors it is connected to. In directed graphs this can be further analyzed to the in-degree and out-degree that are defined as the number of incoming and outgoing links respectively.

Removing the most connected node of the network should logically correspond to the largest possible disconnection.

Table 4.3.1: Repeated attacks on the nodes with the highest degree, by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Exponential degree distribution Undirected and unweighted Processability limitations	Artificial	Change in the order of the giant component	Motter and Lai 2002 Removal of a single node
2-Peak and 3-Peak Undirected and unweighted	Artificial	Critical Threshold	Valente et al. 2004
Uniform Undirected and Unweighted	Artificial and various real-world	Change in the order of the giant component	Estrada 2006
	Food Webs	Change in the order of the giant component	Estrada 2007
Erdos Renyi Undirected and unweighted	Artificial	Critical Threshold	Albert et al. 2000
		Change in the efficiency	Holme et al. 2002 Initial and adaptive Crucitti et al. 2003
		Average order of the giant component	Schneider et al. 2010 Adaptive only
		Change in the diameter	Albert et al. 2000
		Change in the local efficiency	Crucitti et al. 2003
		Change in the order of the giant component	Albert et al. 2000 Holme et al. 2002 Initial and adaptive Chen et al. 2008 Morone and Makse 2015 Initial and adaptive Morone et al. 2016 Adaptive only
Erdos Renyi (Q-augmented) Undirected and unweighted	Artificial	Change in the order of the giant component	Costa 2004
Erdos Renyi Directed and weighted	Artificial	AUC of the order of the giant component	Wagner 2015
Random Regular	Artificial	Change in the order of the giant component	Chen et al. 2008 Initial and adaptive

Undirected and unweighted		AUC of the order of the giant component	Morone et al. 2016 Adaptive only
Random with exponential degree distribution Undirected and unweighted	Artificial	Change in the order of the giant component	Iyer et al. 2013 Initial and adaptive
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
Exponential degree distribution Undirected and Unweighted	Artificial and various real-world	Change in the order of the giant component	Estrada 2006
	Food Webs	Change in the order of the giant component	Estrada 2007
Exponential degree distribution Undirected and unweighted Processability limitations	US power grid	Change in the order of the giant component	Motter and Lai 2002 Removal of a single node
	North American Power Grid	% of disconnected nodes	Albert et al. 2004
2-dimentional with exponential degree distribution Undirected and unweighted	Ant Galleries, their minimal spanning trees and their triangulated graphs	Change in the order of the giant component	Buhl et al. 2004
		Fragmentation Threshold ($a = 0.5$)	Buhl et al. 2004
		Change in the efficiency	Buhl et al. 2004
2-dimentional with exponential degree distribution Undirected and weighted	Ant Galleries, their minimal spanning trees and their triangulated graphs	Change in the order of the giant component	Buhl et al. 2004
		Fragmentation Threshold ($a = 0.5$)	Buhl et al. 2004
		Change in the efficiency	Buhl et al. 2004
Scale Free Undirected and unweighted	Artificial	Critical Threshold	Albert et al. 2000 Cohen et al. 2001
		Average order of the giant component	Schneider et al. 2010 Adaptive only
		Change in the efficiency	Holme et al. 2002 Initial and adaptive Crucitti et al. 2003
		Change in the diameter	Albert et al. 2000
		Change in the local efficiency	Crucitti et al. 2003
		Change in the order of the giant component	Albert et al. 2000 Cohen et al. 2000 Holme et al. 2002 Initial and adaptive Chen et al. 2008 Initial and adaptive

			Iyer et al. 2013 Initial and adaptive Morone and Makse 2015 Initial and adaptive
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
		SSSK-elasticity	Sydney et al. 2008
	European Electricity system (power grid) Internet	Average order of the giant component	Schneider et al. 2010 Adaptive only
Scale Free (Q-augmented) Undirected and unweighted	Artificial	Change in the order of the giant component	Costa 2004
Scale Free Undirected and unweighted Processability limitations	Artificial	Change in the order of the giant component	Motter and Lai 2002 Removal of a single node
Scale Free Undirected and unweighted Self-rewiring	Cannabis distribution network	Change in the efficiency	Duijn et al. 2014
		Change in the network density	Duijn et al. 2014
Scale Free Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015 Kasthrinathna and Mahendra 2013
Scale Free Undirected and unweighted	Food Webs	Change in the order of the giant component	Estrada 2007
	Artificial and various real-world	Change in the order of the giant component	Estrada 2006
Clustered Scale Free Undirected and unweighted	Artificial	Change in the efficiency	Holme et al. 2002 Initial and adaptive
		Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
Scale-Free (power law with exponential tail degree distribution) Undirected and unweighted	Food Webs	Change in the order of the giant component	Estrada 2007
	Artificial and various real-world	Change in the order of the giant component	Estrada 2006
Small World Undirected and unweighted	Artificial	Change in the efficiency	Holme et al. 2002 Initial and adaptive
		Change in the order of the giant component	Holme et al. 2002 Initial and adaptive

Small World Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Unspecified structure Undirected and unweighted	Mexican telecommunications network Twitter network	Change in the order of the giant component	Morone and Makse 2015 Initial and adaptive
Bipartite with unspecified structure Undirected and unweighted	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017
Unspecified structure Undirected and weighted	Metro networks (Beijing, Shanghai, Guangzhou)	Change in the efficiency	Zhang et al. 2018
		Functionality loss	Zhang et al. 2018
	Shanghai Subway	Connectivity of a line	Zhang et al. 2011
		Change in the efficiency	Zhang et al. 2011
		Functionality loss	Zhang et al. 2011
		Change in the order of the giant component	Zhang et al. 2011
	Colaboration Network Computer Network from internet traffic	Change in the efficiency	Holme et al. 2002 Initial and adaptive
		Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
	Beijing Subway	Change in the efficiency	Yin et al. 2016
		Change in the local efficiency	Yin et al. 2016
	Global Salafi Jihad terrorist network (provided by third party) Meth World, Gang-related criminals (Xu and Chen 2003) Terrorist web site network (created by the authors, government data sources)	Change in the order of the giant component	Xu and Chen 2008
		Critical Threshold	Xu and Chen 2008
	Various Real-World	Change in the order of the giant component	Iyer et al. 2013 Initial and adaptive
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
	Workplace network Internet (Autonomous System)	Change in the order of the giant component	Chen et al. 2008 Initial and adaptive

	High energy particle physics station network Metabolic network		
Unspecified structure Directed and Unweighted	Food Webs	Change in the order of the giant component	Sole and Montoya 2001
Unspecified structure Directed and weighted Structural limitations	Food Webs	Extinction Area	Allesina and Pascual 2009
		Fragmentation Threshold ($a = 0.5$)	Dunne et al. 2002*
		Secondary extinctions and extinction area	Dunne et al. 2002*
		Network density	Dunne et al. 2002*
Bipartite with unspecified structure Undirected and weighted	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Pollinator networks	Fragmentation Threshold ($a = 0.5$)	Santamaria et al. 2014
		Secondary extinctions and extinction area	Memmott et al. 2004
	Various real world and artificial	Extinction area of the second nodal group	Dominguez-Garcia and Munoz 2015
Bipartite with unspecified structure Undirected and unweighted Structural limitations Self-rewiring	Pollinator networks	Percentage of nodes of the opposite group remaining	Kaiser-Bundury et al. 2010
HOT for SSSK-elasticity (Heuristically Optimal Topology) (Sydney et al. 2008) Undirected and unweighted	Artificial	SSSK-elasticity	Sydney et al. 2008

*In the study by Dunne et al. 2002, the attack was run twice. The second time, basal species (species with predators but no prey) were not included in the ranking.

Table 4.3.2: Repeated attacks on the nodes with the lowest degree, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Pollinator networks	Fragmentation Threshold ($a = 0.5$)	Santamaria et al. 2014
		Secondary extinctions and extinction area	Memmott et al. 2004
Bipartite with unspecified structure Undirected and unweighted Structural limitations Self-rewiring	Pollinator networks	Percentage of nodes of the opposite type remaining	Kaiser-Bundury et al. 2010

Edge attacks based on the degrees of the neighboring vertices

Holme et al.2002 ranked both vertex and edge attacks under similar terms. The measure for edge ranking corresponding to the vertex degree that was proposed is the edge-degree.

Definition 4.3.3: *Edge-degree (Holme et al. 2002)*

The edge-degree of an edge is defined as the product of the vertex degree of the vertices it connects.

Additionaly, Wu et al. 2010 and Wu et al. 2011 studied the edge attacks based on the possible combinations of the vertices they connect. Specifically, three such attacks were studied;

- Rich-Rich, meaning edges that connect high degree nodes with high degree nodes are removed first, equating to descending edge degree.

- Rich-Poor, meaning edges that connect high degree nodes with low degree nodes are removed first, edges in descending order of the difference of the degrees of the nodes they connect $|d(v) - d(u)|$.
- Poor-Poor, meaning edges that connect low degree nodes with low degree nodes are removed first, equating to ascending edge degree.

Table 4.3.3: Repeated attacks on the edges with the highest edge-degree, by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Scale Free Undirected and unweighted	Artificial	Change in the efficiency	Holme et al. 2002 Initial and adaptive
		Natural Connectivity	Wu et al. 2010 Unspecified Wu et al. 2011 Unspecified
	Chinese internet	Natural Connectivity	Wu et al. 2011 Unspecified
Erdos Renyi Undirected and unweighted	Artificial	Change in the efficiency	Holme et al. 2002 Initial and adaptive
Small World Undirected and unweighted	Artificial	Change in the efficiency	Holme et al. 2002 Initial and adaptive
Clustered Scale Free Undirected and unweighted	Artificial	Change in the efficiency	Holme et al. 2002 Initial and adaptive
Unspecified structure Undirected and unweighted	Colaboration Network Computer Network from Internet traffic	Change in the efficiency	Holme et al. 2002 Initial and adaptive

Table 4.3.4: Repeated attacks on the edges connecting rich to poor or poor to poor nodes, by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Scale Free Undirected and unweighted	Artificial	Natural Connectivity	Wu et al. 2010 Unspecified Wu et al. 2011 Unspecified
	Chinese internet	Natural Connectivity	Wu et al. 2011 Unspecified

Betweenness based vertex and edge attacks

A notion for the importance of a vertex can be the amount of information that flows through it, and a very good index for it is the number of shortest paths that flow through it, since that's the most logical way for information routing.

Definition 4.3.4: *Betweenness*

The betweenness of a vertex or an edge is defined as the number of shortest paths in the network that it is included in.

It is a rational thought that removing the most well-traversed vertex or edge in a network will cause a large perturbation in its information flow. Intuitively, one only has to think of the traffic jams occurring when a central rode in a city closes down. This thought becomes even more appealing when we consider that a bottleneck (a small set of nodes connecting to large parts of the network) will always have high betweenness, and thus will be found faster via this method, increasing the chance to create large disconnections.

Table 4.3.5: Repeated attacks on the nodes with the highest betweenness, by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Erdos Renyi Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive Chen et al. 2008 Morone and Makse 2015
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Erdos Renyi Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Random Regular Undirected and unweighted	Artificial	Change in the order of the giant component	Chen et al. 2008
Exponential Undirected and unweighted	Artificial and various real-world	Change in the order of the giant component	Estrada 2006
	Artificial	Change in the order of the giant component	Iyer et al. 2013 Initial and adaptive
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive Chen et al. 2008 Iyer et al. 2013 Morone and Makse 2015
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Scale Free Undirected and Unweighted Self-rewiring	Cannabis distribution network	Change in the efficiency	Duijn et al. 2014
		Change in the network density	Duijn et al. 2014
Scale Free Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Clustered Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Scale-Free (power law with exponential tail degree distribution)	Artificial and various real-world	Change in the order of the giant component	Estrada 2006

Undirected and Unweigted			
Small World Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Small World Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Uniform degree distribution Undirected and unweighted	Artificial and various real-world	Change in the order of the giant component	Estrada 2006
Unspecified structure Undirected and unweighted	Colaboration Network Computer Network from Internet traffic	Change in the efficiency	Holme et al. 2002 Initial and adaptive
		Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
	Shanghai Subway	Change in the efficiency	Zhang et al. 2011
		Functionality loss	Zhang et al. 2011
		Change in the order of the giant component	Zhang et al. 2011
	Metro networks (Beijing, Shanghai, Guangzhou)	Change in the efficiency	Zhang et al. 2018
		Functionality loss	Zhang et al. 2018
	Mexican telecommunications network Twitter network	Change in the order of the giant component	Morone and Makse 2015
	Workplace network Internet (Autonomous System) High energy particle physics station network Metabolic network	Change in the order of the giant component	Chen et al. 2008
	Various real world	Change in the order of the giant component	Iyer et al. 2013 Initial and adaptive
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
Unspecified structure Undirected and weighted	Beijing Subway	Change in the efficiency	Yin et al. 2016
		Change in the local efficiency	Yin et al. 2016
Unspecified structure Directed and weighted	Food Webs	Extinction Area	Allesina and Pascual 2009

Structural limitations			
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Various real world and artificial	Extinction area of the second nodal group	Dominguez-Garcia and Munoz 2015

Table 4.3.6: Repeated attacks on the edges with the highest betweenness, by network structure, measure used and network type.

Network Structure	Network Origin	Measure of assessment	Reference
Erdos Renyi Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Erdos Renyi Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Scale Free Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Clustered Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Small World Undirected and unweighted	Artificial	Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
		Change in the efficiency	Holme et al. 2002 Initial and adaptive
Small World Directed and unweighted	Artificial	AUC of the order of the giant component	Wagner 2015
Unspecified structure Undirected and unweighted	Colaboration Network Computer Network from Internet traffic	Change in the efficiency	Holme et al. 2002 Initial and adaptive
		Change in the order of the giant component	Holme et al. 2002 Initial and adaptive
Unspecified structure Undirected and weighted	Beijing Subway	Change in the efficiency	Yin et al. 2016
		Change in the local efficiency	Yin et al. 2016

Eigencentrality (and other similar measures) based vertex attacks

Eigencentrality as a centrality measure has evolved as a more elaborate measure to assess the importance of a node not only by how *many* neighbors he has but by how *important* its neighbors are.

Definition 4.3.5: *Eigenvector centrality/Eigencentrality (Bonacich 1987)*

For a given graph, with an adjacency matrix $A = (a_{i,j})$ the *eigencentrality* or *eigenvector centrality* is defined as

$$C_{eig}(v) = \frac{1}{\lambda} \sum_{j \in V(G)} a_{v,j} x_j \qquad (4.3.1)$$

In general, there will be many different eigenvalues λ for which a non-zero eigenvector solution exists. However, the additional requirement that all the entries in the eigenvector be non-negative implies (by the Perron–Frobenius theorem) that only the greatest eigenvalue results in the desired centrality measure.

Table 4.3.7: Repeated attacks on the nodes with the highest eigencentrality, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Erdos Renyi Undirected and unweighted	Artificial	Change in the order of the giant component	Morone and Makse 2015
Exponential Undirected and unweighted	Artificial	Change in the order of the giant component	Iyer et al. 2013
		AUC of the order of the giant component	Iyer et al. 2013
Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Iyer et al. 2013 Morone and Makse 2015
		AUC of the order of the giant component	Iyer et al. 2013
Unspecified structure Undirected and unweighted	Mexican telecommunications network	Change in the order of the giant component	Morone and Makse 2015

	Twitter network		
	Various real world	Change in the order of the giant component	Iyer et al. 2013
		AUC of the order of the giant component	Iyer et al. 2013
Bipartite Unspecified structure Undirected and unweighted Structural limitations	Various real world and artificial	Extinction area of the second nodal group	Dominguez-Garcia and Munoz 2015
	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017
Bipartite Unspecified structure Undirected and weighted Structural limitations	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017

Definition 4.3.6: *PageRank*

$$PR(i) = \frac{1-d}{N} + d \sum_{i \in N_{to}(i)} \frac{PR(i)}{E_{out}(i)} \qquad (4.3.2)$$

Where $0 < d < 1$ is a dampening factor, N is the number of nodes, $N_{to}(i)$ are the nodes that point to i and $E_{out}(i)$ is the number of outgoing edges of node i.

Table 4.3.8: Repeated attacks on the nodes with the highest PageRank, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Erdos Renyi Undirected and unweighted	Artificial	Change in the order of the giant component	Morone and Makse 2015
Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Morone and Makse 2015

Unspecified structure Undirected and unweighted	Mexican telecommunications network Twitter network	Change in the order of the giant component	Morone and Makse 2015
Unspecified structure Directed and weighted Structural limitations	Food Webs	Extinction Area	Allesina and Pascual 2009 (out-pagerank)
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Various real world and artificial	Extinction area of the second nodal group	Dominguez-Garcia and Munoz 2015

*Note that in the study by Allesina and Pascual 2009, the PageRank algorithm was modified to consider important the outgoing instead of the incoming connections, meaning that N_{to} and E_{out} are replaced with N_{from} and E_{in}.

Highest collective influence vertex attacks

The collective influence of a node is a measure designed to find important actors in fake news spreading. These models usually account for probabilistic transitions and do not generally study complete disconnections, thus they are not directly relevant to this work. Nevertheless, there have been some simulations related to our work that seem to imply that the collective influence might be an excellent ranking method for our purposes.

Definition 4.3.7: *Collective influence*

Collective influence of a node v at l steps is defined as follows.

$$CI_l(v) = (C_d(v) - 1)\sum_{u \in S}(C_d(u) - 1), \qquad S = \{u \in V(G) | d(v,u) = l\} \qquad (4.3.3)$$

Where $d(v,u)$ is the distance between v and u

Table 4.3.9: Repeated attacks on the nodes with the highest collective influence, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Erdos Renyi Undirected and unweighted	Artificial	Change in the order of the giant component	Morone and Makse 2015 Adaptive only Morone et al. 2016 Adaptive only
Random Regular Undirected and Unweighted	Artificial	AUC of the order of the giant component	Morone et al. 2016 Adaptive only
Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Morone and Makse 2015 Adaptive only
Unspecified structure Undirected and unweighted	Mexican telecommunications network Twitter network	Change in the order of the giant component	Morone and Makse 2015 Adaptive only

Highest MusRank vertex attacks

Dominguez-Garcia and Munoz 2015 proposed a method of measuring and ranking the complexity of products and the fitness of countries to produce them, proposed by Tacchella et al. 2012, was adapted for biological bipartite networks by considering plants, seeds and anemones as products and pollinators, dispersers and fishes. The algorithm was also tested by reversing those roles.

Definition 4.3.8: *MusRank (Tacchella et al. 2012, modified by Dominguez-Garcia and Munoz 2015)*

Let A_{cp} be the binary country-product matrix with elements 1 if country c produces product p and 0 otherwise. Let also F_c symbolize the fitness of a country and Q_p the complexity of a product. The iterative method of calculating these values starts by assigning them values of $\tilde{F}_c^{(0)} = \tilde{Q}_p^{(0)} = 1$, and then calculating the complexity of each product by the inverse of the sum of the number of countries exporting said product

weighted by their fitness, and the fitness of each country by the number of products they export weighted by their complexity.

$$\begin{cases} \widetilde{F}_c^{(n)} = \sum_p A_{cp} Q_p^{(n-1)} \\ \tilde{Q}_p^{(n)} = \dfrac{1}{\sum_c A_{cp} \dfrac{1}{F_c^{(n-1)}}} \end{cases} \quad (4.3.4)$$

We call MusRank (Mutualistic Species Rank) the application of this complexity index to mutualistic passive-active ecosystem networks. For example, plant-pollinator networks.

Products are replaced by plants, seeds, anemones etc. (passive) while countries are replaced by pollinators, birds, fish etc. (active), although research has been done with these roles (passive-active) reversed as well.

Table 4.3.10: Repeated attacks on the nodes with the highest MusRank, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Various real world and artificial	Extinction area of the second nodal group	Dominguez-Garcia and Munoz 2015
	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017
Bipartite with unspecified structure Undirected and weighted Structural limitations	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017

It is noteworthy, that the MusRank ranking, provides a near-optimal attack strategy in such networks.

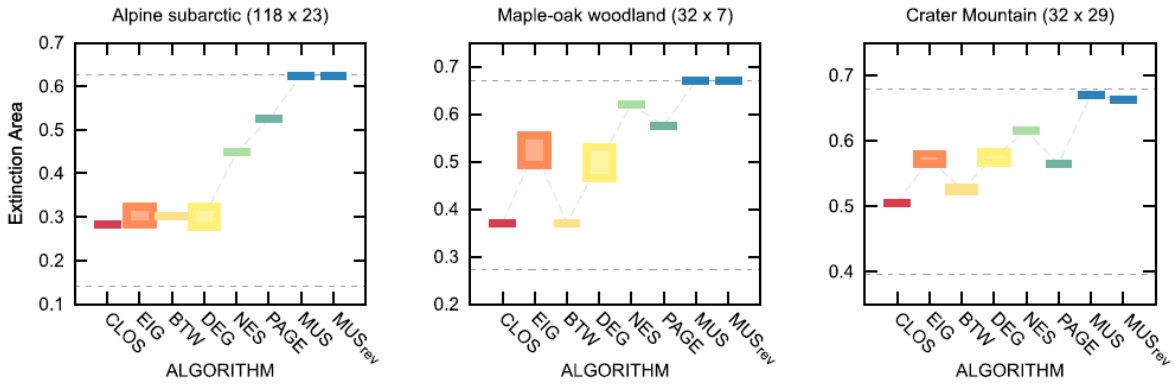

Figure 4.3.1: Extinction areas for three different mutualistic networks as obtained employing the different ranking schemes. The upper dashed line shows the optimal performance according to a genetic algorithm, the lower dashed line shows the null-expectation, that is the averaged area obtained when targeting nodes in random order. The different algorithms are respectively: Closeness, Eigencentrality, Betweenness, Degree, Nestedness, PageRank, MusRank and reversed MusRank (meaning the active-passive indexing is reversed).

(Domingues-Garcia and Munoz 2015)

Highest closeness centrality attacks

Definition 4.3.9: *Closeness centrality*

In a connected graph, *closeness centrality* or *closeness* of a node is defined as the inverse of the sum of the length of the shortest paths between the node and all other nodes, multiplied by the number of nodes in order to be defined in [0,1].

$$C_c(v) = \frac{|V(G)|}{\sum_{x \in V(G)} d(v,x)} \tag{4.3.5}$$

Removing the nodes with the highest closeness centrality leads to larger travel times for information in the network, since in theory it is the optimal spreading point. These nodes are considered very important in news and disease spreading models, but this is out of the scope of this work. Nevertheless, some studies concerning the capacity of an attack method to disconnect the network have been conducting and are presented below.

Table 4.3.11: Repeated attacks on the nodes with the highest closeness centrality, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Exponential degree distribution Undirected and unweighted	Artificial	Change in the order of the giant component	Iyer et al. 2013 Initial and adaptive
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
Erdos Renyi Undirected and unweighted	Artificial	Change in the order of the giant component	Morone and Makse 2015
Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Iyer et al. 2013 Initial and adaptive Morone and Makse 2015
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
Unspecified structure Undirected and unweighted	Various real world	Change in the order of the giant component	Iyer et al. 2013 Initial and adaptive
		AUC of the order of the giant component	Iyer et al. 2013 Initial and adaptive
Unspecified structure Directed and weighted Structural limitations	Food Webs	Extinction Area	Allesina and Pascual 2009
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Various real world and artificial	Extinction area of the second nodal group	Dominguez-Garcia and Munoz 2015

Attacks based on number of dominations

Definition 4.3.10: *Food web*

A food web is a network created by ordaining creatures or sets of creatures from an environment as nodes and connecting them with directed edges representing the relationship A is eaten or consumed by B as $A \rightarrow B$.

Definition 4.3.11: *Root of a food web (Allesina and Pascual 2009)*

The root in a food web is an artificially induced node that represents the environment and has outgoing links towards all primary producers. Also, every species has an intrinsic loss of matter which is represented as an edge connecting it to the root.

Definition 4.3.12: *Dominating node (Allesina and Pascual 2009)*

A node x is said to dominate node y (x is a dominating node of y) if all the paths from the root towards y pass through x.

Allesina and Pascual 2009 also studied the removal of the node that dominates the most nodes in food webs.

Table 4.3.12: Repeated attacks on the nodes that dominate the highest number of nodes, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Unspecified structure Directed and weighted Structural limitations	Food Webs	Extinction Area	Allesina and Pascual 2009

Lowest nestedness vertex attacks

A bipartite network is said to be nested when the nodes of group A that are connected to a few nodes of group B (locations with few species, species with few interactions) have a subset of the group A nodes with more connections to group B.

Definition 4.3.13: *Nestedness (Bastolla et al. 2009)*

(A measure of) *Nestedness* of a network is defined locally for each nodal pair as

$$\eta_{ij} = \frac{(A^2)_{ij}}{d(i)\, d(j)} \tag{4.3.6}$$

With A being the adjacency matrix, and $d(v)$ the degree of node v.

The same measure is defined per node as a centrality measure

$$\eta(j) = \frac{1}{N} \sum_{i \in V(G)} \eta_{ij} \tag{4.3.7}$$

And globally as the average

$$\eta(G) = \frac{1}{N^2} \sum_{i,j \in V(G)} \eta_{ij} \tag{4.3.8}$$

Table 4.3.13: Repeated attacks on the nodes with the lowest nestedness, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Various real world and artificial	Extinction area of the second nodal group	Dominguez-Garcia and Munoz 2015

Attacks based on the flow that goes through each node

Albert et al. 2004 studied the removal of the nodes, in the north American power grid, that the largest amount of electricity flow passes through, both based on the initial values, and by recalculating the flow in the network every 10 removals. In their model, any node above his capacity is considered to fail.

Table 4.3.14: Repeated attacks on the nodes through which the largest amount of flow passes through, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Exponential degree distribution Undirected and unweighted Processability and structural limitations	North American Power Grid	% of disconnected nodes	Albert et al. 2004 Initial and adaptive every 10 iterations

Vertex removals based on the edges that connect to or from them

Highest bottleneck ratio edges

Wagner 2015 proposed bottleneck ratio as a measure to assess the importance of an edge in a flow network by the traffic that flows through it. Two different attacks were then proposed, the first was ranking the nodes by the average bottleneck ratio of their edges, and the second was ranking the nodes by the sum of their edges' bottleneck ratio.

Definition 4.3.14: *Minimum (edge) cut set between a source-destination pair*

Minimum (edge) cut set between a source-destination pair i, j is defined as a set of nodes that if removed disconnects the pair i, j and has the minimum sum of capacity over all such edges.

Definition 4.3.15: *Bottleneck ratio (Wagner 2015)*

If we identify the flow and minimum edge cut set for every possible nodal pair i, j in the network, the *bottleneck ratio* of an edge is defined as the sum of the fraction of flow that was routed across it over the total amount of flow between every pair i, j. Formally,

$$BR(e) = \sum_{e \in MC} \frac{f(e)}{\sum_{\forall MC} f(MC)} \quad (4.3.9)$$

Table 4.3.15: Repeated attacks on the nodes with the highest initial average or sum of bottleneck ratios, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Erdos Renyi Directed and weighted Processability limitations	Artificial	Change in the order of the giant component	Wagner 2015
Scale free Directed and weighted Processability limitations	Artificial	Change in the order of the giant component	Wagner 2015
Small World Directed and weighted Processability limitations	Artificial	Change in the order of the giant component	Wagner 2015

k-shell decomposition

Definition 4.3.16: *k-core*

The k-core of a network is a maximal connected sub-network of degree greater or equal than k. that means that each node in the subnetwork is tied to at least k other nodes in the same subnetwork. The *k-shell* of a network is the set of nodes that belong to the k-core but do not belong to the (k+1)-core.

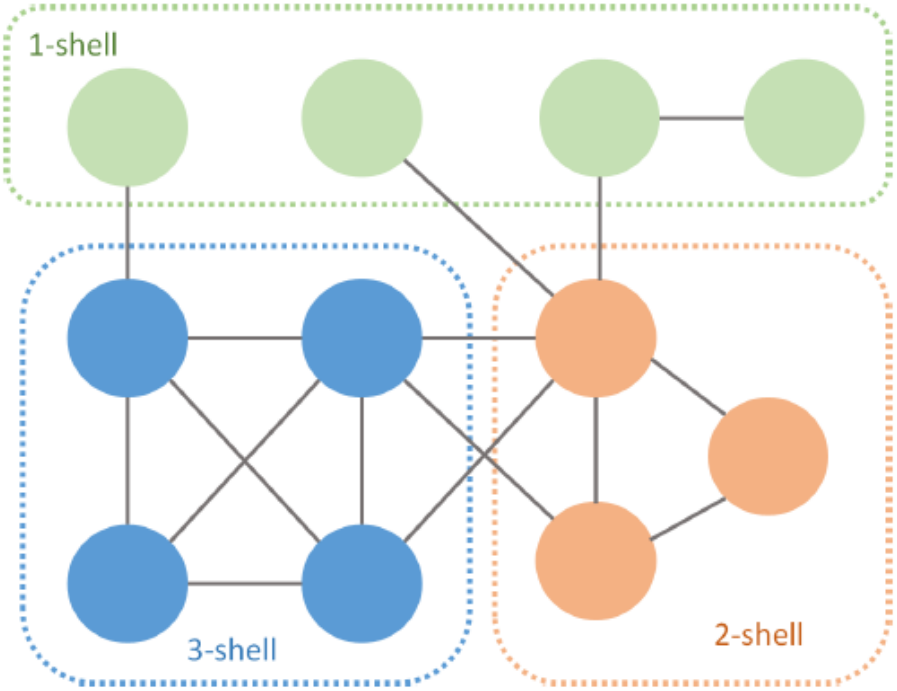

Figure 4.3.2 k-core decomposition of a fictional network. Green nodes are removed during the first iteration, orange during the second and blue during the last one.

(Garcia-Algarra et al. 2017)

The simplest algorithm for the k-shell decomposition is by recursively removing all nodes of degree equal to or less than$k$, starting with $k = 1$ and increasing it. The remaining nodes for each k compose each k-core.

Table 4.3.16: Repeated attacks on the vertices of a network in decreasing k-shell order, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Unspecified Undirected and unweighted	Mexican telecommunications network Twitter network	Change in the order of the giant component	Morone and Makse 2015

Garcia-Algarra et al. 2017 proposed two measures for ranking vertices in bipartite graphs, namely, the k-degree and k-risk.

Definition 4.3.17: *k-radius, k-degree, k-risk (Garcia-Algarra et al. 2017)*

- In a bipartite network, the k-radius of a vertex is defined as the average distance of node j (group A) towards each of the species of the innermost k-shell of group B nodes.

$$k_{radius}^{A}(v) = \frac{1}{|V^{B}(G)|} \sum_{j \in V^{B}(G)} d(v,j)\,, v \in V^{A}(G) \tag{4.3.10}$$

- The *k-degree* of a vertex is defined as

$$k_{d}^{A}(v) = \sum_{j} \frac{a_{vj}}{k_{radius}^{B}(j)} \tag{4.3.11}$$

  Where a_{vj} are the elements of the binary adjacency matrix.

- The k-risk of a vertex is defined as

$$k_{risk}^{A}(v) = \sum_{j} \left(k_{shell}^{A}(v) - k_{shell}^{B}(j)\right) + \varepsilon k_{shell}^{A}(v) \tag{4.3.12}$$

The last part of the equation is meant to solve ties and ε equals a very small value.

Table 4.3.17: Repeated attacks on the nodes with the highest k-degree and k-risk, by network structure, measure used and network type.

Network Class	Network Origin	Robustness Index	Reference
Bipartite with unspecified structure Undirected and unweighted Structural limitations	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017
Bipartite with unspecified structure Undirected and weighted Structural limitations	Networks from the web of life collection	Fragmentation Threshold ($a = 0.5$)	Garcia-Algarra et al. 2017
		AUC of the order of the giant component	Garcia-Algarra et al. 2017
		Extinction area of the second nodal group	Garcia-Algarra et al. 2017

4.4. Optimized attack methods

Few studies have calculated the absolutely optimal attack strategy for some specified concept. Latora and Marchiori 2004 studied the reduction of the efficiency in the network caused by the deletion of each node. It would also be interesting to see the efficiency reduction caused by the removal of sets of nodes and compare the results. Matisziw et al. 2012 studied the removal of nodes that causes the maximal flow disruption. Schieber et al. 2015 attacked the network by targeting the nodes and edges whose removal causes the largest disturbance in the JS-robustness are chosen for deletion.

Another attempt for a heuristic algorithm to create the most damaging attack in the network has been proposed by Arulselvan et al. 2008. Their approach is to find a maximal independent set of nodes in the network and start recreating the network by adding nodes, until the nodes remaining to be added are equal to the order of the attack we intended to make. Nodes are added whilst minimizing $\sum_{h\in M_j}\frac{\sigma_h(\sigma_h-1)}{2}$, where σ_h is the number of connected nodes and M_j is the set of all maximal connected components in the subgraph that has been recreated to this point.

An interesting approach was taken by Chen et al. 2008, where they modified a nested dissection algorithm and created the equal graph partitioning algorithm. The nested dissection algorithm separates the network into two components of equal order with a minimum number of nodes removed. The equal graph partitioning algorithm is designed to separate the network into two components of arbitrary relative orders, thus one can use it to separate the network into many small components by applying the algorithm iteratively. This approach has been shown to be less effective for heterogenous networks by Morone and Makse 2015.

Table 4.4.1: Repeated optimized attacks on the nodes of a network, by attack method network structure, measure used and network type.

Attack target	Network Class	Network Origin	Robustness Index	Reference
Maximal efficiency reduction if removed from the original network	Unspecified structure Undirected and weighted	Infonet internet backbone (US) Infonet internet backbone (EU) 9/11 terrorist network	Change in the efficiency	Latora and Marchiori 2004
Maximal total flow disruption	Unspecified structure Undirected and weighted	US Internet Traffic	Elasticity of MGG-Robustness	Matisziw et al. 2012
Maximal criticality according to the JS-robustness	Unspecified structure Undirected and unweighted	Western Power Grid of the US	JS-robustness and its change over the attack	Schieber et al. 2015
	Unspecified structure Undirected and weighted	Dolphin Social Network	JS-robustness and its change over the attack	Schieber et al. 2015
Maximal Independent Set optimization	Unspecified structure Undirected and weighted	Terrorist network	JS-robustness and its change over the attack	Arulselvan et al. 2008
Equal Graph Partitioning	Erdos Renyi Undirected and unweighted	Artificial	Change in the order of the giant component	Chen et al. 2008
	Scale Free Undirected and unweighted	Artificial	Change in the order of the giant component	Chen et al. 2008 Morone and Makse 2015
	Random Regular Undirected and unweighted	Artificial	Change in the order of the giant component	Chen et al. 2008 Morone and Makse 2015
	Unspecified structure Undirected and unweighted	Workplace network Internet (Autonomous System) High energy particle physics station network	Change in the order of the giant component	Chen et al. 2008

		Metabolic network		
Optimization via genetic algorithm	Unspecified structure Directed and weighted Structural limitations	Food Webs	Extinction Area	Allesina and Pascual 2009

Table 4.4.2: Repeated optimized attacks on the links of a network, by attack method network structure, measure used and network type.

Attack target	Network Class	Network Origin	Robustness Index	Reference
Maximal criticality according to the JS-robustness	Unspecified structure Undirected and unweighted	Western Power Grid of the US	JS-robustness and its change over the attack	Schieber et al. 2015
	Unspecified structure Undirected and weighted	Dolphin Social Network	JS-robustness and its change over the attack	Schieber et al. 2015

4.5. Mixed attack strategies

In this category belongs any type of attack that combines in some way two other (e.g. removing some of the highest degree nodes first, then targeting the ones with the highest betweenness).

Tanizawa et al. 2005 studied the behavior of scale free networks when subjected to both highest degree and random attacks. A 2-value degree distribution was proposed as the most robust when subjected to both types of attacks.

Wu et al. 2007 took a different approach, with the idea that the attacker will not know perfectly the structure of the network, but he will remove the highest degree nodes of the known structure and then follow with random attacks. Results were given for the various level of knowledge of the attacker.

Table 4.5.1: Repeated mixed-strategy attacks on the nodes of a network, by attack combination, network structure and origin, and measure used.

Attack target	Network Class	Network Origin	Robustness Index	Reference
Degree then random	Scale Free Undirected and weighted	Artificial	Critical Threshold	Tanizawa et al. 2005
	Bimodal degree distribution Undirected and weighted	Artificial	Critical Threshold	Tanizawa et al. 2005
Degree then random based on incomplete information on the network	Scale Free Undirected and weighted	Artificial	Change in the order of the giant component	Wu et al. 2007
			Critical Threshold	Wu et al. 2007

4.6. Probabilistic models with limited cost

Instead of comparing the damage inflicted on a network against the removals made on it, it might be more realistic to consider the inequality on the level of difficulty for each attack. Deng and Wu 2015 and Deng and Wu 2016 proposed a method to adjust the cost of removing a node by a network measure (in their examples they used the degree of each node).

Simultaneously, they proposed a probabilistic method of attack. Specifically, an attack strategy where they determine the nodes to be removed by turning it into an unequal probability sampling problem without replacement. The selection probability that a node v_i is sampled to attack in each sample is defined as

$$\nabla_i = \frac{p_i^x}{\sum_{i=1}^{N} p_i^x} = \frac{p_i^{\tan\left(\frac{\pi}{2}\delta\right)}}{\sum_{i=1}^{N} p_i^{\tan\left(\frac{\pi}{2}\delta\right)}}, \qquad \delta \in [-1,1] \tag{4.6.1}$$

With p_i being a certain property of the node v_i (e.g. degree).

For $\delta = 1$ the nodes are removed with descending property value while for $\delta = -1$ with descending and for $\delta = 0$ they are removed uniformly at random.

In their results, in a limited cost model for various exponents to the cost of removal, it was shown that in scale free networks with varying exponents, the optimal values for δ are

usually close to the extremes. Specifically, as the exponent of the scale free model drops and as the cost of removal rises, optimal δ drops to -1.

Table 4.6.1: Repeated probabilistic node attacks with limited cost based on the degree, by network structure and origin, and measure used.

Network Class	Network Origin	Robustness Index	Reference
Scale Free Undirected and unweighted	Artificial	Proportion of nodes in the largest component	Deng and Wu 2015 Deng and Wu 2016

4.7. Attacks unrelated to the network structure

Another very important factor to be considered when studying network functionality is the context of the network's existence. Bad weather may be considered as random failures of edges or vertices for transportation networks, but there is always the question of scale. For example, the explosion of Eyjafjallajökull in 2010 caused great perturbations in the European air traffic network, and it has been shown (Wilkinson et al. 2011, Dunn and Wilkinson 2015) that this network is vulnerable when subjected to *spatial hazards*.

This type of attack can be modelled by a different variable, unrelated to the network structure, such as geographical location or ranking in an organization. It is reasonable to assume that an interloper will not be able to acquire access to every one of his targets in the network, at least not with the same ease. It is thus a promising idea, although case specific, to consider weighing the removal cost in limited cost models, in respect to geographical, or hierarchical data as well.

Another concept that might not be related immediately to the network, is the interconnection of the roles that the actors (vertices) have. Duijn et al. 2014 recreated not only the network with the actors as nodes, but the hierarchical structure of the cannabis distribution network as well. Then they attempt, among others, two attacks based on the second structure. The first is attacking the role that needs the highest/rarest expertise, while the second is attacking actors based on their degree in the role network

Table 4.7.1: Repeated attacks on the nodes of a network, unrelated to the structure.

Method of attack	Network Class	Network Origin	Robustness Index	Reference
Specific role	Scale Free Undirected and Unweighted Self-rewiring	Cannabis distribution network	Change in the efficiency	Duijn et al. 2014
			Change in the network density	Duijn et al. 2014
Highest degree in role network	Scale Free Undirected and Unweighted Self-rewiring	Cannabis distribution network	Change in the efficiency	Duijn et al. 2014
			Change in the network density	Duijn et al. 2014
Spatial Hazard	Truncated power law degree distribution Undirected and Unweighted Structural limitations Self-rewiring	European air traffic and similar artificial	Percentage of destroyed air-routes	Wilkinson et al. 2011 Dunn and Wilkinson 2015
			Change in the order of the giant component	Wilkinson et al. 2011 Dunn and Wilkinson 2015
Top-down on the hierarchy ranking	Hierarchy Undirected and unweighted Structural limitations?	Artificial	Connectivity Robustness	Dodds et al. 2003
			Maximal congestion centrality	Dodds et al. 2003

5. Example

5.1. Methods

In this chapter we will measure the robustness of four different directed networks, and subsequently we shall test these measurements by attacking these networks in various ways and comparing the results of these attacks.

The networks we have used are four different food webs, specifically the dry season of: the Florida bay ecosystem (Ulanowicz et al. 1998), the Cypress wetland ecosystem (Ulanowicz et al. 1997), the Mangrove ecosystem (Ulanowicz et al. 1999) and the Graminoid ecosystem (Ulanowicz et al. 2000). The data was obtained by Batagelj and Mrvar 2006 and they are licensed under a Creative Commons-NonCommercial-ShareAlike 2.5 License https://creativecommons.org/licenses/by-nc-sa/2.5/.

We are interested only in the network properties; therefore, we did not account for the weights or other properties (such as the biomass of the various species) of the network as they represent elements irrelevant to our binary approach of the components of the network being connected or not.

Furthermore, we have arbitrarily considered the networks without self-loops (we have removed five self-loops from the Mangrove ecosystem). This is done because the existence of such edges does not affect any of our evaluation methods, but can affect both the methods of attack and the measures of robustness.

We use three different measures of robustness:

- The normalized natural connectivity of the network (definition 3.5.4). As the natural connectivity strictly increases as edges are added, we can normalize it by dividing the natural connectivity of a network with that of a complete network without self-loops of equal order.

- The difference between the absolute value of the first two largest eigenvalues. We wish to check if the principle behind the deviation from the ideal good expander (definition 3.5.9), holds for directed networks. Specifically, that if the gap between the first two eigenvalues is sufficiently large the network has good expansion properties and is, thus, more robust.
- The assortativity coefficient (definition 2.3.2) based on the degrees of the network, as defined by Newman 2003. We wish to check if in these networks the assortativity coefficient is positively or negatively correlated with the robustness to disconnection.

All three of the robustness measures are calculated both for the whole network and for its largest strongly connected component. However, the difference of the eigenvalues was the same for both components across all three networks.

We measure eight different methods of attack.

- Random attacks: We create five different random enumerations of the vertices of the network and we remove them sequentially. In the end we take the average of all 5 measurements and treat it as the result of a single attack. This is done in order to avoid one lucky hit, or a large series of unlucky ones, producing skewed results. This method should be used as a baseline in order to judge the effectiveness of the rest of the attacks, as for any method of attack to be successful it should perform better than random.
- Degree attacks: We remove the nodes of the network in descending order of degree. In case of ties the order is determined randomly. This method is performed in three different ways:
    - Total degree: Where the number of all edges starting from and ending to each node is used.
    - In-degree: Where the number of all edges ending to each node is used.
    - Out-degree: Where the number of all edges starting from each node is used.
- Betweenness attacks: We remove the nodes of the network in descending order of betweenness. In case of ties the order is determined randomly.

- Collective influence attacks: We remove the nodes of the network in descending order of collective influence. In case of ties the order is determined randomly. This method is performed in two different ways:
    - 2-step collective influence
    - 3-step collective influence
- Eigencentrality: We remove the nodes of the network in descending order of eigencentrality. In case of ties the order is determined randomly.

The evaluation of the attacks is performed by their ability to disconnect the network. This can be seen in the change of the order of the giant component (both the weakly and the strongly connected) over the attack and further quantified by the area under the curve of these plots. It can also be observed as the *a*-fragmentation threshold (definition 3.3.2) of the network both for the weakly and strongly connected giant components, for the various values of a and the different attacks. We present the values for the .75, .50 and .25-fragmentation thresholds of the networks. The thresholds of the giant strongly connected component are in proportion of the whole network. In presenting these *a*-fragmentation thresholds, we have reduced the values exceeding $1 - a$ due to not being divisible by 100, to be equal to $1 - a$. However, there is an interesting case, for the .75-fragmentation threshold on the 2-steps collective influence attack in the Mangrove food web, where more than 25% of the nodes have to be removed to reduce the giant strongly connected component bellow 75% of its original size, this is unexpected, due to the fact that a targeted attack is expected to destroy the network faster, and thus, one would expect it should destroy both the strongly and the weakly connected giant component faster.

We make a further evaluation of the ability of the attacks to break down the communicability in the network by measuring the change of the global efficiency (normalized by the initial order of the network (definition 3.3.8, theorem 3.3.1.) over the attack and its area under the curve.

5.2. Results

Robustness indices

Table 5.2.1: Measures of robustness for the four different networks. Weak and strong indicate if the measurement is for the giant weakly or strongly connected component. Highlighted in grey are the largest values in every category, and the smallest ones in the assortativity coefficient as its positive or negative correlation to robustness is under evaluation.

	Density (weak)	Density (strong)	Assortativity coefficient (weak)	Assortativity coefficient (strong)	Natural connectivity normalized (weak)	Natural connectivity normalized (strong)	Spectral gap
Florida bay	0.131	0.153	-0.234	-0.300	0.050	0.066	3.644
Cypress wetlands	0.129	0.162	-0.334	-0.339	0.043	0.065	3.107
Everglades graminoids	0.194	0.203	-0.408	-0.421	0.107	0.127	5.049
Mangrove estuary	0.160	0.171	-0.311	-0.227	0.105	0.120	5.952

Table 5.2.1 shows that the graminoids food web should be most robust according to density, while the Cypress wetlands should be the least robust according to the density of its weak giant component (with Florida bay being a very close second), and the Florida bay according to the density of the strong giant component. The Natural connectivity suggests, both for the weak and the strong giant components, that the Cypress is the least robust network, again with the Florida bay being a very close second, while the most robust seems to be the Graminoids network with the Mangrove being close. According to the natural connectivity there should be a distinct difference between these two pairs of networks. The gap between the first and second eigenvalues of the adjacency suggests that the Cypress network should be the least robust and the Mangrove should be the most. The assortativity coefficient points on the one hand to the graminoids ecosystem (both its weak and strong giant components) and to the other hand the weak giant component of the Florida bay

network and the strong giant component of the Mangrove network, to be most and least robust.

In summary, the robustness indices point in general towards the Cypress wetlands and the Florida bay to be the least robust of the four networks and to the Everglades graminoids or the Mangrove estuary to be the most robust.

Weak giant component

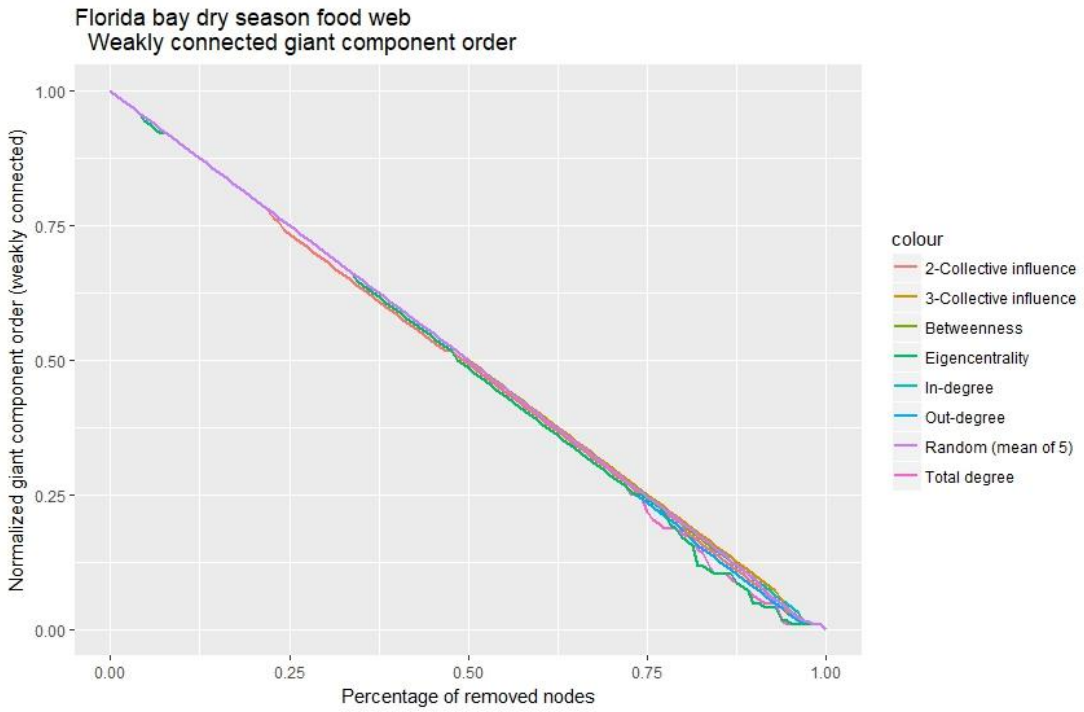

Figure 5.2.1: Giant weakly connected component order of the Florida bay food web when subjected to eight different attacks. The attacks are made on the nodes of the network descending on the ranking provided by 2 and 3 steps collective influence, betweenness, eigencentrality, total, in and out degree. We also present the average impact of five random attacks.

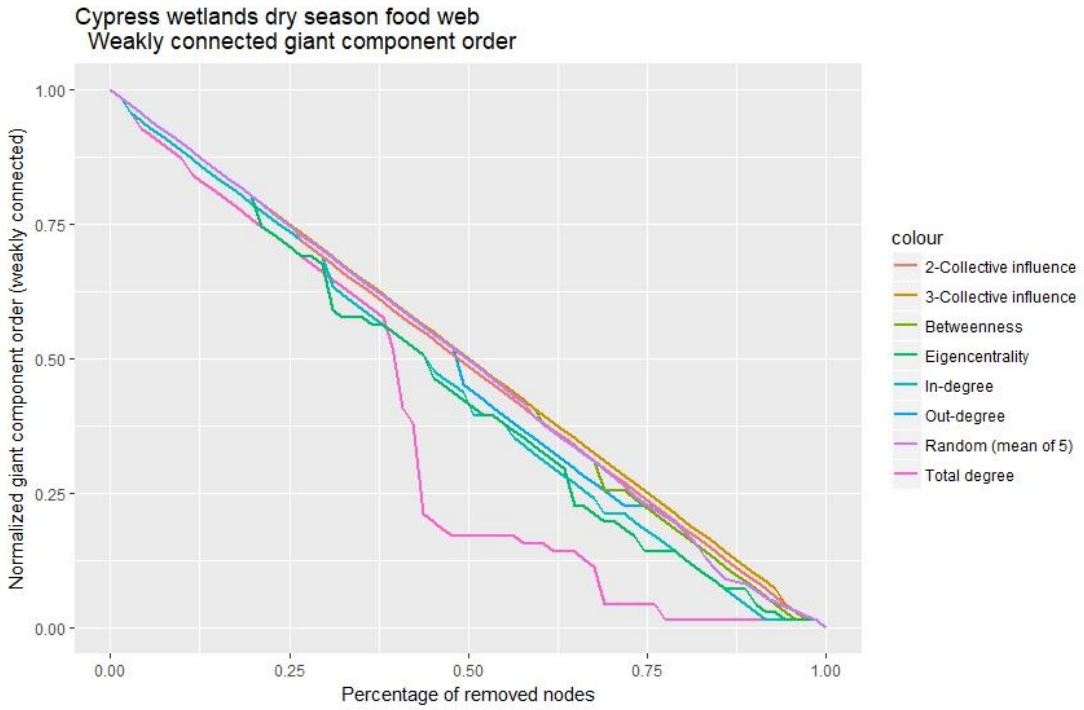

Figure 5.2.2: Giant weakly connected component order of the Cypress wetlands food web when subjected to eight different attacks, as in Figure 5.2.1.

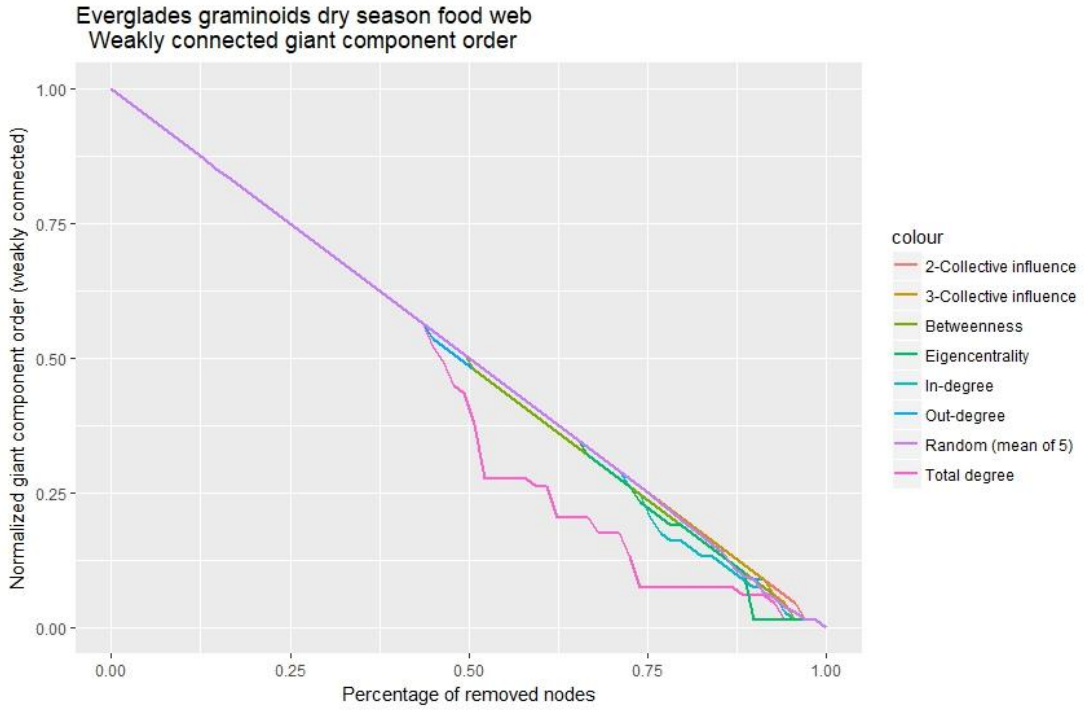

Figure 5.2.3: Giant weakly connected component order of the Everglades graminoids food web when subjected to eight different attacks, as in Figure 5.2.1.

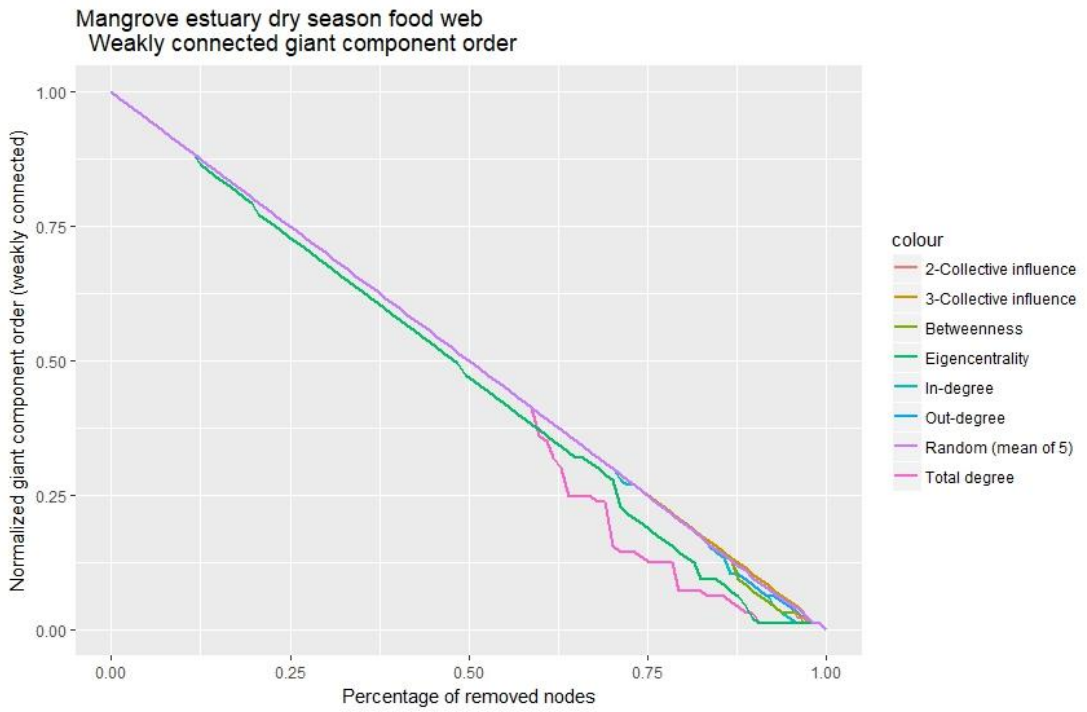

Figure 5.2.4: Giant weakly connected component order of the Mangrove estuary food web when subjected to eight different attacks, as in Figure 5.2.1.

- In Figure 5.2.1. we can see that all methods of attack have a low impact on the weak giant component of the Florida bay food web, with the total degree attack and the eigencentrality attack showing a better performance towards the end.
- In Figure 5.2.2 we can see that the weak giant component of the Cypress wetlands is far more vulnerable to total degree attacks than the other kinds, with eigencentrality and in-degree being second and third, unclear on the order.
- In Figure 5.2.3 we can see that the weak giant component of the Everglades graminoids is also far more vulnerable to total degree attacks than the other kinds, but they stop being effective towards the end of the attack (when almost three quarters of the network have been removed).
- In Figure 5.2.4 we can see that the weak giant component of the Mangrove estuary is more vulnerable to total degree attacks and eigencentrality attacks than the other kinds.

These results can be further explored in table 5.2.2 where we can see that the weak giant component of the Florida bay food web is robust to all attacks (has area under the curve close to 0.5) and all other networks have their minimal area occur when the network is subjected to total degree attacks.

Table 5.2.2: Area under the curve of the normalized giant weakly connected component (rounded to three decimal digits) of four networks when subjected to eight different attacks. Networks are food webs during the dry season of the Florida bay, the Cypress wetlands, the Everglades graminoids and the Mangrove estuary. Attacks are as in Figure 1. Highlighted in grey are the highest areas for each attack. Highlighted in black with white letters are the lowest areas for each attack.

Weak GC AUC	Degree attacks			Betweenness attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.490	0.498	0.491	0.497	0.491	0.499	0.486	0.497
Cypress Wetlands	0.366	0.450	0.483	0.490	0.490	0.499	0.446	0.490
Everglades graminoids	0.441	0.492	0.493	0.493	0.500	0.499	0.492	0.498
Mangrove estuary	0.466	0.497	0.498	0.497	0.500	0.500	0.471	0.499

We can see that the Cypress wetlands is the most vulnerable network to damage to its weak giant component in every instance, while the Mangrove estuary and the Florida bay are the most robust with the Everglades graminoids being close in terms of robustness.

Table 5.2.3: a-fragmentation threshold (rounded to the second decimal digit) of the weakly connected giant component of four networks when subjected to eight different attacks. Networks and attacks are as in Figure 5.2.1. Highlighted in grey are the cases where the threshold is equal to a (no disconnections occur). Highlighted in black and written in white are the lowest values for each attack, unless all are equal.

Weak GC .75 Threshold	Degree attacks			Betweenness attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.25	0.25	0.25	0.25	0.24	0.25	0.25	0.25
Cypress Wetlands	0.21	0.24	0.25	0.25	0.25	0.25	0.21	0.25
Everglades graminoids	0.25	0.25	0.25	0.25	0.25	0.25	0.25	0.25
Mangrove estuary	0.25	0.25	0.25	0.25	0.25	0.25	0.24	0.25
Weak GC .50 Threshold	Degree attacks			Betweenness attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.50	0.50	0.49	0.50	0.49	0.50	0.48	0.50
Cypress Wetlands	0.41	0.45	0.49	0.50	0.49	0.50	0.45	0.50
Everglades graminoids	0.46	0.50	0.49	0.50	0.50	0.50	0.50	0.50
Mangrove estuary	0.50	0.50	0.50	0.50	0.50	0.50	0.48	0.50
Weak GC .25 Threshold	Degree attacks			Betweenness attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.73	0.74	0.73	0.74	0.74	0.75	0.73	0.75
Cypress Wetlands	0.44	0.68	0.70	0.73	0.75	0.75	0.65	0.73
Everglades graminoids	0.62	0.74	0.74	0.74	0.75	0.75	0.74	0.75
Mangrove estuary	0.64	0.75	0.75	0.75	0.75	0.75	0.71	0.75

The *a*-fragmentation thresholds of the weak giant component indicate that the actual damage done to the network with most methods of attack is menial before more than 70% of its nodes have been removed. Apart from the Cypress wetlands network, which as indicated by the robustness indices was the least robust, all other networks' weak giant components had very few, if any, secondary disconnections. The only exception in this is that the Everglades graminoids and the Mangrove estuary networks sustained some disconnections (a little more than a tenth of their nodes) were disconnected from the weak giant component) when the 62 and 64 percent of their nodes being removed.

In summary, the only deviation from the expectations provided by the robustness indices, is the Florida bay network, which was expected to be vulnerable, but was very robust to the disconnection of its weak giant component. The Cypress wetlands network was indeed the least robust in almost every case, while in general the Mangrove estuary was more robust than the everglades graminoids in terms of their weak giant component.

Strong giant component

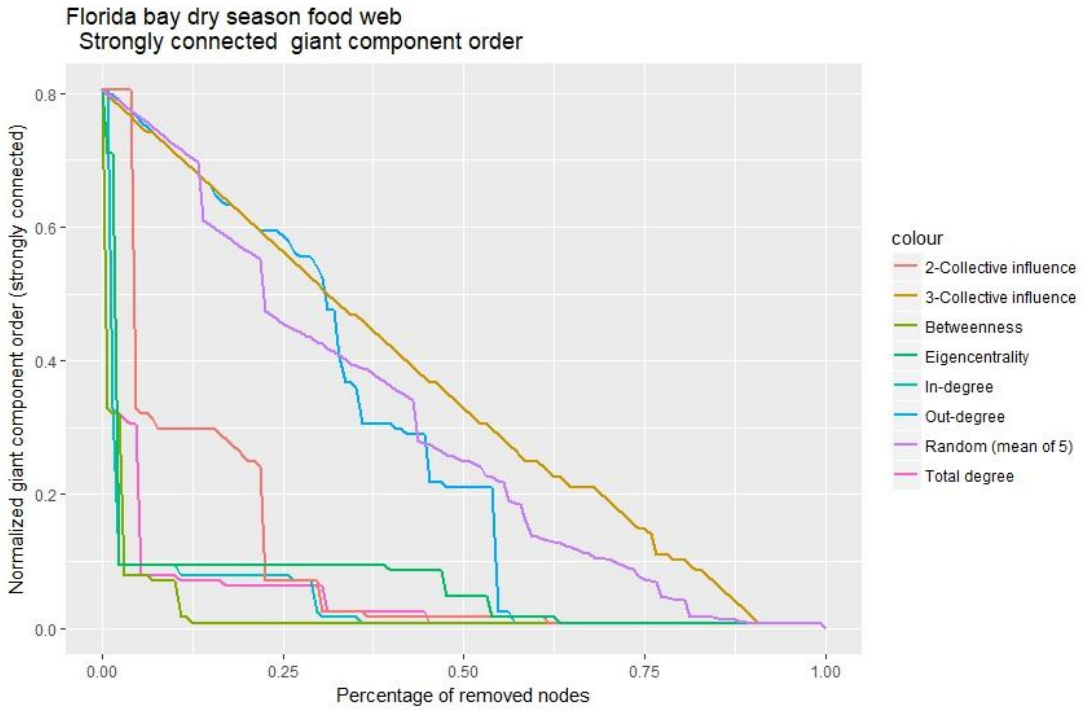

Figure 5.2.5: Giant strongly connected component order of the Florida bay food web when subjected to eight different attacks, as in figure 5.2.1.

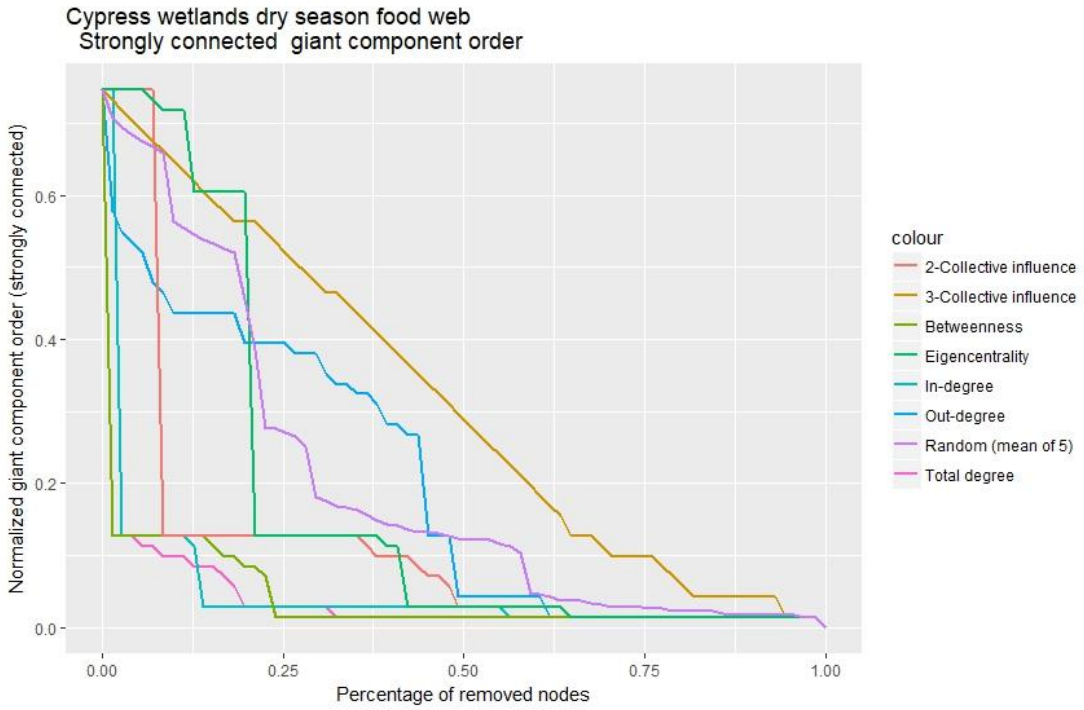

Figure 5.2.6: Giant strongly connected component order of the Cypress wetlands food web when subjected to eight different attacks, as in figure 5.2.1.

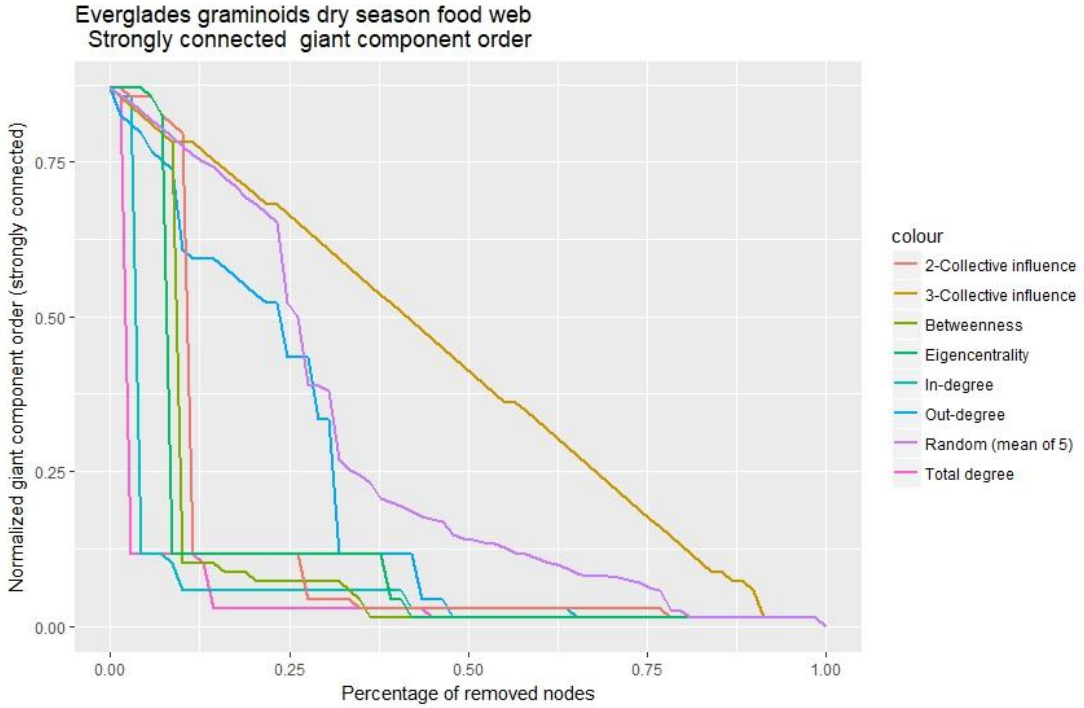

Figure 5.2.7: Giant strongly connected component order of the Everglades graminoids food web when subjected to eight different attacks, as in figure 5.2.1.

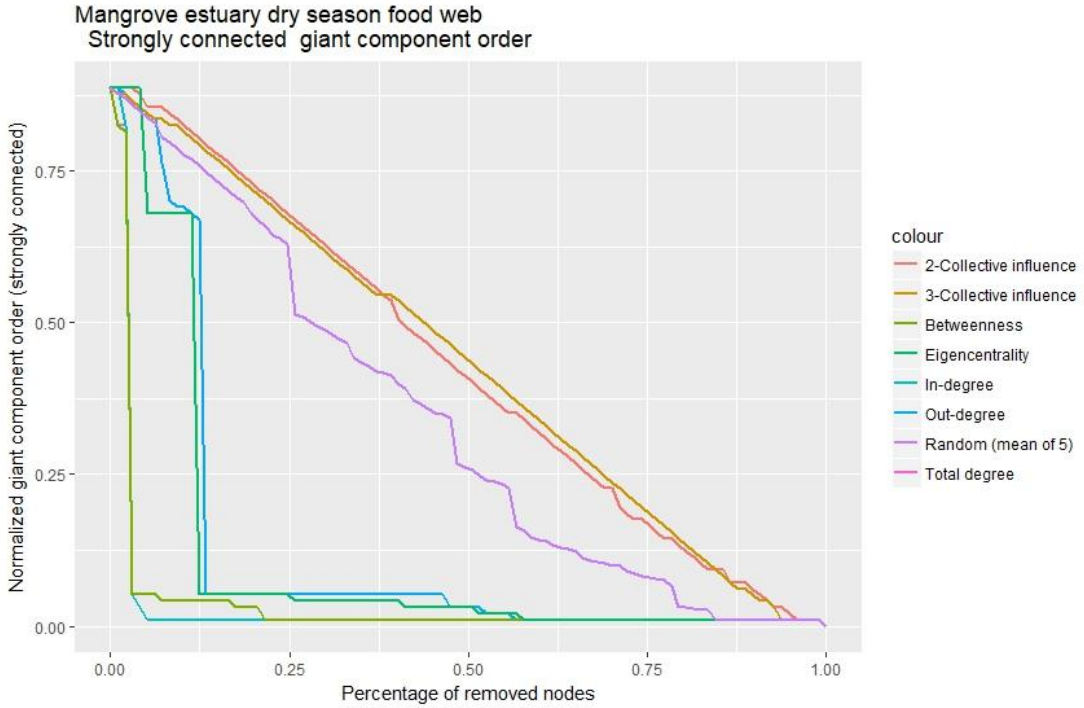

Figure 5.2.8: Giant strongly connected component order of the Mangrove estuary food web when subjected to eight different attacks, as in figure 5.2.1.

- In figure 5.2.5 we can see that the giant strongly connected component of the Florida bay network is very vulnerable to total degree, in-degree, eigencentrality and betweenness attacks, as it breaks down to less than 10% of the network's order with less than 10% of the nodes in the network being removed. 2-steps collective influence is also a relatively successful attack that breaks it down to less than half its order with less than 10% nodes removed and at less than 10% of the network's order at a bit less than 25% of the nodes being removed.
- In figure 5.2.6 we can see that the giant strongly connected component of the Cypress wetlands network breaks down even faster under eigencentrality and betweenness attacks, but to a lesser extent, also more methods of attack seem to be effective against it, even though some of them are less effective than in the Florida bay network.

- In figure 5.2.7 we can see that the giant strongly connected component of the Everglades graminoids network also breaks down very fast, specifically, less than 12.5% of the network needs to be removed to achieve a less than 12.5% participation in the strong giant component for total degree, in-degree, eigencentrality, betweenness and 2-steps collective influence. Also, out-degree and random attacks seem to be effective as well.
- In figure 5.2.8 we can see that the giant strongly connected component of the Mangrove estuary network we can see a similar great effectiveness for total degree, in-degree and betweenness, and an only slightly less effectiveness for eigencentrality and out-degree attacks.

These results can be further explored in table 5.2.4 where we can see that total degree, in-degree and betweenness attacks are the best methods to shatter the strong giant component of most of the networks, with eigencentrality being followed by the rest, while performing very well on the Florida bay network. We note that neither of the two collective influence measures show promising results, especially in the Mangrove estuary network. We believe the cause of this to be the size of the networks.

Table 5.2.4: Area under the curve of the normalized giant weakly connected component (rounded to three decimal digits) of four networks when subjected to eight different attacks. Networks and attacks are as in figure 5.2.1. Highlighted in grey are the highest areas for each attack. Highlighted in black with white letters are the lowest areas for each attack.

Strong GC AUC	Degree attacks			Betweenness attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.049	0.043	0.277	0.027	0.103	0.354	0.066	0.299
Cypress Wetlands	0.050	0.052	0.198	0.046	0.118	0.314	0.180	0.201
Everglades graminoids	0.054	0.073	0.211	0.112	0.133	0.418	0.118	0.283
Mangrove estuary	0.036	0.037	0.129	0.041	0.427	0.432	0.116	0.338

We can see that, surprisingly, the strong giant component of the Mangrove estuary network is the most vulnerable to degree attacks, but the most robust to collective influence and random attacks. Also surprising is the fact that the strong giant component of the Cypress wetlands network is slightly more robust than most other networks (except the Everglades graminoids) in almost all efficient methods of attack. Finally, the Everglades graminoids network is shown to be the most, or second most, robust under all attacks.

Table 5.2.5: a-fragmentation threshold (rounded to the second decimal digit) of the strongly connected giant component of four networks when subjected to eight different attacks. Networks and attacks are as in figure 5.2.1. Highlighted in grey are the cases where the threshold is equal to a (no disconnections occur). Highlighted in black and written in white are the lowest values for each attack, unless all are equal. Highlighted in light grey are the largest values for each attack. Highlighted in orange is the only case where more than 1-a is smaller than the a-fragmentation threshold.

Weak GC .75 Threshold	Degree attacks			Betweennes s attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.02	0.02	0.21	0.01	0.05	0.021	0.02	0.16
Cypress Wetlands	0.03	0.03	0.03	0.01	0.08	0.23	0.21	0.11
Everglades graminoids	0.03	0.04	0.01	0.01	0.12	0.25	0.09	0.23
Mangrove estuary	0.03	0.03	0.13	0.03	0.27	0.25	0.12	0.22
Weak GC .50 Threshold	Degree attacks			Betweennes s attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.02	0.02	0.34	0.01	0.05	0.42	0.02	0.34
Cypress Wetlands	0.03	0.03	0.31	0.01	0.08	0.42	0.21	0.23
Everglades graminoids	0.03	0.04	0.24	0.10	0.12	0.48	0.09	0.28
Mangrove estuary	0.03	0.03	0.13	0.03	0.46	0.49	0.12	0.34
Weak GC .25 Threshold	Degree attacks			Betweennes s attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.05	0.02	0.55	0.03	0.23	0.70	0.02	0.56
Cypress Wetlands	0.03	0.03	0.45	0.01	0.08	0.61	0.21	0.30
Everglades graminoids	0.03	0.04	0.32	0.10	0.12	0.71	0.09	0.38
Mangrove estuary	0.03	0.03	0.13	0.03	0.71	0.72	0.12	0.57

Table 5.2.5 agrees with our previous observations, namely that degree attacks are extremely effective and that most attacks (excluding collective influence and random) are very effective against all strong giant components. However, 2-steps collective influence is a very effective method on the Florida bay and Cypress wetlands networks. Furthermore,

we observe that most of the damage on the strong giant components is achieved early on, as in many cases the *a*-fragmentation threshold is equal, or almost equal for $a = 0.75, 0.50$ and 0.25. This shows that there are a few key nodes holding the strong giant component together, and most attack methods disconnect them very efficiently.

Another thing to point out, is that the .25-threshold, in most cases of attack, agrees with our expectations of the Cypress wetlands being the least robust network followed by the Florida bay, and that the Everglades graminods and the Mangrove estuary are the most robust networks (excluding out-degree and eigencentrality). The .50 and .75 thresholds also somewhat agree, although the Florida bay is shown to be less robust than the Cypress wetlands, the differences are very small in most cases.

In summary, according to the *a*-fragmentation thresholds there are small deviations to the effect the various attacks have on the networks, from the expected effects provided by the robustness indices, but the ordering seems to be following the same patterns, the two networks shown to be less robust, are less robust and the two networks shown to be more robust, are indeed more robust. These results disagree somewhat with the AUC of the graphs, but the differences in the areas under the curves are very small; small enough to be affected by very small strongly connected clusters surviving for a long period of time in the network and biasing the results.

Global efficiency

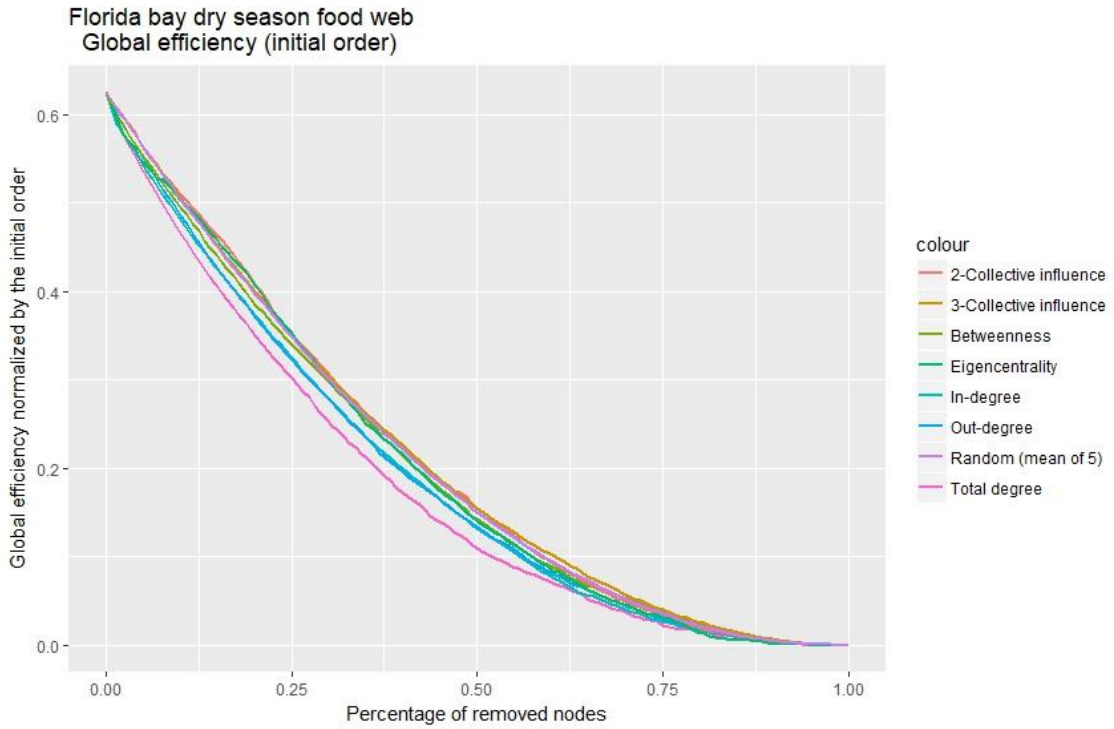

Figure 5.2.9: Global efficiency normalized by the initial network order of the Florida bay food web when subjected to eight different attacks, as in figure 5.2.1.

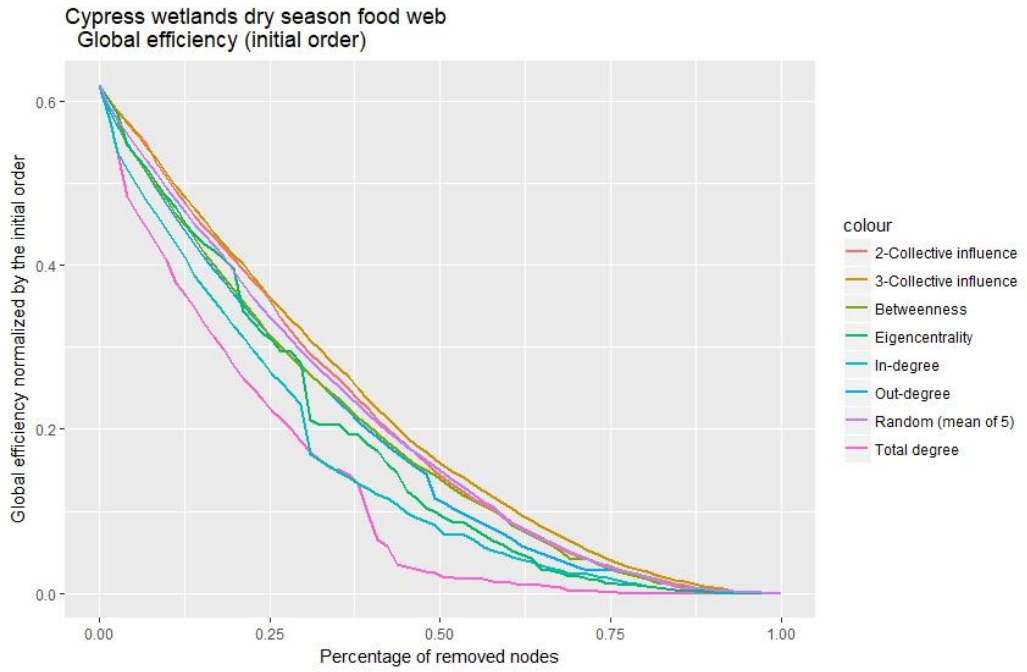

Figure 5.2.10: Global efficiency normalized by the initial network order of the Cypress wetlands food web when subjected to eight different attacks, as in figure 5.2.1.

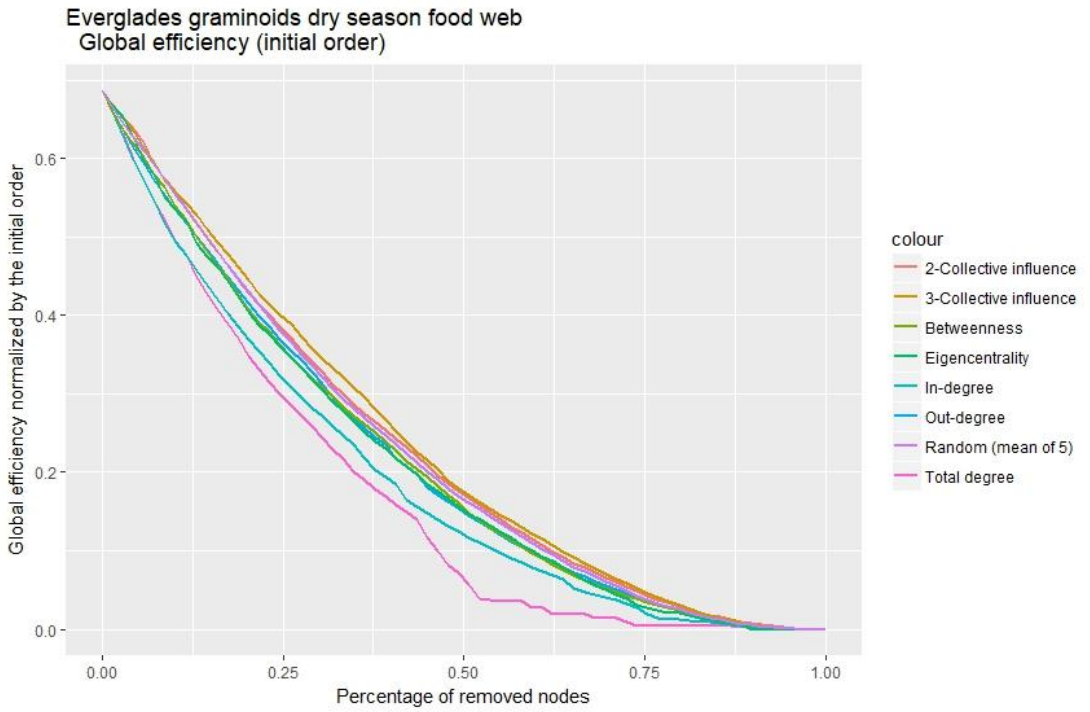

Figure 5.2.11: Global efficiency normalized by the initial network order of the Everglades graminoids food web when subjected to eight different attacks, as in figure 5.2.1.

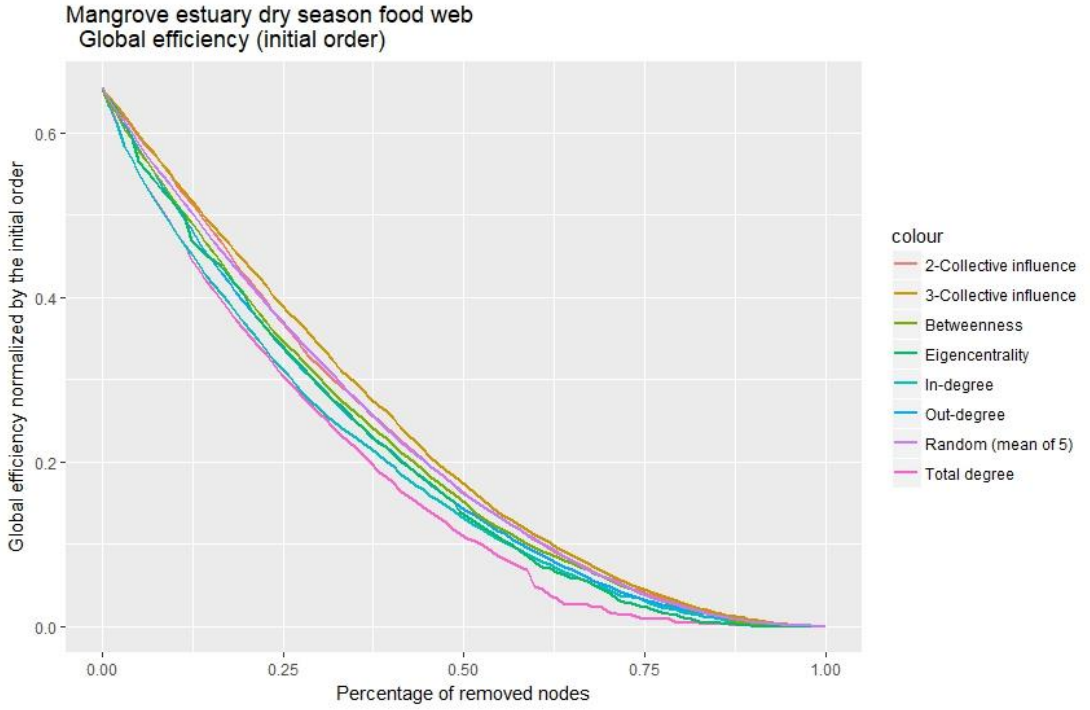

Figure 5.2.12: Global efficiency normalized by the initial network order of the Mangrove estuary food web when subjected to eight different attacks, as in figure 5.2.1.

- In figure 5.2.9 we can see that the global efficiency of the Florida bay network is most vulnerable to total degree attacks, but all of the methods of attack perform in a similar way.
- In figure 5.2.10we can see that the global efficiency of the Cypress wetlands ndetowork is most vulnerable to total degree attacks, followed by in-degree and eigencentrality attacks. Furthermore, they seem to be more effective than in the rest of the networks.
- In figure 5.2.11 we can see that the global efficiency of the Everglades graminoids network is most vulnerable to total degree and in-degree attacks.
- In figure 5.2.12 we can see that the global efficiency of the Mangrove estuary network is most vulnerable to total degree and eigencentrality attacks.

Table 5.2.6: Area under the curve of the global efficiency normalized by the initial network order (rounded to three decimal digits) of four networks when subjected to eight different attacks. Networks and attacks are as in table 5.2.1. Highlighted in grey are the highest areas for each attack. Highlighted in black with white letters are the lowest areas for each attack.

Global efficiency (normalized) AUC	Degree attacks			Betweenness attacks	Collective influence attacks		Eigenvector centrality attacks	Random attacks
	Total	In	Out		2-steps	3-steps		Mean of 5
Florida bay	0.178	0.191	0.192	0.201	0.208	0.209	0.201	0.206
Cypress Wetlands	0.129	0.156	0.186	0.192	0.206	0.214	0.178	0.201
Everglades graminoids	0172	0.193	0.217	0.217	0.230	0.236	0.215	0.226
Mangrove estuary	0.178	0.191	0.204	0.210	0.221	0.229	0.200	0.219

Table 5.2.6 confirms the observations from the plots. In summary, all networks seem more vulnerable to total degree, in-degree and eigencentrality attacks, with the Cypress wetlands seemingly losing efficiency faster than the others. Furthermore, we see that it agrees almost entirely with the robustness indices. Specifically, the Cypress wetlands network seems to be the most vulnerable in every attack except one, the Florida bay network is in most cases the second or first least robust network, the mangrove estuary is

in most cases the second most robust, and the Everglades graminoids is in all cases except one the most robust network.

5.3. Conclusions

We note that neither of the two collective influence measures show promising results, especially in the Mangrove estuary network.

The anomalies in the ranking of the networks according to their robustness compared to the rankings provided by the impact on the attacks on their strong giant components, lead us to assume that the robustness indices, whether they are calculated for the strong giant component or the weak, do not necessarily encompass the robustness of the strong giant component to attacks on the whole of the network. On the one hand, a more thorough research is required to verify if this is true in general and not for only a few networks, and on the other, research is also required to explore the degree that the property of strong connection affects such indices.

The very low values for the AUCs of the strong giant component for all networks under most attacks, suggest that there are a few key nodes holding the strong giant components together, and most attack methods detect them very efficiently, breaking the network very early on. Furthermore, the ranking of the AUCs and the *a*-fragmentation thresholds disagree somewhat, but the differences in the areas under the curves are very small; small enough to be affected by tiny strongly connected clusters surviving for a long period of time in the network and biasing the results. Because of this, we find the more direct approach of the *a*-fragmentation threshold to be a better criterion than the change in the size of the giant component. Our belief is enhanced by the fact that it agrees with the selected indices of robustness.

Moreover, the fact that the strong giant component of the Cypress wetlands network is slightly more robust than most other networks (except the Everglades graminoids) in almost all efficient methods of attack, leads us to assume that the robustness indices, whether they are calculated for the strong giant component or the weak, do not necessarily

encompass the robustness of the strong giant component to attacks on the whole of the network. On the one hand, a more thorough research is required to verify if this is true in general and not for only a few networks, and on the other, research is also required to explore the degree that the property of strong connection affects such indices.

Finally, the fact that the AUC of the global efficiency shows a near perfect agreement on the ranking of the networks according to their robustness. This leads us to assume that the robustness indices we have studied are correlated more with the ability of the network to maintain efficient communications, rather than with its ability to remain strongly, or weakly connected when subjected to a variety of attacks.

Concerning the methods of attack, we have seen that the total degree method is consistently the most efficient, followed by the in-degree, eigencentrality and sometimes betweenness attacks. This is in general expected, as attacks based on the initial degree have been shown to be more effective than those based on the initial betweenness on undirected networks.

Surprisingly, the collective influence methods were heavily outperformed by the rest. This might have to do with one of two things. One possibility is that they perform better on larger networks, as they have been demonstrated to work to networks of with nodes numbering in the order of 10^7. Or it might have to do that it is related as a measure by definition to the out-degree, which in these networks has been shown to not be an efficient method of attack. This requires further investigation.

6. Results

6.1. Introduction

In this chapter we will present the most important of the known results, as well as compare them to our findings in chapter 5.

6.2. Summary of known results

In this chapter we will present a summary of the known results in the literature, as well as point out the issues that have been observed in the literature.

Network structure

Intuitively, star graphs or similar structures are the most robust networks against random removals of edges or nodes, because a very small proportion (in a pure star graph, only one) of its nodes is vital, and thus, for any removal, the probability of it being picked is $1/N'$.

Scale free networks have similar structure to star graphs. They have a few central (core) nodes, and many peripheral nodes. Scale free networks have been shown to be robust to random attacks (Albert et al. 2000, Cohen et al. 2000, Cruciti et al. 2003), to the point that the changes in their global and local efficiency are barely detectable for a few removals (Cruciti et al. 2003). However, they are vulnerable to targeted degree attacks (Albert et al. 2000, Cohen et al. 2001, Cruciti et al. 2003).

Furthermore, for attacks with incomplete information on the network, very small increases in the knowledge of the network can increase the effectiveness of the attack a lot (Wu et al. 2007). However, hiding just a small fraction of nodes can prevent the network to break down under an intentional attack to the hubs. This is a surprising result, as randomly hiding a fraction of nodes in a scale-free network, should correspond to hiding

preferentially low-degree nodes, and thus, one should expect an attack targeting the hubs to still be able to detect most of them. It can be explained that for scale-free networks with inhomogeneous degree distributions, there are few highly connected hubs which dominate a network, so even hiding a few of them can protect the whole network (Wu et al. 2007).

Core-periphery hierarchies and multiscale (high connectivity both at local and at global scale) hierarchies are more robust to congestion, but as the size of the network grows, layered hierarchies (the layers are well connected communities and the communities form a hierarchical network) get more robust (Dodds et al. 2003).

Erdos-Renyi networks have been shown to be among the most robust models as they have no structural bias (Holme et al. 2002). This agrees with Albert et al. 2000, who showed that random attacks and degree targeting attacks have similar effects on these networks, as well as Wagner 2015 who showed that they are more robust than scale free or small world networks against targeted attacks.

Networks with homogeneous degree distributions, are very robust against cascades caused by a single attack in networks with processability limitations (Motter and Lai 2002). They are also more robust against combinations of degree and random attacks, as according to Valente et al. 2004, the configurations that maximize the percolation threshold under such combined attacks have at most three distinct node degrees (d_1, d_2 and d^*, with d^* being the largest degree a node may have after a fraction f_a of the most connected nodes in the network has been removed). This agrees with Tanizawa et al. 2005, who showed that a 2-value degree distribution is the most robust to combined degree and random attacks.

2-peak and 3-peak degree distribution networks get the robustness similar to scale-free networks against random attacks, while lacking large structural biases against random attacks. However, they have not been tested under other methods of attack, especially recalculated measures.

Estrada et al. 2006 pointed out that networks with good expansion properties are more robust as they have less bottlenecks. Later, Estrada 2007, ranked networks having and lacking either a good expansion or a non-skewed degree distribution as follows:

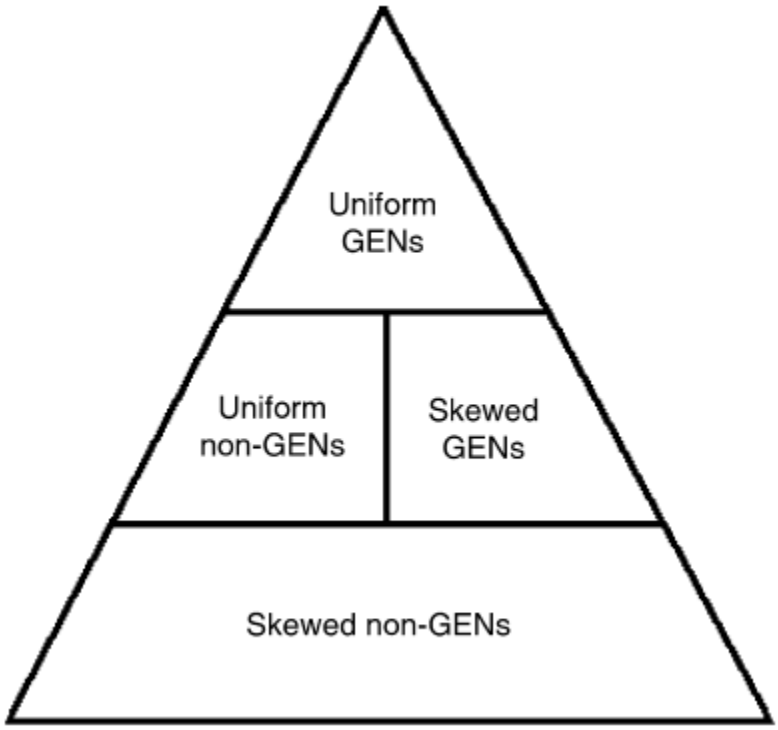

Figure 6.2.1: A simplified representation of the robustness in undirected food webs. At the bottom are the networks with power law or exponential degree distributions which lack good expansion property, they are the most vulnerable of the networks. In the middle, the networks are either skewed good expanders, or non-skewed non-good expanders. At the top are the networks with both non-skewed degree distribution and good expansion properties, these are the most robust.
Estrada et al. 2007.

Real world networks come from all backgrounds, but they are usually shown to be robust. Buhl et al. 2004 showed ant galleries to be robust to random failures. Estrada 2006 showed that naturally evolving networks have good expansion properties more frequently than technological networks. Newman and Goshal 2007 observed that real world networks appear to have large, and usually exceptionally robust giant bicomponents. Memmott et al. 2004 showed that bipartite plant-pollinator networks are generally robust, but Santamaria et al. 2014 showed that they span a wide range of robustness and are not all robust.

Criminal networks have been shown to be vulnerable to the removal of links that connect communities (attacks that target betweenness) (Xu and Chen 2008). Duijn et al. 2014 also studied such networks, with the addition of self-rewiring properties. In this study, criminal networks were shown to be more robust than expected, and even got more efficient when subjected to attacks that left their value chain (role structure) intact.

Air traffic networks, studied with structural limitations and rewiring properties, have been shown to be vulnerable to natural hazards due to many hub airports being close together in the geographical center, thus making them vulnerable to disproportionately small spatial hazards (they are technically core-periphery networks, not on their connectivity, but on the geography of the physical world) (Wilkinson et al. 2012). The possibility of a re-wiring algorithm is suggested as a solution as most structural changes are unrealistic due to the nature of the networks, as countries both need and want hub airports (Wilkinson et al. 2012). However, a well-designed network with the same limitations can be more robust than a network with a good rewiring mechanism (Dunn and Wilkinson 2015).

Subway networks have been shown to be robust to random failures and most vulnerable to betweenness and degree attacks (Zhang et al. 2011, Yin et al. 2016, Zhang et al. 2018). This is expected as they generally are sets of path and cycle graphs with a few common nodes spread around. This turns them to low-connectivity bimodal-like structures, where betweenness and degree attacks target exactly these common nodes, or their bridge neighbors.

Finally, two very interesting results for real world networks are that they do not behave very similarly to modelled networks (Holme et al. 2002), and that, in weighted networks, taking the weights into account when attacking the network, might reduce the effectiveness of the attack (Kaiser-Bunbury et al. 2010).

Change in the efficiency and change in the giant component give similar results in ranking attacks in most cases (Holme et al. 2002).

Vertex and edge connectivity are very loos bounds to algebraic connectivity, especially for large networks, as can be seen from (figures 3.5.1, 3.5.2 and 3.5.3, Jamakovic and Uhlig 2007). This is generally expected, as the addition of a node with equal to or slightly larger than average vertex and edge connectivity, does not reduce these values but might reduce the total average robustness of the network.

Although [undirected] algebraic connectivity is unable to consistently capture the robustness of networks, it provides an avenue to easily capture the robustness of networks with non-aparent elements utilizing key components of the Laplacian spectrum (Sydney et al. 2008).

Natural connectivity is negatively correlated with efficiency (Peng et al. 2016).

Schieber et al. 2015 noted that although JS-robustness can be used with any probability distribution, the use of distances has shown to be more consistent in capturing structural deviations. In addition, the distance probability distribution is able to acknowledge disconnected pairs of nodes. Heuristic algorithms are proposed, but the errors of these algorithms can be important as for $N = 10^7$ the changes in the info-theoretic robustness are of the order of 10^{-15} (Schieber et al. 2015).

SSSK-elasticity and degree assortativity are positively correlated (Sydney et al. 2008). Newman 2002 pointed out that degree assortativity is negatively corelated with network robustness against random attacks, but positively correlated against degree attacks. This agrees with Iyer et al. 2013 who showed that high values of negative degree assortativity seem to indicate vulnerability to degree attacks. This comes in contrast to Vazquez and Moreno 2003, who found assortative power law networks to be more robust than non-

assortative even at random failures. It also comes in contrast with Zhou et al. 2012 who found degree assortativity to be negatively correlated with robustness overall.

Another important remark, although not directly relevant to this work is by Scala and D'Agostino 2012, who claimed that disassortative networks are more robust since they have a higher failure threshold, but in assortative networks there is more time for intervention before total breakdown.

There is disagreement on whether degree assortativity is positively or negatively correlated with the robustness of networks. This requires further extensive research. Further research is also called to examine other forms of degree assortativity.

Ellens 2011 showed that a variety of measures (vertex/edge/algebraic connectivity, reliability polynomial, clustering coefficient and number of spanning trees) do not detect all edge additions/deletions. Furthermore, maximum betweenness may increase with edge addition, thus making the network seem more vulnerable. This final observation can be extended to congestion centrality and other flow measures.

Ellens 2011 and Ellens et al. 2011 showed that normalized effective resistance gives the same information as ER, but is a measure of robustness instead of vulnerability. Also, as its values are in [0,1], it can be used comparatively between networks.

Finally, Van der Meer 2012 showed that many robustness indices will order various graph structures differently in terms of how robust they are (edge connectivity, average distance, efficiency, clustering coefficient, algebraic connectivity, number of spanning trees, effective resistance, natural connectivity, percolation limit, resilience factor, graph diversity).

Methods of attack

Random attacks are very ineffective against scale free networks (Albert et al. 2000, Cohen et al. 2000). They have also been found to be less effective than degree attacks in many real-world networks (Sole and Montoya 2001, Buhl et al. 2004).

Attacks targeting the highest degree nodes are very effective against scale free networks (Albert et al. 2000). In real networks (NA power grid, targeting specific nodes) they are very effective for relatively large (larger than 8% of the nodes) attacks.

However, some cases of plant-pollinator bipartite networks, with structural limitations, have been found to be more vulnerable to attacks that target degree in ascending order, as such nodes are more specialized and can cause secondary extinctions. Nevertheless, they found high degree attacks to also be very effective (Memmott et al. 2004, Santamaria et al. 2014). Similar results have been shown by Dunne et al. 2002 in food webs, where they ranked node attacks generally as such: high degree > random > low degree, with the low degree strategy sometimes being very effective.

Allesina and Pascual 2009 found eigencentrality attacks to be better than degree attacks in various food webs.

Attacks targeting the highest betweenness nodes have been shown to be effective against Criminal networks (Xu and Chen 2008). They are less effective (on scale free, exponential degree distribution and some real-world networks) than degree attacks, although in assortative networks, for the first part of the attack (approximately 25% nodes) betweenness is better than degree (Iyer et al. 2013). Also, (in scale free and exponential degree distribution networks) re-estimated betweenness is similar to re-estimated closeness attacks and slightly better than re-estimated degree or eigenvector attacks, and in real world networks it is significantly better. This agrees with Holme et al 2002, who ranked the effectiveness of nodal attacks as such: re-estimated betweenness > re-estimated degree > initial degree > initial betweenness.

Attacks targeting the highest betweenness edges, are in most cases more effective over those targeting the highest edge-degree edges, both for their initial and recalculated values (Holme et al. 2002).

Albert et al. 2004 showed that removing transmission stations (nodes, detected by high voltage transferring) in the NA power grid, in descending order of the load that passes through them is more effective than degree initially (<8% nodes removed). But it is by far the most effective method when the measures are recalculated (every 10 steps). Load attacks where also found to be very effective in causing cascades in scale free, homogeneous and some real-world networks (Motter and Lai 2002).

Garcia-Algarra et al. 2017 showed in plant-pollinator, bipartite networks that the k-risk and k-degree attacks are better than other traditionally used ranking methods (degree, betweenness eigenvector). Furthermore, K-degree is exceptionally good at destroying the giant component of the network.

Dominguez-Garcia and Munoz 2015 showed in mutualistic species bipartite networks that MusRank attacks approach near optimality (compared to attacks generated via genetic algorithm) and far better than closeness, eigenvector, betweenness, degree, nestedness and pagerank. It also has the smallest variance in its effectiveness. This aggrees with Garcia-Algarra et al. 2017 who showed it to be the most effective method at ranking pollinators to cause secondary extinctions to plants.

Morone and Makse 2015 found that the best results for collective influence are found at 3 and 4 steps. Targeting high collective influence nodes appears to be a more efficient than re-estimated degree in the Mexican telecommunications network (scale free) and twitter network (unspecified structure).

Chen et al. 2008 showed the equal graph partitioning attack strategy to be better than high degree attacks (both initial and re-estimated) in scale free and Erdos-Renyi networks.

6.3. Comparison with our findings

In our own example, we have found food webs to be vulnerable to degree attacks. More specifically, as we studied them as directed networks, they were vulnerable to total

degree and in-degree attacks. Our results agree with Dunne et al. 2002, in that high degree attacks are better than random. However, they disagree with Allesina and Pascual 2009, as we found two out of the three degree-based attacks to be more effective than the eigencentrality attacks, although eigencentrality attacks are, in general, effective as well.

We also found the principles suggested by Estrada 2006 and Estrada 2007 to hold true; networks with a larger difference between their first two larger eigenvalues were indeed more robust in general. In addition, the natural connectivity proposed by Wu et al 2008 to provide a near-identical ranking of the networks according to their robustness when the impact of the attacks was measured on their global efficiency.

Attacks based on collective influence where not as effective as was expected, especially in the fourth network tested, given their effectiveness shown in Morone and Makse 2015.

7. Conclusions

A lot of effort has been invested in researching methods to measure a networks robustness, and every approach has advantages and disadvantages.

Studying measures derived only from the network is a general approach that can give a direct insight for a networks robustness, as well as being able to compare different networks. On the other hand, different methods of attack can have extremely different effects on a network. Thus, such a measure will always be a generalization, containing less information than any study that actually tests the impact of various attack methods on the network.

The disadvantages of studying attack-dependent measures outline the advantages of studying network-based measures. The tradeoff comes firstly from the computational complexity of attack dependent measures, as a lot of instances of the network have to be measured, while network-dependent measures only need to be computed once. Secondly, attacks on a network can be unpredictable. There are many efficient ways to disconnect a network, and many of them are very different to one another, it is nigh impossible to predict all the possibilities.

The advantages and disadvantages of attack-dependent measures are amplified in the class of worst case scenario measures (excluding simplistic measures as the vertex and edge connectivity). On one hand, they give a very informative estimate of the impact an attack can have on the network, but on the other, calculating all possible attacks that create disconnections is a very difficult task for large networks. Moreover, they omit an aspect of the problem; how many different worst-case (or near-worst case) attacks are there, as finding a close to perfect attack when there are very few good ones can be a very difficult task, while finding it among a much greater set heuristically may not be as hard.

The computational complexity is not a problem only in this kind of measure. Total graph diversity for example, requires the calculation of every possible path between every

pair of nodes in the network, and the proportion of non-overlapping paths. Something which would yield a probably ideal measure, but is extremely taxing computationally.

An important problem is the lack of consistent testing. The ability of a lot of measures to portray the robustness of networks, have been tested only in specific sets of networks, and under specific attacks. For example, the result that hiding a small part of a scale free network can protect it very effectively (Wu et al. 2007), might be biased by the existence of a few low degree nodes linking a lot of high degree ones, which is not necessarily a property for scale free networks. Furthermore, some indices of robustness, such as effective resistance, or results concerning such indices, are not tested at all. For example, Ghosh and Boyd 2006 showed that optimizing the algebraic connectivity requires relatively few edge additions, but never showed if these edge additions affect the rate that the network degrades under attack.

Peng et al. 2016, showed that natural connectivity is negatively correlated with efficiency. We consider the cause of this to be that the most efficient structures are star-like or core-periphery like networks, while natural connectivity is a measure based on the closed walks of all nodes, thus preferring more homogeneous networks.

The findings of Van der Meer 2012 show that different robustness indices measure different aspects of the networks. This shows that there is a lot of research needed in order to identify what are the different elements these indices measure. The results in chapter 5 agree with this, as we found some variation in the rankings of the networks according to the effects of the attacks, but the natural connectivity seemed to predict almost exactly the ranking of the effect on the efficiency of the network of almost all methods of attack.

Finally, concerning the rankings of collective influence as an attack method, our findings were greatly contrasted by those of Morone and Makse 2015. We believe the cause of this to be either the order of the networks, as collective influence was shown to be extremely effective by Morone and Makse 2015, but in networks with a number of vertices in the order of 10^7, or the structure of the networks, as they were shown to be that effective mostly in scale free networks. This is another point that require further research.

7.1. Discussion

We would like to draw attention to the lack of attack-independent indices of network robustness. In fact, algebraic connectivity seems to be the only directly available measure, and it has issues with networks that are not strongly connected. The other measures that were available were the natural connectivity, after we modified its method of calculation and assortativity, which requires further research before it can be used individually as a measure of robustness. Total effective resistance has not been tested in undirected networks, although theoretically it shows promise, the work done by Young et al. 2016a and Young et al. 2016b, opens another field of research here.

Another critical issue is that the results on network robustness lack coherence. Findings that are widely accepted are sometimes disregarded, while a variety of avenues remains unexplored due to most of the research being specifically targeted on other projects, or because of conflicting findings. Most studies refer to few real networks and to a handful of models. This may lead to bias. For this reason, we believe the creation of a large enough database of networks, with various combinations of properties, is required, in order for most of the knowledge in the robustness of networks to be thoroughly tested.

Beside the lack of testing in networks of various structures, there is a lack of consideration in the choosing of most methods of attack. Ranking correlations (such as centrality correlations) are an important aspect that has been widely ignored in the planning of network attacks. We believe that in order to find an efficient method of attack, one needs to know how similar the methods of attack are. It would, therefore, be an interesting point of research to study the correlation of the various measures according to the topology of a graph. Such correlations have, only partly, been explored by Kasthurinathna et al. 2013.

In conclusion, the answer that can be given to our original questions, is that further research is required, and that, currently, it seems more informative to examine the impact various attacks have on a network, than to attempt to quantify its robustness.

Originality

There have been some overviews of network robustness indices before, but they usually focus to specific measures, or groups of measures. Our work is the most complete collection of indices of robustness and methods of attack. Furthermore, it contains one of the very few attempts to study a robustness measure for directed networks, as well as to understand which aspect of robustness of the network it represents. Our study of the global efficiency normalized by the order of the initial network, allows us to compare networks of different orders, and mitigates the increasing impact of node removals on the network as the attack progresses.

Further research

This work is a survey of robustness indices and methods of attack, with respect to network structure. A logical next step is to create a thorough database, containing many networks that have as many combinations of properties as possible, in order to create a set of samples big and diverse enough to enable more thorough network studies.

In terms of methods of attack, the next steps would be (a) to examine the correlations in the various methods of vertex and edge rankings. (b) To investigate probabilistic attacks and attacks with incomplete information individually and more thoroughly, following the work of Wu et al. 2007. (c) Double down on the research of combined attacks, especially by studying the correlations previously mentioned.

In terms of indices of robustness, the next steps would likely be (a) a thorough sensitivity analysis of the existing ones in undirected networks. (b) Exploring the correlations of the various robustness indices further in an attempt to understand the type of structures and resistance to attacks they represent, especially in comparison with the various methods of attacks. (c) The generalization of those indices to directed, weighted and eventually signed networks, both in and out of the various possible contexts.

Finally, the transition of everything contained in this research, as well as the further research proposed, in belief propagation, knowledge transferring, or disease spreading networks, generalizing the ideas of robustness to true disconnection, to robustness to real-world changes.

Bibliography

Albert R., Barabasi A.L. (2002) Statistical mechanics of complex networks. Reviews of Modern Physics 74, 47. https://doi.org/10.1103/RevModPhys.74.47

Albert R., Hawoong J., Baarabasi A.L. (2000) Error and attack tolerance of complex networks, Letters to Nature 406, 378–382. https://doi.org/10.1038/35019019

Albert R., Jeong H., Barabasi A.L. (1999) Internet: Diameter of the World-Wide Web. Nature 401, 130-131. https://doi.org/10.1038/43601

Allesina S. Pascual M. (2009) Googling food webs: can an eigenvector measure species' importance for coextinctions? PLoS Computational Biology. https://doi.org/10.1371/journal.pcbi.1000494

Arulselvan A. Commander C.W. Elefteriadou L. Pardalos P.M. (2008) Detecting critical nodes in sparse graphs. Computers & Operations Research 36, 2193-2200. https://doi.org/10.1016/j.cor.2008.08.016

Aslan E., Bacak-Turan G. (2016) Mean rupture degree of graphs. UPB Scientific Bulletin, Series A: Applied Mathematics and Physics 78, 233-242. https://www.researchgate.net/publication/305521464_Mean_rupture_degree_of_graphs

Bagga K.S. Beineke L.W. Goddard W.D. Lipman M.J. Pippert R.E. (1992) A survey of integrity. Discrete Applied Mathematics, 37–38, 13-28. https://doi.org/10.1016/0166-218X(92)90122-Q

Bagga K.S. Beineke L.W. Lipman M.J. Pippert R.E. (1994) Edge-Integrity: a survey. Discrete Mathematics 124, 3-12. https://doi.org/10.1016/0012-365X(94)90084-1

Baker W.E., Faulkner R.R. (1993). The social organization of conspiracy: Illegal networks in the heavy electrical equipment industry. American Sociological Review, 58, 837-860. https://doi.org/10.2307/2095954

Baras J.S. Hovareshti P. (2009) Efficient and Robust Communication Topologies for Distributed Decision Making in Networked Systems. Proceedings of the joint 48th IEEE Conference on Decision and Control (CDC) and 2009 28th Chinese Control Conference, 3751-3756. https://doi.org/10.1109/CDC.2009.5400448

Bastolla U., Fortuna M.A., Pascual-Garcia A., Ferrera A., Luque B., Bascompte J. (2009) The architecture of mutualistic networks minimizes competition and increases biodiversity. Nature 458, 1018-1020. https://doi.org/10.1038/nature07950

Batagelj V., Mrvar A. (2006): Pajek datasets. http://vlado.fmf.uni-lj.si/pub/networks/data

Bates D., Maechler M. (2017). Matrix: Sparse and Dense Matrix Classes and Methods. R package version 1.2-12. https://CRAN.R-project.org/package=Matrix

Beineke L.W. Goddard W.D. Hamburger P. Kleitman D.J. Lipman M.J. Pippert R.E. (1991) The integrity of the cube is small. Journal of Combinatorial Mathematics and Combinatorial Computing 9. https://www.researchgate.net/publication/240317606_The_Integrity_of_the_Cube_is_Small

Bonacich, P. (1987) Power and Centrality: A Family of Measures. *American Journal of Sociology* 92, 1170-1182. https://doi.org/10.1086/228631

Buhl J. Gautrais J. Sole R.V. Kuntz P. Valverde S. Deneubough J.L. Theraulaz G. (2004) Efficiency and robustness in ant networks of galleries. European Physical Journal B 42, 123-129. https://doi.org/10.1140/epjb/e2004-00364-9

Chen Y., Paul G., Havlin S., Liljeros F., Stanley H.E. (2008) Finding a better immunization strategy. Physical review letters 101. https://doi.org/10.1103/PhysRevLett.101.058701

Chvátal V. (1973) Tough graphs and Hamiltonian circuits. Discrete Mathematics 5, 215-228. https://doi.org/10.1016/0012-365X(73)90138-6

Cohen R. Erez K. ben-Avraham D. Havlin S. (2000) Resilience of the internet to random breakdowns. Physical Review Letters 85. https://doi.org/10.1103/PhysRevLett.85.4626

Cohen R. Erez K. ben-Avraham D. Havlin S. (2001) Breakdown of the internet under intentional attack, Physical Review Letters 86. https://doi.org/10.1103/PhysRevLett.86.3682

Colbourn C.J., Harms D.D., Myrvold W.J. (1993) Reliability Polynomials Can Cross Twice, Journal of the Franklin Institute 330, 629-633. https://doi.org/10.1016/0016-0032(93)90102-Z

Costa L.da F. (2004) Reinforcing the resilience of complex networks. Physical Review E 69. https://doi.org/10.1103/PhysRevE.69.066127

Cozzens, M.B., Moazzami, D. and Stueckle, S. (1995) The Tenacity of a Graph. Graph Theory. Proceedings of the Seventh International Conference on the Theory and Applications of Graphs.

Crucitti P., Latora V., Marchiori M., Rapisarda A. (2003) Efficiency of scale-free networks: error and attack tolerance. Physica A 320, 622-642. https://doi.org/10.1016/S0378-4371(02)01545-5

Csardi G., Nepusz T. (2006): The igraph software package for complex network research. InterJournal, Complex Systems 1695. http://igraph.org

Deng Y. Wu J. (2015) Optimal attack strategy based on limited cost model on complex networks. IEEE International Conference on Systems, Man, and Cybernetics. https://doi.org/10.1109/SMC.2015.31

Deng Y. Wu J. (2016) Optimal Attack Strategy with Heterogeneous Costs in Complex Networks. Annual IEEE Systems Conference (SysCon). https://doi.org/10.1109/SYSCON.2016.7490518

Dodds P.S. Watts D.J. Sabel C.F. (2003) Information exchange and the robustness of organizational netoworks. Proceedings of the National Academy of Sciences of the

United States of America 100, 12516-12521. https://doi.org/10.1073/pnas.1534702100

Domiunguez-Garcia V., Munos M. (2015) Ranking species in mutualistic networks. Scientific Reports 5. https://doi.org/10.1038/srep08182

Duijin P.A.C., Kashirin V., Sloot P.M.A. (2014) The Relative Ineffectiveness of Criminal Network Disruption. Scientific reports 4. https://doi.org/10.1038/srep04238

Dunkum M. Lanphier D. (2014) Vulnerability of nearest neighbor graphs. Discrete Applied Mathematics 171, 42–52. https://doi.org/10.1016/j.dam.2014.02.021

Dunn S. Wilkinson S.M. (2015) Increasing the resilience of air traffic networks using a network graph theory approach. Transportation Research Part E: Logistics and Transportation Review 90, 39-50. https://doi.org/10.1016/j.tre.2015.09.011

Dunne J.A., Williams R.J., Martinez N.D. (2002) Network structure and biodiversity loss in food webs: robustness increases with connectance. Ecology Letters 5, 558-567. https://doi.org/10.1046/j.1461-0248.2002.00354.x

Ellens W. (2011) Effective resistance and other graph measures. Master Thesis in the Mathematical Institute of the University of Leiden. https://www.universiteitleiden.nl/binaries/content/assets/science/mi/scripties/ellensmaster.pdf

Ellens W. Spieksma F.M. Van Mieghem P. Jamakovic A. Kooij R.E. (2011) Effective graph resistance. Linear Algebra and its Applications 435, 2491-2506. https://doi.org/10.1016/j.laa.2011.02.024

Ellens W. Kooij R.E. (2013) Graph measures and network robustness. arXiv:1311.5064

Estrada E. (2006) Network robustness to targeted attacks. The interplay of expansibility and degree distribution. The European Physical Journal B - Condensed Matter and Complex Systems 52, 563-574. https://doi.org/10.1140/epjb/e2006-00330-7

Estrada E. (2007) Food webs robustness to biodiversity loss: The roles of connectance, expansibility and degree distribution, Journal of Theoretical Biology 244, 296-307. https://doi.org/10.1016/j.jtbi.2006.08.002

Erdos P., Renyi A. (1959) On Random Graphs I. Publicationes Mathematicae 6, 290-297.

Fiedler M. (1973) Algebraic connectivity of Graphs. Czechoslovak Mathematical Journal 23, 298-305.

Freeman L.C. (1979) Centrality in Social Networks, Conceptual clarification, Social Networks 1, 215-239. https://doi.org/10.1016/0378-8733(78)90021-7

Garcia-Algarra J., Pastor J.M., Iriondo J.M., Galeano J. (2017) Ranking of critical species to preserve the functionality of mutualistic networks using the k-core decomposition. PeerJ 5. https://doi.org/10.7717/peerj.3321

Ghosh A. Boyd S. (2006) Growing Well-Connected Graphs, Proceedings of the 45th IEEE Conference on Decision and Control 6605-6611. https://doi.org/10.1109/CDC.2006.377282

Gilbert E.N. (1959) Random Graphs. Annals of Mathematical Statistics 30, 1141-1144. https://doi.org/10.1214/aoms/1177706098

Goddard W. (1994) Measures of Vulnerability – The integrity Family, Networks 24, 207-213. https://doi.org/10.1002/net.3230240403

Goddard W. Sward H.C. (1990) The Integrity in Graphs: Bounds and Basics. Journal of Combinatorial Mathematics and Combinatorial Computing 7.

Grubesic T.H., MAtisziw T.G., Ripley D.A.J. (2011) Approximating the geographical characteristics of Internet activity. Journal of Urban Technology 18, 51-71. https://doi.org/10.1080/10630732.2011.578409

Harary F. (1983) Conditional Connectivity. Networks 13, 347-357. https://doi.org/10.1002/net.3230130303

Holme P., Beom J.K., Chang N.Y., Seung K.H. (2002) Attack vulnerability of complex networks. Physical Review E 65. https://doi.org/10.1103/PhysRevE.65.056109

Holme P., Kim J.B. (2002) Growing scale-free networks with tunable clustering. Physical Review E 65. https://doi.org/10.1103/PhysRevE.65.026107

Iyer S., Killingback T., Sundaram B., Wang Z. (2013) Attack Robustness and Centrality of Complex Networks. PLoS One. https://doi.org/10.1371/journal.pone.0059613

Jamakovic A. Uhlig S. (2007) On the relationship between the algebraic connectivity and graph's robustness to node and link failures. Next Generation Internet Networks 96-102. https://doi.org/10.1109/NGI.2007.371203

Jonhson S., Dominguez-Garcia V., Munoz M.A. (2013) Factors determining nestedness in complex networks. PLoS ONE. https://doi.org/10.1371/journal.pone.0074025

Jung H.A. (1978) On a Class of Posets and the Corresponding Comparability of Graphs. Journal of Combinatorial Theory, Series B 24, 125-133. https://doi.org/10.1016/0095-8956(78)90013-8

Jung E. Gouda M.G. (2006) Vulnerability analysis of certificate graphs. International Journal of Security and Networks 1, 13-23. https://doi.org/10.1504/IJSN.2006.010820

Kaiser-Bundury C.N., Muff S., Memmott J., Muller C.B., Caflisch A. (2010) The robustness of pollination networks to the loss of species and interactions: a quantitative approach incorporating pollinator behaviour. Ecology Letters 13, 442-452. https://doi.org/10.1111/j.1461-0248.2009.01437.x

Kasthurinathna D., Piraveenan M., Thedchanamoorthy G. (2013) Network robustness and topological characteristics in scale-free networks. IEEE Conference on Evolving and Adaptive Intelligent Systems. https://doi.org/10.1109/EAIS.2013.6604114

Kelmans A.K. (1981) On graphs with randomly deleted edges. Acta Mathematica Academiae Scientarum Hungarica 37, 77-88. https://doi.org/10.1007/BF01904874

Kirlangic A. (2002) A measure of graph vulnerability: scattering number. International Journal of Mathematics and Mathematical Sciences 30, 1-8. https://doi.org/10.1155/S0161171202012607

Kirlangic A. (2009) The rupture degree and gear graphs. Bulletin of the Malaysian Mathematical Sciences Society, Series 2, 32, 31-36.

Klein D.J., Randic M. (1993) Resistance Distance. Journal of Mathematical Chemistry 12, 81-95. https://doi.org/10.1007/BF01164627

Krebs V.E. (2002) Uncloacking Terrorist Networks. First Monday 7(4). http://firstmonday.org/ojs/index.php/fm/article/view/941/863">Network

Krishnamoorthy M.S. (1987) Fault diameter of interconnection networks. Computers & Mathematics with Applications 13, 577-582. https://doi.org/10.1016/0898-1221(87)90085-X

Latora V., Marchiori M. (2004) How the science of complex networks can help developing strategies against terrorism. Chaos, Solitions and Fractals 20, 69-75. https://doi.org/10.1016/S0960-0779(03)00429-6

Laskar R. Stueckle S. Piazza B. On the edge-integrity of some graphs and their complements. Discrete Mathematics 122, 245-253. https://doi.org/10.1016/0012-365X(93)90299-9

Li F., Ye Q., Sheng B. (2009) On Integrity of Harary Graphs. International Conference on Combinatorial Optimization and Applications 269-278. https://doi.org/10.1007/978-3-642-02026-1_25

Li Y. Zhang S. (2010) Extremal graphs with given order and the rupture degree. Computers and Mathematics with Applications 60, 1706-1710. https://doi.org/10.1016/j.camwa.2010.07.001

Li Y. Zhang S. Li X. (2005) Rupture degree of graphs. International Journal of Computer Mathematics 82, 793-803. https://doi.org/10.1080/00207160412331336062

Li Y. Zhang S. Zhang Q. (2008) Vulnerability parameters of split graphs. International Journal of Computer Mathematics 85, 19-23. https://doi.org/10.1080/00207160701365721

Lin J. (1991) Divergence measures based on the Shannon entropy. IEEE Transactions on Information Theory 37, 145-151. https://doi.org/10.1109/18.61115

Maas C. (1987) Transportation in graphs and the admittance spectrum. Discrete Applied Mathematics 16, 31-49. https://doi.org/10.1016/0166-218X(87)90052-7

Mahde S.S. Mathad V. Sahal A.M. (2010) Hub-Integrity of Graphs. Bulletin of the Intenational Mathematical Virtual Institute 5. https://www.researchgate.net/publication/308901018_HUB-INTEGRITY_OF_GRAPHS

Matisziw T.C. Grubesic T.H. Guo J. (2012) Robustness Elasticity in Complex Networks. PLOS ONE. https://doi.org/10.1371/journal.pone.0039788

Memmott J., Waser N.M., Price M.V., (2004) Tolerance of pollination networks to species extinctions. Proceedings of the Royal Society B 271, 2605–2611. https://doi.org/10.1098/rspb.2004.2909.

Moazzami D. (1999) Vulnerability in Graphs - a Comparative Survey. Journal of Combinatorial Mathematics and Combinatorial Computing 30, 23-32.

Moazzami D. (2008) On the Edge-Tenacity of Graphs. International Mathematical Forum 3. https://www.researchgate.net/publication/241594595_On_the_Edge-Tenacity_of_Graphs

Moazzami D. (2010) Tenacity of a graph with maximum connectivity. Discrete Applied Mathematics 159, 367-380. https://doi.org/10.1016/j.dam.2010.11.008

Moazzami D. (2011) An algorithm for the Computation of Edge Integrity, I'(T). International Journal of Contemporary Mathematical Sciences 6, 507-516.

Moazzami D. Salehian B. (2008) On the Edge-Tenacity of Graphs, International Mathematical Forum 3.

https://www.researchgate.net/publication/241594595_On_the_Edge-Tenacity_of_Graphs

Mohar B. (1989) Isoperimetric Numbers of Graphs. Journal of Combinatorial Theory 47, 274-291. https://doi.org/10.1016/0095-8956(89)90029-4

Moore E.F., Shannon C.E. (1956) Reliable Circuits Using Less Reliable Relays. Journal of the Franklin Institute 262, 191-208. https://doi.org/10.1016/0016-0032(56)90559-2

Moreno Y. Pastor-Satorras R. Vasquez A. Vespignani A. (2003) Critical load and congestion instabilities in scale-free networks. Europhysics Letters 62. https://doi.org/10.1209/epl/i2003-00140-7

Morone F., Makse H.A. (2015) Influence maximization in complex networks through optimal percolation. Nature 524, 65-8. https://doi.org/10.1038/nature14604

Morone F., Min B., Bo L., Mari R., Makse H.A. (2016) Collective influence algorithm to find influencers via optimal percolation in massively large social media. Scientific Reports 6. https://doi.org/10.1038/srep30062

Motter A.E. Lai Y.C. (2002) Cascade-based attacks on complex networks. Physical Review E 66. https://doi.org/10.1103/PhysRevE.66.065102

Newman M.E.J. (2001) Clustering and preferential attachment in growing networks. Physical Review E 64. https://doi.org/10.1103/PhysRevE.64.025102

Newman M.E.J. (2003) Mixing patterns in networks. Physical review E 67. https://doi.org/10.1103/PhysRevE.67.026126

Newman M.E.J. Choshal G. (2007) Bicomponents and the robustness of networks to failure. Physical Review Letters 100. https://doi.org/10.1103/PhysRevLett.100.138701

Page L.B. Perry J.E. (1994) Reliability Polynomials and Link Importance in Networks. IEEE Transactions on reliability 43, 51-58. https://doi.org/10.1109/24.285108

Peng G-S. Tan S-Y. Wu J. Holme P. (2016) Trade-offs between robustness and small-world effect in complex networks, Scientific Reports 6. https://doi.org/10.1038/srep37317

R Core Team (2017). R: A language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. https://www.R-project.org/

Raab J., Milward H.B. (2003). Dark networks as Problems. Journal of Public Administration Research and Theory 13, 413-439. https://doi.org/10.1093/jpart/mug029

Reingold O. (2008) Undirected connectivity in log-space. Journal of the ACM 55. https://doi.org/10.1145/1391289.1391291

Rohrer J.P., Sterbenz J.P.G. (2011) Predicting Topology Survivability using Path Diversity. 2011 3rd International Congress on Ultra Modern Telecommunications and Control Systems and Workshops (ICUMT), 1-7. https://www.researchgate.net/publication/261045842_Predicting_topology_survivability_using_path_diversity

RStudio Team (2016) RStudio: Integrated Development for R. RStudio, Inc. Boston, MA. http://www.rstudio.com/

Salles R.M., Marino D.A.Jr. (2011) Strategies and Metric for Resilience in Computer Networks. The Computer Journal 55, 728-739. https://doi.org/10.1093/comjnl/bxr110

Santamaria S., Galeano J., Pastor J.M., Mendez M. (2014) Robustness of alpine pollination networks: effects of network structure and consequences for endemic plants. Arctic, Antarctic and Alpine Research 46, 568-580. https://doi.org/10.1657/1938-4246-46.3.568

Schieber T.A. Carpi L. Frery A.C. Rosso O.A. Pardalos P.M. Ravetti M.G. (2015) Information theory perspective on network robustness. Physics Letters A 380, 359-364. https://doi.org/10.1016/j.physleta.2015.10.055

Schneider C.M., Moreira A.A., Andrade J.S.Jr., Havlin S., Herrmann H.J. (2010) Mitigation of malicious attacks on networks. Proceedings of the National Academy of Sciences of the United States of America 108, 3838-3841. https://doi.org/10.1073/pnas.1009440108

Seidman S.B. (1983) Network structure and minimum degree. Social Networks 5, 269-287. https://doi.org/10.1016/0378-8733(83)90028-X

Sole R.V., Montoya J.M. (2001) Complexity and fragility in ecological networks, Proceedings of the Royal Society B 268, 2039-2045. https://doi.org/10.1098/rspb.2001.1767

Sydney A. Scoglio C. Schumm P. Kooij R. (2008) ELASTICITY: Topological Characterization of Robustness in Complex Networks. Bionetics 2008. ISBN: 978-963-9799-35-6. arXiv:0811.4040

Tacchella A., Cristelli M., Caldarelli G., Gabrielli A., Petronero L. (2012) A new metrics for countries' fitness and products' complexity. Scientific Reports 2. https://doi.org/10.1038/srep00723.

Tanizawa T. Paul G. Cohen R. Havlin S. Stanley H.E. (2005) Optimization of network robustness to waves of targeted and random attacks. Physical Review E 71. https://doi.org/10.1103/PhysRevE.71.047101

Tsen F.S.P. Sung T.Y. Lin M.Y. Hsu L.H. Myrvold W. (1994) Finding the most vital edges with respect to the number of spanning trees. IEEE Transactions on Reliability 43, 600-603. https://doi.org/10.1109/24.370220

Ulanowicz, R.E., C. Bondavalli and M.S. Egnotovich. (1997 Network Analysis of Trophic Dynamics in South Florida Ecosystems, FY 96: The Cypress Wetland Ecosystem. Ref. No. [UMCES]CBL 97-075. University of Maryland System, Chesapeake Biological Laboratory, Solomons, MD 20688-0038 USA.

Ulanowicz, R.E., C. Bondavalli and M.S. Egnotovich. (1998) Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 97: The Florida Bay Ecosystem. Ref. No. [UMCES]CBL 98-123. University of Maryland System, Chesapeake Biological Laboratory, Solomons, MD 20688-0038 USA.

Ulanowicz, R.E., C. Bondavalli and M.S. Egnotovich. (1999) Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 98: The Mangrove Ecosystem. Ref. No. [UMCES]CBL 99-0073. University of Maryland System, Chesapeake Biological Laboratory, Solomons, MD 20688-0038 USA.

Ulanowicz, R.E., C. Bondavalli and M.S. Egnotovich. (1999) Network Analysis of Trophic Dynamics in South Florida Ecosystem, FY 99: The Graminoid Ecosystem. Ref. No. [UMCES]CBL 00-0176. University of Maryland System, Chesapeake Biological Laboratory, Solomons, MD 20688-0038 USA.

Valente A.X.C.N. Sarkar A. Stone H.A. (2004) 2-Peak and 3-Peak Optimal Complex Networks. Physical Review Letters 92. https://doi.org/10.1103/PhysRevLett.92.118702

Van der Meer E. (2012) Comparing measures of network robustness. Master thesis at the VU University of Amsterdam. https://beta.vu.nl/nl/Images/werkstuk-meer_tcm235-280356.pdf

Van Mieghem P. (2010) Graph Spectra for Complex Networks. Cambridge University Press. https://doi.org/10.1017/CBO9780511921681

Van Mieghem P., Doerr C., Wang H., Herandez J.M., Hutchison D., Karaliopoulos M., Kooij R.E. (2010) A Framework for Computing Topological Network Robustness. Submitted for publication. https://www.nas.ewi.tudelft.nl/people/Piet/papers/RobustnessRmodel_TUDreport20101218.pdf

Vince A. (2004) The integrity of a cubic graph. Discrete Applied Mathematics 140, 223-239. https://doi.org/10.1016/j.dam.2003.07.002

Wagner S.J.F. (2015) Applying Flow Algorithms to Deconstruct Social Networks. Thesis submitted to the Graduate School of the Creighton University (Master of Arts, Department of Political Science). http://dspace.creighton.edu:8080/xmlui/handle/10504/68812

Wang H. Van Mieghem P. (2008) Algebraic connectivity optimization via link addition. Bionetics 2008. ISBN: 978-963-9799-35-6

Watts D.J., Strogatz S.H. (1998) Collective dynamics of 'small-world' networks. Nature 393, 440–442. https://doi.org/10.1038/30918

Wickham H. (2009) ggplot2: Elegant Graphics for Data Analysis. Springer-Verlag New York.

Wilkinson S.M. Dunn S. Ma S. (2011) The vulnerability of the European air traffic network to spatial hazards. Natural Hazards 60, 1027-1036. https://doi.org/10.1007/s11069-011-9885-6

Wu C.W. (2005) Algebraic connectivity of directed graphs. Linear and Multilinear Algebra 53, 203-223. https://doi.org/10.1080/03081080500054810

Wu J. Barahona M. Yue-Jin T. Deng H.Z. (2010) Natural Connectivity of Complex Networks. Chinese Physics Letters 27. https://doi.org/10.1088/0256-307X/27/7/078902

Wu J. Barahona M. Yue-Jin T. Deng H.Z. (2011). Spectral Measure of Structural Robustness in Complex Networks. IEEE Transactions on Systems, Man, and Cybernetics 41, 1244-1252. https://doi.org/10.1109/TSMCA.2011.2116117

Wu J. Deng H.Z. Tan Y.J. Zhu D.Z. (2007) Vulnerability of complex networks under intentional attack with incomplete information. Journal of Physics A: Mathematical and Theoretical 40. https://doi.org/10.1088/1751-8113/40/11/005

Wu J. Tan Y-J. Deng H.Z. Li Y. Liu B. Lv X. (2008) Spectral Measure of Robustness in Complex Networks. arXiv:0802.2564

Xu J, Chen H. (2003) Untangling criminal networks: A case study. International Conference on Intelligence and Security Informatics, 232-248. https://doi.org/10.1007/3-540-44853-5_18

Xu J., Chen H. (2008) The Topology of Dark Networks. Communications of the ACM 51, 58-65. https://doi.org/10.1145/1400181.1400198

Yin H., Han B,. Li D., Wang Y. (2016) Evaluating Disruption in Rail Transit Network: A Case Study of Beijing Subway. Procedia Engineering 137, 49-58. https://doi.org/10.1016/j.proeng.2016.01.233

Young G.F., Scardovi L., Leonard N.E. (2016) A New Notion of Effective Resistance for Directed Graphs - Part I: Definition and Properties. IEEE Transactions on Automatic Control 61, 1727-1736. https://doi.org/10.1109/TAC.2015.2481978

Young G.F., Scardovi L., Leonard N.E. (2016) A New Notion of Effective Resistance for Directed Graphs - Part II: Computing Resistances. IEEE Transactions on Automatic Control 61, 1737-1752. https://doi.org/10.1109/TAC.2015.2481839

Zhang J., Wang S., Wang X. (2018) Comparison analysis on vulnerability of metro networks based on complex network. Physica A: Statistical Mechanics and its Applications 496, 72-78. https://doi.org/10.1016/j.physa.2017.12.094

Zhang S. Wang Z. (2001) Scattering Number in Graphs. Networks 37, 102-106. https://doi.org/10.1002/1097-0037(200103)37:2<102::AID-NET5>3.0.CO;2-S

Zhang J., Xu X., Hong L., Wang S., Fei Q. (2011) Networked analysis of the Shanghai subway network in China. Physica A: Statistical Mechanics and its Applications 390, 4562-4570. https://doi.org/10.1016/j.physa.2011.06.022

Zhou S., Zhang G.Q., Zhang G.Q. (2007) Chinese Internet AS-level Topology. IET Communications.

INDEX