THESIS

Quantifying supply chain vulnerability using a multilayered complex network perspective

Nadia M. Viljoen

Ph.D. (Industrial Engineering)

Supervisor:
Prof Johan W. Joubert
Room 2-11, Engineering Building II
Department of Industrial and Systems Engineering
Faculty of Engineering, Built Environment, and Information Technology
University of Pretoria
Pretoria
0002
South Africa

Student personal information:
Nadia M. Viljoen
Room 2-8, Engineering Building II
Department of Industrial and Systems Engineering
Faculty of Engineering, Built Environment, and Information Technology
University of Pretoria
Pretoria
0002
South Africa
nadia.viljoen@up.ac.za
+27 (0)71 610 6051

February 1, 2018

Contents

List of Figures

List of Tables

Acronyms and abbreviations

BA	Barabási-Albert
CAS	Complex Adaptive Systems
CBD	Central Business District
CNT	Complex Network Theory
CoCT	City of Cape Town
DH	Double Hub
Elemental-B	Elemental link betweenness
EDF	Empirical Distribution Function
ER	Erdös-Rényi
ET	eThekwini Metropolitan Municipality
FC	Fully Connected
FON	First Order Neighbourhood
GPS	Global Positioning System
GT	Gauteng Province
IT	Information Technology
KS-test	Kolmogorov-Smirnov test
LCC	Largest Connected Component
LFSN	Largest Functional Subnetwork
MCGC	Mutually Connected Giant Component
OSM	OpenStreetMap
Overall-B	Overall link betweenness
Overall-S	Link salience
SCM	Supply Chain Management
SCRes	Supply Chain Resilience

SCRM	Supply Chain Risk Management
SCS	Supply Chain Systems
SCVI	Supply Chain Volatility Index
SH	Single Hub
VGI	Volunteer Geographic Information
WEF	World Economic Forum
WS	Watts-Strogatz

Abstract

Today's supply chains face increasing volatility on many fronts. From the shop-floor where machines break and suppliers fail to the boardrooms where unanticipated price inflation erodes profitability. Turbulence is the new normal.

To remain competitive and weather these (daily) storms, supply chains need to move away from an efficiency mindset towards a resilience mindset. For over a little more than a decade industry and academia have awakened to this reality. Academic literature and case studies show that there is no longer a shortage of resilience strategies and designs. Unfortunately, industry still lacks the tools with which to assess and evaluate the effectiveness of such strategies and designs. Without the ability to quantify the benefit it is impossible to motivate the cost.

This thesis adds one piece to the puzzle of quantifying supply chain vulnerability. Specifically, it focussed on supply chains within urban areas. It addresses the question:

> *"How does a supply chain's network design (internal configuration) and its dependence on the underlying road network (external circumstances) make it more or less vulnerable to disruptions of the road network?"*

Multilayered Complex Network Theory (CNT) held promise as a modelling approach that could capture the complexity of the dependence between a *logical* supply chain network and the *physical* road network that underpins it. This approach addressed two research gaps in complex network theory applications. In the supply chain arena CNT applications have reaped many benefits but the majority of studies regarded single-layer networks that model only supply chain relations. There were no studies found where the dependence of supply chain layers on underlying physical infrastructure was modelled in a multilayered manner. Road network applications offered many more multilayered applications but these primarily focussed on passenger transport, not freight transport.

The first artefact developed in the thesis was a multilayered complex network formulation representing a *logical* (supply chain) layer placed on a *physical* (road infrastructure) layer. The individual layers had predefined network characteristics and on their own could not hint at the inherent vulnerability that the system as a whole might have. From the multilayered formulation, the collection of shortest paths emerged. This is the collection of all shortest path alternatives within a network. The collection of shortest paths is the unique fingerprint of each multilayered network instance. The key to understanding vulnerability lies within the characteristics of the collection of shortest paths.

Three standard supply chain network archetypes were defined namely the Fully Connected (FC), Single Hub (SH) and Double Hub (DH) archetypes. A sample of 500 theoretical multilayered network instances was generated for each archetype. These theoretical instances were subjected to three link-based progressive targeted disruption simulations to study the vulnerability characteristics of the collection of shortest paths. Two of the simulations used relative link betweenness to prioritise the disruptions while the third used

the concept of network skeletons as captured by link salience. The results from these simulations showed that the link betweenness strategies were far more effective than the link salience strategy.

From these results three aspects of vulnerability were identified. *Redundancy* quantifies the number of alternative shortest paths available to an instance. *Overlap* measures to what degree the shortest path sets of an instance overlap and have road segments in common. *Efficiency step-change* is a measure of the magnitude of the "shock" absorbed by the shortest paths of an instance during a disruption. For each of these aspects one or more metrics were defined. This suite of vulnerability metrics is the second artefact produced by the thesis.

The design of the artefacts itself, although novel, was not considered research. It is the insights derived during analysis of the artefacts' performance that contributes to the body of knowledge. Link-based progressive random disturbance simulations were used to assess the ability of the vulnerability metrics to quantify supply chain vulnerability. It was found that none of the defined vulnerability aspects are good stand-alone predictors of vulnerability. The multilayered nature and random disturbance protocol result in vulnerability being more multi-faceted than initially imagined. Nonetheless, the formulation of the multilayered network proved useful and intuitive and even though the vulnerability metrics fail as predictors they still succeed in capturing shortest path phenomena that would lead to vulnerability under non-random protocols.

To validate the findings from the theoretical instances, link-based random disturbance simulations were executed on 191 case study instances. These instances were extracted from real-life data in three urban areas in South Africa, namely Gauteng Province (GT), City of Cape Town (CoCT) and eThekwini Metropolitan Municipality (ET). The case study instances showed marked deviations from the assumptions underlying the theoretical instances. Despite these differences, the multilayered formulation still enables the quantification of the relationship between supply chain structure and road infrastructure. The performance of the vulnerability metrics in the case study corroborates the findings from the theoretical instances.

Although the suite of vulnerability metrics was unsuccessful in quantifying or predicting vulnerability in both the theoretical and case study instances, the rationale behind their development is sound. Future work that will result in more effective metrics is outlined in this thesis. On the one hand the development of a more realistic disruption strategy is suggested. Road network disruptions are neither completely random nor specifically targeted. Important segments with greater traffic loads are more *likely* to be disrupted, but the reality is that disruptions such as accidents, equipment failure or road maintenance could really occur anywhere on the network. A more realistic disruption strategy would lie somewhere on the continuum between targeted and random disruptions. Other future work suggests the refinement of both artefacts by incorporating link weights in both the logical and physical layers.

An unanticipated finding from this thesis is that future research in the field may be expedited if theory-building emanates from real-life empirical networks as opposed to theoretically generated networks. Expanding the scope of the case study, characterising the true network archetypes found in practice and increasing the number of case study samples is a high priority for future work.

Chapter 1

Introduction

Ask any operations manager on the warehouse floor, in the distribution centre's cross-dock or at the receiving bay of the retail store and they will confirm that disruptions in supply chain activities are far less exceptional and far more costly than commonly believed. It is not only the natural disasters and terrorist attacks that cost supply chains billions in turnover, the less extraordinary realities of traffic congestion, power outages, internet network failure and even industrial action can be just as harmful. The impact of these disruptions could range from missed delivery time windows to stock-outs or unplanned overtime. Although these impacts are of a smaller scale, over time they could lead to *death by a thousand cuts* for any supply chain.

1.1 Turbulence is the new normal

It is difficult to find a supply chain management paper published after 2001 that does not bemoan the perplexing turbulence of the global business community. From natural disasters to oil price volatility, political unrest to market (mis)behaviour — the anecdotes abound. In contrast to the neutral definition adopted by classical decision theory and even the international standard ISO3100:2009 (Purdy, 2010), *risk* in the Supply Chain Management (SCM) arena is regarded as negative (Wagner and Bode, 2006).

> **Supply chain risk** is "anything that [disrupts or impedes] the information, material or product flows from original suppliers to the delivery of the final product to the ultimate end-user" — Peck (2006).

This definition is in character for a management science where *risky* is synonymous with *dangerous* and where the assessment of risk blends qualitative information and gut feel with statistics (March and Shapira, 1987).

Christopher and Holweg (2011) were the first to propose a quantitative metric to measure and trace supply chain turbulence. The first version of the Supply Chain Volatility Index (SCVI) included parameters relating to financial indicators, raw material availability, stock market volatility and maritime shipping costs. This first iteration of the SCVI showed that "[a]s of 2008, we have left an almost 30-year lasting period of stability behind and are now entering a period of turbulence that was last seen during the oil crisis of 1973" (Christopher and Holweg, 2011). The magnitude of the peaks and troughs post-2008 were, however, larger than those in the 70's. Four pertinent trends have burgeoned supply chain volatility: an increase in natural and man-made disasters; a rise in supply

chain complexity; heightened financial pressure on supply chain operations; and fiercer global competition (Wagner and Neshat, 2010). Six years later, Christopher and Holweg (2017) recalculated a much refined version of the SCVI. They showed that volatility had reduced somewhat from the unprecedented levels post global financial crisis but had not returned to its pre-crisis state. The authors warned that current volatility is the new normal. Turbulence is now the rule, not the exception.

1.1.1 Supply chain vulnerability

It is not just the specific disruption that determines the magnitude of its damaging effect, but also the degree to which a supply chain is susceptible to that damage (Wagner and Bode, 2006). A retail chain with multiple distribution centres in a metropolitan area would be less susceptible to a power failure at one of the centres than a similar chain with only one consolidated distribution centre. The former would be able to fill orders from the other centres during the downtime while the latter would be debilitated. This is a typical example of susceptibility arising from the *internal configuration* of the supply chain. Let's now consider a scenario where the internal configuration of two chains are identical, but one operates in a City A where snowstorms are a common occurrence and the other in a City B where it hasn't snowed in twenty years. Should a blizzard hit both cities, the infrastructure and municipal services in City A would cope far better than that of City B (which probably doesn't have more than one working snow plough). In this case the susceptibility is due to *external circumstances*. Wagner and Bode (2006) discussed the works of numerous contributors to the definition of supply chain vulnerability and the definition adopted in this thesis echoes that of the main proponents they mentioned.

> **Supply chain vulnerability** is the degree to which the supply chain's internal configuration and external circumstances make it susceptible to the damaging effects of a disruption.

Although there is a significant body of literature addressing Supply Chain Risk Management (SCRM), reviewers noted that most work fixates on definitions, frameworks and taxonomies. These theoretical constructs are too vague and ambiguous to offer practical solutions to managers who wish to account for vulnerability in their decision making (Heckmann et al., 2015; Hohenstein et al., 2015; Rao and Goldsby, 2009). Nonetheless, some frameworks can be a useful starting point. One such framework is that of Peck (2005). From their study of critical sectors in the UK economy, they classified supply chain vulnerability drivers into four levels:

Level 1: Value stream/product or process. This level adopts a typical engineering-based process flow view of vulnerability. It is concerned with the smooth execution of the sequential processes in the value chain. Typical risks include machine breakdown, product failure or an inability to respond to volatile demand or changing market needs.

Level 2: Asset and infrastructure dependencies. Vulnerability on this level pertains to the dependence of the supply chain on facilities, communication & Information Technology (IT) infrastructure, power grids and transportation networks. These infrastructures most often do not fall within the supply chain's ambit of control.

Level 3: Organisations and inter-organisational networks. This level zooms out even further to view supply chains as an interactive network of individual organisations. Vulnerability on this level relates to corporate strategy, business relations and micro-economics.

Level 4: The environment. Finally the fourth level regards the wider macro-economic, socio-political and natural environments within which the supply chain operates.

Different modelling paradigms are better suited to each level of vulnerability drivers. Thus, this framework is useful in determining the gaps relating to the quantification of supply chain vulnerability drivers.

While the stream of SCRM research grapples with the classification and avoidance of supply chain risk and vulnerability, Supply Chain Resilience (SCRes) formulates strategies and approaches to overcome those disruptions that do occur. The discourse around SCRes began after the terrorist attack of 9/11 with seminal works by Christopher and Peck (2004), Rice and Caniato (2003), and Sheffi and Rice (2005) and has surged since 2011 (Kilubi, 2016b; Tukamuhabwa et al., 2015).

1.1.2 Building resilient supply chains

A resilient supply chain, as Sheffi and Rice (2005) plainly put it, is one that can *bounce back*. Recent reviews summarised how definitions had evolved to include the phases of readiness, response, recovery and growth (Hohenstein et al., 2015; Kamalahmadi and Parast, 2016) and even the concept of cost effectiveness (Tukamuhabwa et al., 2015). We[1] drew from these definitions but returned to the simplicity of the original proposition.

> **Supply chain resilience** is the degree to which a supply chain, after being disrupted, can return to a state of operation where it provides the same or better service to its customers at an equal or greater level of efficiency.

To build resilient supply chains, companies need to know both what implementation strategies are available to them and how to weigh these against other corporate objectives such as cost efficiency and short-term profits.

Resilience strategies

The very first strategies to make supply chains more resilient focussed on *flexibility* and *redundancy* (Rice and Caniato, 2003; Sheffi and Rice, 2005). A multi-skilled workforce or production process capable of quick product changes are examples of flexibility. It creates the capability to respond and change quickly to adapting needs. Redundancy focusses on maintaining current capacity during a disturbance. This is possible only if additional capacity is available in the form of safety stock, overtime or unutilised production capacity. Investment in flexibility and redundancy has to be made before a supply chain faces trouble. While both strategies represent additional costs, the authors argued that redundancy is more costly to implement and maintain.

[1] In this thesis I make many references to myself, noting decisions and approaches that I am ultimately responsible for. Dealing with these references in the third-person reads awkwardly, and does not reflect the blood, sweat, swearing and tears of a real person. Hence, I will make reference to "we". The efforts and decisions attributed to the royal "we" reflect my own inputs, and not that of others in the research group, unless explicitly acknowledged.

As the discourse of SCRes continued, researchers elaborated many more resilience strategies and investigated their relationships. In their review of 100 SCRes papers, Kamalahmadi and Parast (2016) identified four prevailing SCRes principles and mapped the strategic themes to these principles as shown in Figure 1.1.

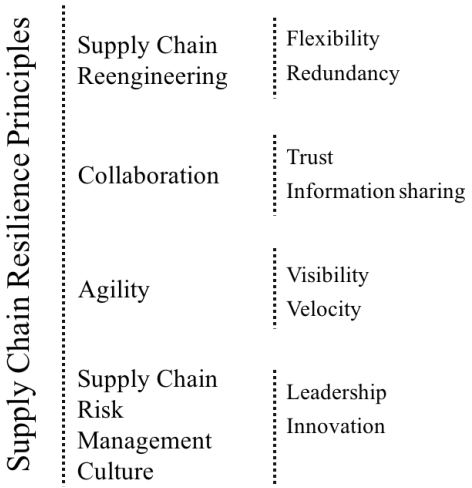


Figure 1.1: SCRes Strategies (adapted from Kamalahmadi and Parast (2016)).

Although it is widely acknowledged that adopting resilience strategies often challenges prevailing management culture, few authors have been as explicit as Christopher and Holweg (2011, 2017). They dismissed altogether the traditional supply chain management style premised on stability. They argued that in this era of turbulence the quest for control and emphasis on streamlined systems actually does more harm than good. The rigidity of control mechanisms often serve to amplify, instead of dampen variability. Popular examples of such philosophies include lean, SIX SIGMA and consolidation. These approaches prioritise cost efficiencies under the assumption of stability, failing which their benefit is short-lived. The long term costs of a traditional SCM approach can be crippling when, not if, disruption strikes.

Christopher and Holweg (2011) differentiated between *dynamic* and *structural* flexibility. Dynamic flexibility is reactive: when facing disruption a supply chain is able to contract in additional resources or temporarily adjust transport options — in short, make a quick plan. But these solutions are temporary departures from the state-of-practice and cannot be sustained indefinitely. Structural flexibility is proactive: the very make-up of the supply chain — its *internal configuration* — is adaptable to changes in the business environment. They cited the case of Zara, a global clothing label that could respond within a week or two to changes in fashion and demand. Their "rapid-fire" supply chain was reported to utilise small modular factories in Northern Spain that made them adaptable, not just reactive. The progression towards resilience is from an efficient supply chain (traditional view) to a stable supply chain (dynamic flexibility) and finally an adaptable supply chain (structural flexibility). But this progression requires a very real mental and cultural shift. Table 1.1 contrasts the mindsets behind efficient and adaptable supply chains.

Table 1.1: Efficient versus adaptable supply chains (Christopher and Holweg, 2011).

	Efficient supply chain	**Adaptable supply chain**
Focus	Establish control to reduce variability and thus cost to compete.	Embrace volatility and develop superior ability to adapt.
Decision time horizon	Short-term, quarterly results.	Long-term viability, while maintaining positive cash flow.
View on turbulence	Bad, as it causes instability and cost.	Inevitable, hence the need to pre-empt it by creating adaptable structures.
Approach to dealing with turbulence	Use SIX SIGMA and other tools to eradicate it where possible.	Use tools to increase flexibility "bandwidth" to cope.

Weighing efficiency against resilience

In the early days of SCRes, Tang (2006) lamented that while managers acknowledged supply chain risks, they failed to make proportionate investments in resilience. Along with other earlier works (Rice and Caniato, 2003; Zsidisin et al., 2000, 2004) the author proposed three reasons for this observation:

Assessment: In the absence of an accurate assessment of prevailing risks and a firm's inherent vulnerability, managers underestimated the problem.

Ingenuity: Should firms have acknowledged the need to act, they were uncertain of what to do.

Evaluation: Existing accounting methods could not properly capture the cost/benefit or return on investment of resilience strategies.

More recent feedback from industry showed that the attitude towards resilience has changed. A World Economic Forum survey reported in 2011 that more than 90% of respondents said that supply chain and transport risk had been elevated as a top priority in the preceding five years (Chacon et al., 2012). The following year the survey's focus shifted to "building resilience". This time more than 80% of companies were concerned about resilience and wanted to act. Most of the experts surveyed believed that resilience and efficiency could coexist and that it didn't have to be an either-or decision (Bhatia et al., 2013). A recent study of the FMCG sector in South Africa observed that managers were willing to invest significantly in resilience strategies even in spite of the potential short-term losses this may cause (Agigi et al., 2016). These sentiments contrast the reluctance voiced by industry a few years earlier.

So what are the current impediments to creating resilient supply chains? The wealth of resilience strategies emanating from case studies and surveys indicate that ingenuity is no longer the issue. Companies know the options available to them. Assessment and evaluation remain the primary hindrances.

1.1.3 Quantitative tools for a new management mindset

Although a number of quantitative models have been proposed in the fields of SCRM (Heckmann et al., 2015) and SCRes (Kamalahmadi and Parast, 2016; Kilubi, 2016a), not all levels of vulnerability drivers have been equally addressed.

Vulnerability drivers on the first level have traditionally received the most attention. It is an established field with frequent applications of Monte Carlo type models, stochastic programming (Klibi and Martel, 2012) and simulation models (Wu et al., 2013). Any mathematical model that can test outcomes over a range of probabilistic scenarios can quantify risk and vulnerability on this level. The quality of the assessment depends on the appropriate choice of model and the reliability of risk information.

The other level that has received increasingly more attention as new modelling techniques developed is the third level. Complex Network Theory (CNT) and agent-based modelling have been applied with much success (Bellamy and Basole, 2013). In CNT organisations are modelled as nodes and their relationships as links. The topology, community structure and hierarchy of nodes and links then offer good insights of a business landscape that otherwise can seem undecipherable. Agent-based simulation is then capable of modelling the behaviour of autonomous agents that interact within these networks. Section 1.2.2 elaborates on such studies.

While the first and third levels have received considerable scrutiny, the second and fourth levels pose verifiable gaps in assessment. In both cases the difficulty is in understanding the relationship between the supply chain and the vulnerability driver. What impact would political unrest in Europe really have on a vehicle manufacturer in the USA? To what degree would damage to SEACOM's fibre-optic submarine cable on the East coast of Africa impede communication in a retail supply chain in South Africa? And how would this affect the coordination of deliveries? What is required is models that can capture the relationship between these phenomena external to the supply chain and the supply chain's operations.

1.1.4 Problem statement

Supply chain networks in urban areas often face volatile demand with short notice periods, tight transport lead times, fierce competition and fickle customers. Urban centers also frequently experience traffic congestion and disruptions due to roadworks or accidents, infrastructure or equipment failure. The reality of the supply chain's daily dependence on urban transport infrastructure makes it vulnerable. But how vulnerable exactly?

Executives concerned with this dependence on the urban road network may ask: *"How vulnerable is our current distribution network design in City A? By what margin could we lessen this vulnerability by changing our facility locations? Alternatively, by what margin could we reduce vulnerability if we moved operations to neighbouring City B? What impact would these changes have on our bottom line in the short, medium and long term?"* To answer these questions requires a model that can quantify the dependence.

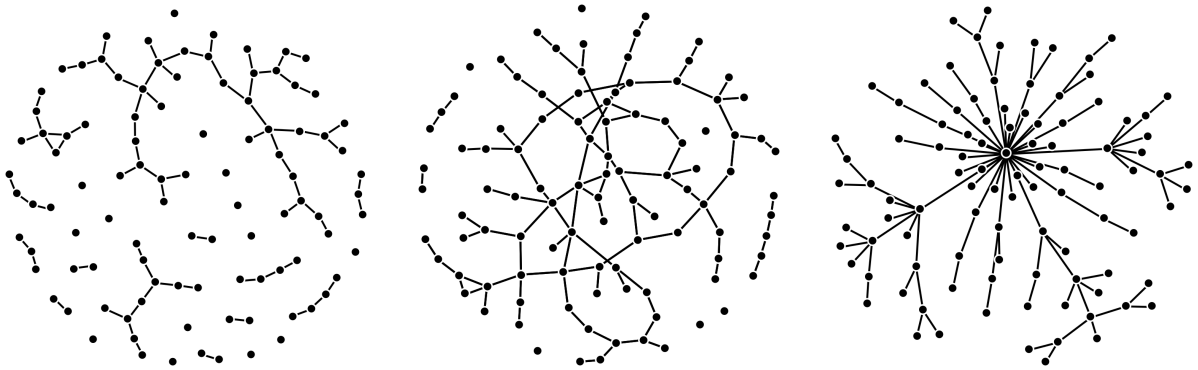
1.2 A complex network theory perspective

Graph theory, which is the study of networks, originated with Euler's famous solution of the Seven Bridges of Königsberg in 1736. Since then it has been pivotal in disciplines such as mathematics, quantitative geography and operations research. In the late 1950s

Erdös and Rényi (1959) proposed the famous random network model, since called the Erdös-Rényi (ER) network. In the ER network each node is connected to a random set of remaining nodes. For decades, before the advent of Global Positioning System (GPS) data and other technologies, the capability to study large real-life networks without extensive simplification to either a completely regular or completely random network representation did not exist.

In 1998, Watts and Strogatz (1998) tried to find a middle ground by rewiring regular networks to induce randomness. What they observed was a topology that has dense clustering, like that of a regular network, but also relatively short path lengths, similar to ER networks. They called this the *small-world* network, because every node is only a few links away from every other node, regardless of network size. The small-world network is also referred to as the Watts-Strogatz (WS) network.

One year later, Barabási and Albert (1999) observed that for many diverse real-world systems the connectivity of the nodes follow a scale-free, power-law distribution. This meant that the majority of the network nodes only have a few connections while a few nodes are highly connected. These highly connected nodes are also called *hubs*. The authors developed an algorithm to grow networks that replicate this real-world topology. The algorithm adds nodes to an existing network using preferential attachment to nodes that are already well connected. This finally gave shape to the intuition that large, ungoverned networks are self-organising. This topology became known as the *scale-free* or Barabási-Albert (BA) network. (Figure 1.2 show examples of each of the three principal topologies.)



(a) Erdös-Rényi/ Random (b) Watts-Strogatz/ Small-world (c) Barabási-Albert/ Scale-free

Figure 1.2: The three foundational CNT topologies (reproduced from Basole and Bellamy (2014) with permission).

Within seven years comprehensive reviews examined the structure and dynamics of a host of theoretical and real-world networks (Albert and Barabási, 2002; Boccaletti et al., 2006; Dorogovtsev and Mendes, 2002; Newman, 2003). Topics covered in these reviews included: *structural characteristics* such as centrality measures, clustering coefficients and shortest path metrics; *growth models* relating to the ER, WS and BA networks; *spreading processes* such as percolation theory and disease spread; and *error and attack tolerance*. These works set the stage for the surge of complex network theory applications to systems in sociology, neurology, biology, ecology, transportation, computer science and critical infrastructures.

A next wave of research probed community structure in networks. In his review Fortunato (2010) discussed a library of algorithms and techniques for identifying and extracting

communities. Some authors have referred to communities as meso-level characteristics.

Another meso-level characteristic recently identified is that of network skeletons. To reduce the complexity of large networks scientists have developed methods that reduce the "noise" in the network while preserving key characteristics. Network skeletons contain only the nodes and links that remain once "unimportant" nodes and links have been filtered out according to some rationale (Grady et al., 2012; Serrano et al., 2009; Shekhtman et al., 2014).

In 2011, Barthélemy (2011) focussed a review on the subset of spatial networks. These networks are embedded in space and the links represent a physical distance as opposed to a purely relational distance. Empirical networks discussed in the review included transportation networks, infrastructure networks, origin-destination matrices, mobility networks and neural networks. He asserted that the geographic distance affects many metrics and processes in a significant way and thus warrant a separate discussion.

CNT revolutionised the study of real-world systems. It gave tools and paradigms with which to make sense of dense complexity.

1.2.1 Multilayered complex network theory

Infrastructure systems are tangible, measurable, definitive and (often) immovable. Supply chains, on the other hand, are a conglomeration of interconnected business relationships that cannot be easily defined or measured and are subject to change. These are two separate networks — one *physical* and one *logical* — a perfect scenario for *multilayered* CNT.

Multilayered CNT captures interdependent intricacies between individual networks. Consider the social network example of relationships between colleagues all working together in a department. Some may be no more than work acquaintances, others may work together on projects and therefore have a closer relationship, and then there are those who have become friends outside of work and compete together in the volleyball league. This scenario describes three distinct relationship networks. Consider now the same group of colleagues but focus only on the project team relationships. As time goes by project teams change, giving rise to a unique set of relationships. Once again there are two distinct networks, separated by time. All of these distinct networks can be combined and studied simultaneously using techniques developed by multilayered complex network theory. The network science community regard the study of multilayered networks as a new frontier in many areas of science and a rapidly expanding movement that will stimulate interdisciplinary research (Boccaletti et al., 2014). Similarly, it offers promise in getting a handle on supply chain vulnerability, especially in terms of its underlying infrastructure.

1.2.2 Supply chain applications

Supply chains are interconnected networks of multiple entities (agents) that exhibit adaptive action in response to change in both the environment and system of entities itself (Choi et al., 2001). It has *structural complexity* (the interconnectedness of firms) and *adaptivity* (dynamic learning) (Pathak et al., 2007). These are typical characteristics of Complex Adaptive Systems (CAS). There is a growing body of knowledge proposing more appropriate models to emulate CAS (Bellamy and Basole, 2013) and CNT is fast becoming a favourite.

CNT is a well-suited approach to making sense of the *structural complexity* of supply chains. Hearnshaw and Wilson (2013) set a standard in applying CNT to supply chains by comprehensively mapping different facets of a supply chain to complex network constructs. They investigated which topological characteristics make for an efficient supply chain. Another study by Kim et al. (2011) modelled material flow and relationships in the automotive industry using a complex network framework. The analysis showed that this approach provided additional quantitative insights that the initial case study by Choi and Hong (2002) missed. These and similar studies provided good descriptions of the structure of a supply chain, but to study the topic of vulnerability one needs to understand the *adaptivity* of a supply chain.

Recent studies have used two CNT approaches to studying supply chain vulnerability. The first approach uses simulations of random errors and targeted attacks. Thadakamalla et al. (2004) defined a resilient supply chain as one that

> "maintains connectivity between the majority of its nodes; does not suffer a significant increase in the average shortest path length between nodes; has well-defined clusters that offer many alternative shortest paths and can automatically re-wire itself after disruption to establish functionality".

According to these criteria they evaluated the resilience of the random, small-world and scale-free topologies. Their results showed that a hybrid topology is more resilient. Nair and Vidal (2011) extended this purely topological perspective by adding inventory flow to the random and scale-free formulations. They found that there is a significant relationship between topological vulnerability and vulnerability from an inventory level, backorder disruption and total cost point of view. In a study customised for military supply chains Zhao et al. (2011) proposed alternate vulnerability metrics that took into account the fact that different facilities performed distinct functions. Their customisation provided a more accurate analysis of the real-life vulnerability compared to approaches that did not consider facility function.

The second approach to studying vulnerability uses concepts from epidemiology to study the spread of risk or damage through a network. Basole and Bellamy (2014) studied risk diffusion in networks where nodes represented individual firms and links represented business relationships. Their results showed that small-world networks are more robust than scale-free networks when it comes to risk diffusion. A following study evaluated financial risk diffusion amongst supply networks from the electronics industry spread across North America, Asia and Europe (Basole et al., 2016). Their results showed that networks that were not as dependent on a few central hubs (scale-free) but instead had relationships that connected distant neighbours (small-world) reduced the impact of risk propagation and increased overall network health.

Considering the four levels of vulnerability drivers defined by Peck (2005) we observed that the work of Basole and Bellamy (2014); Basole et al. (2016) and Thadakamalla et al. (2004) addressed third level drivers relating to organisations and inter-organisational networks. Nair and Vidal (2011) and Zhao et al. (2011) addressed a combination of third level and first level drivers by respectively adding the inventory and supply-demand perspectives. None of these studies addressed the multilayered dependency between supply chains and infrastructure.

1.2.3 Road network applications

Road transport systems have a more mature repertoire of CNT applications than supply chain management. Standard network representations and methodologies have emerged, making results comparable. Applications have either focussed on understanding and comparing the topological characteristics of networks or quantifying vulnerability through the simulation of targeted attacks and random errors. Multilayered applications in urban transportation systems have also become popular.

Many studies focussed exclusively on the topological structure of the road network, disregarding transport activity. Insights from such applications complemented the understanding of urban structures, densities and dynamics (Barthélemy and Flammini, 2008; Crucitti et al., 2006; Jiang and Claramunt, 2004; Jiang, 2007; Masucci et al., 2009; Porta et al., 2012; Strano et al., 2009). However, there are very few studies that have incorporated transport activity in a road-only network. Most studies that incorporated transport activity focussed on public transit systems that combined multiple modes.

One very relevant exception that modelled commercial vehicle activity on a road-only network is that of Joubert and Axhausen (2013). They developed an innovative methodology to study the complex network characteristics of commercial vehicle movement in Gauteng, South Africa. The activity chains of commercial vehicles were extracted from GPS data (Joubert and Axhausen, 2011) and used to pinpoint logistics facilities (nodes) using clustering algorithms (Joubert and Axhausen, 2013). The approach was refined in Joubert and Meintjes (2015a). These authors built a complex network of commercial vehicle movement and used centrality metrics to identify key logistics players.

A significant restriction of single layer CNT is that it is limited in terms of capturing the multimodal nature or multiple service layers characteristic of real-life transportation systems. We considered a supply chain network layered on an urban road network. The supply chain layer represented a *logical* layer defined by the supply chain network design. The links represented inter-firm or intra-firm relations. Freight could not travel along such ethereal links but instead travels along the urban road network which represents the *physical* layer of the network.

Notably two other multilayer transportation studies also placed a *logical* layer on a *physical* layer (Kurant and Thiran, 2006b; Zhuo et al., 2011). However, these studies addressed passenger transport, not freight.

Van Heerden and Joubert (2014) described how the understanding and modelling of freight traffic lags far behind that of passenger traffic. Although freight traffic may constitute a small percentage of traffic within an urban environment, it has a disproportionate effect on economic activity. Traditional freight modelling focuses on commodity flow, not vehicular movement. Most traffic models add commercial vehicle activity as background noise or by multiplying passenger traffic by some factor. However, recent studies of disaggregate commercial vehicle movement have shown explicitly that commercial vehicles simply do not behave like passenger vehicles (Joubert and Axhausen, 2011, 2013; Van Heerden and Joubert, 2014). Although these studies have made inroads in addressing the freight transport research gap, they still focus on a single-layer view of freight connections, not explicitly incorporating the physical infrastructure. So while there have been numerous CNT and multilayer CNT studies relating to the road network, none directly address the problem statement in this thesis.

1.3 Road vulnerability studies

Robust and reliable road transport systems are essential to economic activity and welfare. Consequently, the vulnerability of road networks has been a topic of ardent research since the 1960s (Mattsson and Jenelius, 2015). In their review, the authors identified two traditions of road network vulnerability research. The one tradition, termed *topological vulnerability studies*, regards only the topological properties of road transport networks. The second tradition, termed *system-based vulnerability studies*, takes into account the demand and supply profiles of the transport system and how these are affected. Although related, these two traditions have limited interaction, distinct authors and are even published in different types of journals.

Advocates of system-based vulnerability studies contend that it is a far more realistic representation of the real-life system's response to disruptions. Because the metrics are also more intuitive, expressed in terms of costs or time or percentage served, it is easier for planners to see the path to practical application. Another benefit (data permitting) is that studies can focus on discrete population segments such as the elderly or commuters living in a specific feeder town outside a major city. Unfortunately, the drawbacks of this approach can be quite significant. System-based studies are data intensive, computationally burdensome and have a lower level of standardisation and comparability (Mattsson and Jenelius, 2015).

In this thesis we were limited by the availability of supply chain data. For the data regarding commercial vehicle movement that were available to us, not enough was known about the travel demand, behaviour and user value of commercial vehicles in those urban settings to posit generalisable assumptions.

In topological vulnerability studies real-life transport systems are represented as abstract networks. CNT is the prevailing approach used to model these networks. Vulnerability is then investigated using targeted attack and random error simulations.

Although the bulk of these studies focussed on subway systems and other public transit systems, there are examples that have dealt exclusively with road network vulnerability. Demšar et al. (2008) studied the road network of the Helsinki Metropolitan Area in Finland while Duan and Lu (2014) compared vulnerability across three levels of geographic granularity using six real city road networks from Asia, Europe and North America.

The relative simplicity of the topological approach and its limited data requirements make it easy to execute large experiments for multiple vulnerability scenarios. This is desirable for studies that are exploring new network metrics as the validity of the metrics can be compared across a range of topologies with relative ease.

1.4 Topological vulnerability studies using CNT

Topological vulnerability studies are concerned with identifying the critical nodes and/or links in a network that, if removed, will cause greatest damage to the network.

There are two pertinent reasons that make prioritising critical elements non-trivial. Firstly, the very structural complexity that defines these networks makes it impossible to find analytic answers. Relationships are emergent, not predefined. Therefore, simulation studies have become state-of-practice in testing different prioritisation schemes. Secondly, the response variables defined to measure network damage in these simulation studies differ. If there are different viewpoints on how to measure damage, then there will be different ways of prioritising critical elements.

Three common dimensions of network damage have evolved. *Robustness* refers to the connectedness of a network, whether it has broken into disconnected sub-networks or still functions as a cohesive unit. *Efficiency* (referred to by some as *responsiveness*) determines how quickly a message (or people or freight) can travel from one node to another. *Flexibility* refers to the network's inherent ability to find alternative paths if the shortest paths are destroyed.

In multilayered networks, the concept of cascading failures and damage in terms of network robustness have been the foremost phenomena in vulnerability studies. If the layers of a multilayered network are interdependent then the links between the layers are called dependency links. Removing any node in one layer would then trigger the removal of all nodes in all other layers that are connected via dependency links. Such failures cascade back and forth until all dependent nodes are removed.

Simulations either emulate *random errors* or *targeted attacks*. In random error simulations the nodes or links to be removed are selected randomly. Targeted attack simulations use some predefined strategy to prioritise nodes or links for removal (Albert et al., 2000).

1.5 Research design

The following thesis statement was formulated:

> *Metrics related to the shortest path sets of the multilayered supply chain/road network formulation can quantify the inherent vulnerability of a specific supply chain to its choice of internal configuration and the underlying urban road network's integrity.*

In order to evaluate this statement the following objectives had to be achieved:

1. Development of a multilayered complex network model that captures the dependence of a supply chain network on an urban road network.
2. Identification of the characteristics of this model that describe the nature of the supply chain's vulnerability to the integrity of the urban road network.
3. Development of metrics that could quantify a supply chain's inherent vulnerability based on its internal configuration and the underlying road network.
4. Evaluation of the validity of the suite of vulnerability metrics through statistical analysis and a real-life case study.

These objectives were achieved using a *design research* approach as described by Manson (2006).

1.6 Research methodology

Design research is

> "a process of using knowledge to design and create useful artefacts, and then using various rigorous methods to analyse why, or why not, a particular artefact is effective" (Manson, 2006).

Design itself is not considered research but it is through the insights derived during analysis of the designed artefact's performance that the body of knowledge in a field grows. The phases of the design research approach are shown in Figure 1.3. In this thesis the *awareness of* CNT, particularly its application to system vulnerability, led to the *suggestion* that it could be applied to supply chain vulnerability to address the practical knowledge gap outlined in the problem statement. This led to the *development* of two artefacts: a multilayered complex network formulation of the problem, and a suite of vulnerability metrics that were proposed to quantify the inherent vulnerability of the supply chain.

The artefacts were *evaluated* in two ways. Using the multilayer formulation we generated large samples for three different supply chain network archetypes. The distributions of the topological characteristics were investigated and compared to verify the formulation. We then used random error simulations and statistical correlation tests to assess the performance of the suite of vulnerability metrics.

After the evaluation we tested the thesis statement and formalised what we had learnt (circumspection). Further feedback of the artefacts' utility was obtained when we applied these to a case study of three South African urban areas. Feedback from the case study added to the operation and goal knowledge obtained through the study.

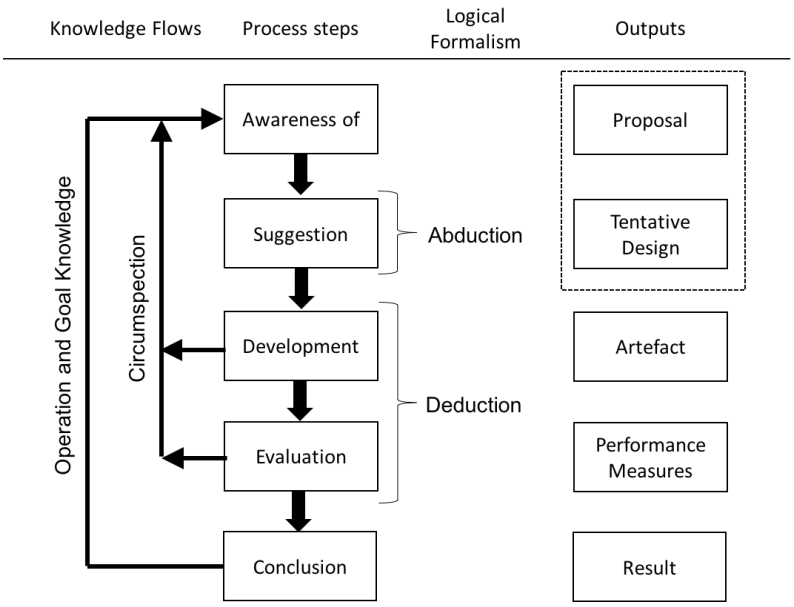


Figure 1.3: Design research methodology (reproduced from Manson (2006)).

1.6.1 Multilayered network formulation and theoretical datasets

The first objective of this thesis was:

> 1. Development of a multilayered complex network model that captures the dependence of a supply chain network on an urban road network.

The multilayered network consisted of a *logical* layer and a *physical* layer. The *logical* layer represented the inter-firm business rules and intra-firm relationships that defined how freight could be transported between logistics facilities in an urban environment. The *physical* layer represented a typical urban road network that commercial vehicles could use to ship freight.

The *physical* network was represented as a regular grid network and the *logical* network was one of three predefined supply chain network archetypes. There was thus little mystery regarding these networks individually. It was how the *logical* network was layered on the *physical* network that created distinct multilayered network instances. Depending on the size of the individual networks, the possible layering permutations could have ranged from thousands to millions.

A representative sample of 500 multilayered network instances was generated for each of the three network archetypes. One instance was generated by associating each node in the *logical* layer to a randomly selected node in the *physical* layer.

What distinguished one multilayered instance from the next was its unique collection of shortest paths. The shortest *relational* path between any two nodes in the *logical* network was predefined. However, freight cannot move along relational paths, it requires the road infrastructure. Thus, when calculating the shortest *freight movement* path between two nodes in the *logical* network, one had to take into account the position of these nodes on the *physical* road grid. The collection of shortest paths that resulted from each multilayered instance was its distinct fingerprint. The characteristics that would describe the supply chain's vulnerability therefore emanated from these shortest path sets.

1.6.2 Link-based targeted attack simulation

Armed with large samples of randomly generated multilayered instances for each of three supply chain network archetypes, the second and third objectives of the thesis could be tackled.

> 2. Identification of the characteristics of this model that describe the nature of the supply chain's vulnerability to the integrity of the urban road network.

> 3. Development of metrics that could quantify a supply chain's inherent vulnerability based on its internal configuration and the underlying road network.

While it is possible that in reality a disruption could damage or destroy an entire intersection — like a catastrophic multi-vehicle accident, it is unlikely. It is more likely that disruptions will disable individual road segments. The same rationale has applied in other transport-related vulnerability studies (Viljoen and Joubert, 2016). Therefore, removing links was more appropriate than removing nodes.

To identify the characteristics of the collection of shortest paths that are indicative of network vulnerability we used three link-based targeted attack simulations. Each simulation used a different characteristic of the shortest path sets to prioritise links for removal. The resultant damage was tracked and compared across the three strategies to show which

characteristics were more telling than others. Based on the performance, a suite of vulnerability metrics was formulated. The proposition was that this suite of metrics could quantify the inherent vulnerability of a supply chain to its internal configuration and the integrity of the underlying road network. To test this proposition, the suite of metrics was evaluated using a random error simulation and statistical tests.

1.6.3 Link-based random error simulation and statistical tests

The urban road network is frequently disturbed by congestion, roadworks, accidents, and infrastructure failure. This could be described as a combination of *random errors* and *targeted attacks*. Streets that are more central could be more likely to experience congestion or higher traffic loads could increase the likelihood of accidents (targeted attacks). However, accidents, equipment failure and roadworks could really happen anywhere on the grid at any time (random error).

The suite of vulnerability metrics were deduced from the outcomes of targeted strategies, therefore testing their validity using further targeted strategies would not have yielded much insight. If a metric is a good quantifier and predictor of vulnerability under completely random link disruption, it is expected that its power will be even greater in circumstances where disruptions have both random and targeted elements[2]. Thus, validating the metrics using random link-based disruptions was the most conservative method of evaluation. A random error simulation was used to track the performance of the suite of vulnerability metrics. Statistical tests based on these results addressed the first part of the fourth research objective:

> 4. Evaluation of the validity of the suite of vulnerability metrics through **statistical analysis** and a real-life case study.

Pairwise correlation tests were used to determine the strength, direction and significance of correlation of the vulnerability metrics to efficiency loss and robustness (number of disruptions endured before breaking). Those vulnerability metrics that were found to have strong, significant correlations to efficiency loss and robustness were then further tested to determine whether they could be good discriminators of disconnection or efficiency loss in the next disruption. These two tests provided a comprehensive evaluation of whether the suite of vulnerability metrics was able to quantify the inherent vulnerability of the multilayered network instances.

Up until this stage of the thesis, the multilayered network instances were *theoretical*. To establish the utility and performance of the artefacts in real-life scenarios, they had to be applied to *case study* network instances.

1.6.4 Case study validation

The case study addressed the last part of the final research objective:

> 4. Evaluation of the validity of the suite of vulnerability metrics through statistical analysis and a **real-life case study**.

[2]Note that regardless of whether a link was selected by a targeted or random strategy, the impact of the disruption was equal — i.e. the removal of the link.

A total of 191 case study network instances were extracted from real-life data in three urban areas in South Africa namely the City of Cape Town (CoCT) in the Western Cape, eThekwini Metropolitan Municipality (ET) in the KwaZulu-Natal province and the Gauteng Province (GT).

The complex network of commercial vehicle movement was used as a proxy for supply chain relationships. In these networks the nodes were logistics facilities and the links represented commercial vehicles travelling between the facilities. The rationale was that if freight is frequently shipped between two facilities, there is a supply chain relationship between these facilities.

The methodologies used to transform commercial vehicle GPS logs into complex networks are documented in Joubert and Axhausen (2011, 2013) and Joubert and Meintjes (2015a,b). The dataset used in the case study contained the complex network of freight movement for each of the three urban areas as extracted from the February 2014 GPS logs. From these three urban networks, the *logical* layers of 191 case study instances were extracted.

Each of these *logical* layers had an underlying road network. The *physical* layers that represented these road networks were extracted from OpenStreetMap (OSM) data. For each of these 191 instances, the *logical* layer was placed on the *physical* layer to create a multilayered network instance.

The link-based random error simulation described in Section 1.6.3 was executed on the real-life multilayered instances. Their response to the simulation and the performance of the vulnerability metrics were compared to the results of the theoretical instances.

Based on the simulation results of both the theoretical and case study instances, we could complete the evaluation of the artefacts. The *design research* loop was closed by commenting on the utility of the artefacts and suggesting future improvements.

1.7 Thesis overview

The thesis is structured as follows: Chapter 2 grounds the proposed research design by reviewing the body of knowledge that promotes CAS techniques for modelling supply chain systems, surveying other relevant applications of CNT to road networks and investigating the most prevalent parameters used in designing topological vulnerability studies. Chapter 3 presents the mathematical formulation of the multilayered network, describes how the 500 theoretical instances were generated for each of the three supply chain network archetypes. The chapter concludes with an analysis of the characteristics of the undisturbed instances. Chapter 4 presents the targeted attack simulations executed on the theoretical instances and identifies key vulnerability characteristics in these networks. Based on these discoveries, Chapter 5 details the development of the suite of vulnerability characteristics. Chapter 6 describes the random error simulation, details its results and presents the statistical evaluation of the vulnerability metrics. Chapters 7 and 8 showcase the case study. The first chapter details the data preparation and case study instances extracted from the real-life data. It also highlights the (influential) differences between the case study instances and the theoretical instances. The results of the random error simulation performed on the case study instances is given in Chapter 8. This chapter compares the case study and theoretical results in detail before concluding with a discussion regarding the validity of the vulnerability metrics in real-life scenarios. Finally, Chapter 9 concludes the thesis by summarising its approach and findings, stating its research contribution and limitations and providing an outlook of future research.

Chapter 2

Literature review

This review provides a deeper discussion of how the supply chain has been modelled as a Complex Adaptive Systems (CAS). It particularly elucidates why Complex Network Theory (CNT) is an up-and-coming technique for modelling supply chain systems. Thereafter CNT applications to road transport systems are described in more detail, showing the scope of research questions tackled in the domain using CNT as well as the most prevalent techniques and metrics used. Finally, common design parameters for topological vulnerability studies and their applicability to this thesis are discussed.

2.1 Modelling the complex adaptive nature of supply chains

Initially it was convenient to imagine the dynamics of supply chains exactly as the name implies: a *chain* of supply activities. A linear sequence of processes and partnerships. Until recently this was also how supply chains were modelled (Hearnshaw and Wilson, 2013). But supply chains are not really chains. They are networks: complex networks with autonomous actors. Cause-and-effect is non-linear. Hierarchies and structure are ill-defined. And just when you think you have it mapped, everything changes. Supply chains are CAS in every sense (Bellamy and Basole, 2013; Choi et al., 2001; Hearnshaw and Wilson, 2013; Kim et al., 2011; Pathak et al., 2007; Tukamuhabwa et al., 2015).

The term CAS has its roots in complexity theory, a field concerned with the "emergence of order in dynamic and non-linear systems that operate at the edge of chaos" (Tukamuhabwa et al., 2015). CAS are both *structurally complex* and *adaptive* (Pathak et al., 2007). Anderson (1999) defined four essential characteristics of CAS:

Schema: Aggregate trends are the result of autonomous entities that interact with other autonomous entities and their environment by following simple decision-making rules.

Self-organisation: Patterns and regularity emerge from the recursively applied rules of the autonomous agents. There is no central controlling mechanism or parameters that govern the system.

Co-evolution: Autonomous agents will continue adapting to improve their payoff. When the adaptation of many agents pushes a system to chaos, the agents will self-regulate to re-establish equilibrium. This process is recursive.

System evolution: Cycles of co-evolution among agents cause the system as a whole to evolve. Although this evolution is not predictable using parametric models, prototypical behaviours emerging recursively over time can be observed.

In their seminal work, Choi et al. (2001) mapped these four characteristics to supply chains. A number of case studies followed that proved that the CAS approach to solving supply chain problems reaped at least four benefits over previous approaches namely: increased efficiency, better preparedness for external uncertainty, increased awareness of markets and competition; and overall improved decision making (Pathak et al., 2007).

Bellamy and Basole (2013) conducted a comprehensive review of of 126 papers published between 1995 and 2011 that considered the concept of network analysis in the supply chain context. Although they did not use the term CAS, it is clear that their perspective of network analysis was synonymous. They identified three themes of CAS applications to supply chain systems:

Network structure (system architecture). This considers the node-level, link-level and network-level properties of a system. These properties are typically quantified by traditional CNT metrics such as centrality, embeddedness, clustering and tie strength.

Network dynamics (system behaviour). These studies analyse the initial formation of a network, its evolution over time as well as the emergence or propagation of phenomena. At this level, concepts from CNT, epidemiology, systems thinking and Supply Chain Risk Management (SCRM) are combined to study system behaviour. Time-based simulations are used to monitor the response of the system architecture metrics to network growth or to specific stimuli or disruptions.

Network strategy (system control and policies). This stream of research studies how the implementation of strategies and protocols regarding the network structure and behaviour impact performance. In this field researchers have drawn heavily from the fields of systems engineering, SCM and SCRM to study scope, intent and governance of the supply chain system.

CNT has been used to address problems relating to both *network structure* and *network dynamics*. Entities are modelled as *nodes* within the network. The relationships that connect them — be these business contracts, infrastructure or even the flow of goods — are the *links*. CNT offers an interdisciplinary lens that aggregates these elementary relationships in such a way that one can appreciate the structures and behaviours that emerge from the system. On an elementary level micro-level metrics such as centrality and betweenness illuminate hierarchical structures. Community structures and network skeletons come to the fore using meso-level techniques. On the macroscopic level one can differentiate between different network topologies based on the distributions of metrics. CNT holds great promise in studying supply chains which are known for their scale and complexity and nonlinear behaviour (Basole and Bellamy, 2014; Pathak et al., 2007).

2.2 Relevant applications of CNT to road networks

There are different ways of representing a transport system as a complex network. Although defining nodes can be straightforward, how one defines the links of the network governs the interpretation of the results (Ducruet et al., 2010; Hu and Zhu, 2009; Kurant and Thiran, 2006a).

In their analysis of mass public transit systems, Kurant and Thiran (2006a) defined three types of network representations that are relevant to all constrained-link networks. Quite simply, the stations are the nodes, but the differentiation lies in how the links are defined. In the *space-of-stations* representation two nodes are connected if there is a roadway or rail line linking them directly (i.e. there is no intermediate station). *Space-of-stops* considers two nodes linked if they are *consecutive* stops on a route. Once again it is a direct link with no intermediate stops. *Space-of-changes* considers two nodes linked if the same vehicle visits both on a route, thus there may or may not be intermediate stops.

In constrained-link transport modes there are many studies that only regarded the network of physical infrastructure and disregarded transport activity. Such studies used a space-of-stations representation. Applications of this kind are particularly popular in road transport where insights from network modelling complement the understanding of urban structures, densities and dynamics. The works of Barthélemy and Flammini (2008); Crucitti et al. (2006); Jiang and Claramunt (2004) and Jiang (2007) compared urban structures and growths of world cities using the urban street network. Masucci et al. (2009) focussed more specifically on the growth of London's urban street network and the resultant ease of navigation. Porta et al. (2012) and Strano et al. (2009) approached urban street networks from another perspective, investigating the correlation of street centrality to economic activities in Barcelona, Spain and Bologna, Italy, respectively.

Other studies used either a space-of-stops of space-of-changes representation to formulate single layer complex networks of transport activity. Kurant and Thiran (2006a) mapped the multimodal public transport system of Warsaw, Poland; the railway network of Switzerland; and the regional rail system of central Europe. They used timetable data as a proxy for transport activity and modelled all three networks using different representations. What they found was that the different representations produce fundamentally different networks. Sen et al. (2003) deduced transport activity profiles from India's passenger rail schedules. Using a space-of-changes representation they showed that the system has strong small-world properties. Joubert and Axhausen (2013) used an innovative methodology to study the complex network characteristics of commercial vehicle movement in Gauteng, South Africa. They extracted the activity chains of commercial vehicles from Global Positioning System (GPS) data and used this to pinpoint logistics facilities (nodes) (Joubert and Axhausen, 2011). Using a space-of-stops representation they built a complex network of commercial vehicle movement. Centrality metrics helped to identify key logistics players, offering unique insights to transport planners.

A significant restriction of single layer applications in transportation is that it is limited in terms of capturing the multimodal nature or multiple service layers characteristic of real-life systems.

Multilayered representations

Multilayered network studies in transportation have been more recent. Most studies to date have either used the layered approach to model interdependent transport modes (Gallotti and Barthelemy, 2014, 2015; Parshani et al., 2011; Solé-Ribalta et al., 2016;

Strano et al., 2015) or used different passenger services or cargo flows that utilised the same mode (Cardillo et al., 2013; Ducruet, 2013; Kaluza et al., 2010; Lordan et al., 2015; Tsiotas and Polyzos, 2015).

A context familiar to many is the multimodal interaction between the street network and subway systems of New York and London. Strano et al. (2015) modelled the street and subway networks as interdependent layers and found that these two cities show similar emergent transport topologies despite their different geographies. Network metrics suggested that the subway layer acts as a decentralising force that moves congestion from the centers of the cities to the terminal stations of the subway lines. This multilayered view also highlighted that speeding up subway lines is not always better but can result in an uneven spatial distribution of accessibility. Solé-Ribalta et al. (2016) explored the similar phenomenon of congestion onset. They proved analytically that the very fact that transport layers are interdependent induces congestion and so they developed equations that could approximate this onset based on characteristics of individual layers.

Gallotti and Barthelemy (2014) broadened the scope of multimodal urban mobility when they considered all modes within the British public transport system. They asserted that the growth of multimodal transportation systems has the unintended consequence of an increase in time lost through connections. They compared theoretical shortest trip statistics to 'time-respecting' paths to quantify time spent riding, waiting and walking. The value of their insights to transport network planners is clear. Following up on this work, Gallotti and Barthelemy (2015) documented a comprehensive methodology used to construct a temporal multilayered network dataset of all passenger modes in the United Kingdom for a week in October 2010.

Adding socioeconomic considerations to urban mobility, Lotero et al. (2016) constructed six multilayer representations of transit systems in Bogotá and Medellín for different socioeconomic strata. The strata corresponded to household income. The network for each stratum contained layers corresponding to different transport modes. This novel representation allowed unique insights into the mobility patterns of different socioeconomic classes. The poorest used few and cheap modes to cover large parts of the urban area in a sparse way. The middle income classes showed truly multimodal behaviour and covered nearly all the urban zones. Finally the wealthiest commuters used the most expensive modes and travelled only in certain urban areas. Such insights are invaluable to city planning and policy making.

Of all the multilayered road network studies surveyed, only two other studies placed a *logical* layer on a *physical* layer. Kurant and Thiran (2006b) investigated the centrality metrics of public transport systems in three European geographies by defining one layer as the transport infrastructure and the other as transport intensity extracted from timetables. They asserted that the layered view offers a better estimation of the real traffic load on the network than other commonly used techniques. Zhuo et al. (2011) studied congestion vulnerability on a selection of experimental random and scale-free networks. They also defined one layer as physical infrastructure while the other comprised traffic profiles generated by an algorithm. Their results showed that a homogenous network structure is more tolerant of congestion. While these two studies moved closer to the topic of this thesis, the key difference remains that they addressed passenger transport, not freight.

2.3 Common design parameters for topological vulnerability studies

A wealth of topological vulnerability studies are accumulating in the transport domain. In this section we discuss the most prevalent ways in which network damage has been assessed in single-layer and multilayered complex networks. Then the typical decision process followed by researchers when designing targeted attacks is discussed.

2.3.1 Metrics to assess network damage

In their seminal work on *targeted attacks* and *random errors*, Albert et al. (2000) proposed that the diameter of a network is an indication of connectedness (is it possible to travel between two nodes at all) and efficiency (how easy it is to travel between two nodes). Therefore, they measured the change in the diameter of the network as an indication of damage.

Since this seminal work, three common dimensions of network damage have evolved. *Robustness* refers to the connectedness of a network, whether it has broken into disconnected sub-networks or still functions as a cohesive unit. *Efficiency* (referred to by some as *responsiveness*) determines how quickly a message (or people or freight) can travel from one node to another. *Flexibility* refers to the network's inherent ability to find alternative paths if the shortest paths are destroyed. The most prevalent metrics used to express these dimensions in single layer complex network studies are listed in Table 2.1.

Table 2.1: Common metrics used to measure network damage along the dimensions of robustness, efficiency and flexibility in single-layer complex network studies.

Dimension	**Metrics**
Robustness	% of nodes in the Largest Connected Component (LCC) (P_∞); and critical point (p_c).
Efficiency	Diameter of the LCC; Average shortest path length for all connected node-pairs; and Inverse efficiency indicator.
Flexibility	Average clustering coefficient.

The LCC is the largest subset of network nodes that has an undirected path between all node-pairs. The fraction of all network nodes in the LCC is referred to as P_∞ and is the benchmark for measuring network robustness. Essentially when $P_\infty \to 1$, messages from any node have a high probability of reaching any other node of the network[1]. In this case the LCC is also called the *giant connected component*. As P_∞ reduces, nodes become disconnected and can no longer be reached by other network nodes. Below some value of P_∞ the network would become dysfunctional due to increasing disconnectedness (Danziger et al., 2014). However, the threshold value of P_∞ is instance-specific and depends greatly on the field of application (social systems, transportation, neural networks etc.).

[1]Danziger et al. (2014) use the symbol $\sim$ instead of $\to$ to indicate a variable approaching a value. The intended audience of this thesis would better relate to the use of $\to$ than $\sim$ in this regard.

Percolation theory is concerned with determining $P_\infty(p)$ after a random fraction $1-p$ of nodes have been removed and defining the critical point p_c such that for $p > p_c$, $P_\infty(p) > 0$; as $p \to p_c$, $P_\infty(p) \to 0$; and finally $P_\infty(p) \equiv 0$ when $p < p_c$ (Danziger et al., 2014).

It has been shown, in single layer networks, that $p_c \to 0$ for scale-free networks, meaning that in cases of random failures, nearly *all* of the nodes have to be removed before the LCC becomes disconnected (Cohen et al., 2000). Scale-free networks are therefore highly robust to random failures. Random networks, on the other hand, are highly vulnerable with $p_c = 1/\langle k \rangle$ where $\langle k \rangle$ is the average network degree[2] (Danziger et al., 2014).

A significant caveat is that percolation (and the determination of p_c) refers to random failures and not targeted attacks. Cohen et al. (2001) showed that under targeted attack scale-free networks disintegrate long before the calculated p_c is reached. In fact, Pastoras-Satorras and Vespignani (2001) proved that a p_c that takes into account targeted attack cannot be analytically determined. Despite the fact that p_c is not always a relevant indicator, measuring the size of the LCC remains the most popular robustness metric.

Notably, alternative metrics for measuring robustness are emerging. For example in their study of *network skeletons*, Shekhtman et al. (2014) showed that the size of the skeleton could also be a valid interpretation of robustness. When extracting network skeletons, scientists seek to remove the "noise" by judiciously removing links and/or nodes. The result is a minimal subset of links and nodes that still represents all the key topological characteristics of the original network.

In terms of efficiency, the average shortest path length is only defined as long as all node-pairs are connected. That is, when $P_\infty \equiv 1$. This is hardly the case for real-life networks and certainly not the case as networks undergo successive disruptions. Therefore some studies have reverted to measuring the average shortest path length of the LCC. However, if the LCC shrinks significantly the averages are no longer comparable. The efficiency indicator of Berche et al. (2009) is therefore a good workaround for disconnected networks. It uses a summation of inverse lengths so that disconnected pairs can still affect the measurement.

Of all three dimensions, the metric for flexibility is the most abstract. The clustering coefficient is the ratio of the actual links to the theoretically possible links between the first order neighbours of a node. The average clustering coefficient is then a measure of the tightness of clusters in a network. To date it has been accepted as a measure that indicates the probability of alternative paths in a network (Thadakamalla et al., 2004; Viljoen and Joubert, 2016). Admittedly, there can be exceptions to this rule. In lieu of a better measure of flexibility, the average clustering coefficient remains the standard.

Although the majority of single-layer vulnerability studies used these dimensions, there is merit in deviating from the standard if the context calls for it. The work of Zhao et al. (2011) is a good example. In their application to military supply chains they proposed that the heterogeneity of demand and supply nodes and the function of the network would be disregarded if the traditional dimensions and metrics were used. Instead they suggested *availability*, *connectivity* and *accessibility* and modelled these with encouraging results.

Multilayered networks behave very differently to single-layer networks. Researchers have shown, both analytically and empirically, that one cannot simply deduce multilayered characteristics by studying the single-layer components in isolation. This is particularly true for the topics of resilience and spreading processes (Boccaletti et al., 2014; Danziger

[2]Danziger et al. (2014) and many other network scientists use $\langle k \rangle$ to indicate average network degree. We recognise that for the audience of this thesis it may be an unusual convention.

et al., 2014; Kivelä et al., 2014; Lee et al., 2015; Salehi et al., 2015).

In multilayered networks, the concept of cascading failures has been the foremost phenomenon in vulnerability studies. If the layers of a multilayered network are interdependent then the links between the layers are called dependency links. Removing any node in one layer would then trigger the removal of all nodes in all other layers that are connected via dependency links. Such failures cascade back and forth until all dependent nodes are removed. In studying multilayered vulnerability, scientists have focussed on the calculation of the *Mutually Connected Giant Component (MCGC)* and the critical percolation properties of cascading failures ((Boccaletti et al., 2014; Kivelä et al., 2014), and references therein).

Studies that have used the concepts of cascading failures and percolation to investigate vulnerability made three *a priori* assumptions:

1. the layers of the multilayered network were interdependent, resulting in cascading phenomena;

2. nodes were removed during failures or attacks; and

3. a network was considered robust when P_∞ was larger than some fraction.

If these *a priori* assumptions are not present in a specific problem context, it is necessary to find different means to measure the robustness of a multilayered network.

So far the discourse on multilayered vulnerability has centred only around robustness. Efficiency and flexibility have not been discussed as pertinently.

Once it has been established how network damage will be measured, simulations must be designed to progressively disrupt the networks. Firstly, one must determine whether the simulation will use *random errors* or *targeted attacks*. Random error simulations are mostly used to establish a baseline or control of a network's vulnerability. Targeted attacks are used either to assess a network's vulnerability to a plausible real-life threat or when scientists want to investigate the best way to identify and prioritise the critical elements of a network.

2.3.2 Targeted attack strategies

There are three decisions to be made when designing a targeted attack:

1. Will nodes or links be removed in each disruption?

2. Which metric will be used to prioritise critical elements for removal?

3. Will this prioritisation be dynamic (i.e. recalculated after each disruption) or static?

The most prevalent method of prioritisation has been based on the concept of *centrality*. Centrality has been interpreted in many ways for network nodes with metrics like degree centrality, betweenness, closeness, eccentricity and so on. This abundance of node-based metrics is part of the reason why the majority of vulnerability studies have concentrated on node removal. The other reason is that node removal makes sense, contextually, in relational networks such as those found in sociology, epidemiology and even supply chain networks. Link removal studies have recently picked up the pace in spatial networks. In transportation networks especially it is far more likely that a link would be removed/damaged than a node (Viljoen and Joubert, 2016).

Initially defined for nodes, some centrality measures have been adapted for links such as link betweenness and degree product (Girvan and Newman, 2002; He et al., 2009; Holme, 2002; Travieso and da Fontoura Costa, 2012). Other approaches that use node-based centrality metrics to prioritise links have either inverted the network to a dual representation where urban streets are modelled as nodes while intersections are modelled as edges (Demšar et al., 2008; Jiang and Claramunt, 2004; Porta et al., 2006; Tomko et al., 2008) or have prioritised those links associated with central nodes (Zhang et al., 2007). In the former technique node-based metrics were directly applied to streets while in the latter streets were "guilty by association" with their incident nodes.

Apart from centrality metrics, some authors have used metrics relating to distance (Ortigosa and Menendez, 2014) or geodesic range (Motter et al., 2002).

Another method of prioritisation stemmed from the study of network skeletons (Grady et al., 2012; Serrano et al., 2009; Shekhtman et al., 2014). Shekhtman et al. (2014) were the first to use network skeletons — specifically the salience network and disparity backbone — to prioritise links in vulnerability studies. Their study concluded that while skeletons remained capable to present global network statistics after perturbation, the way in which the skeletons morph in response to disruptions remained to be understood. Viljoen and Joubert (2016) compared the efficacy of targeting salient links versus links with high betweenness in the global container shipping industry. They found that removing salient links quickly reduced commonality between shortest paths, which meant reducing consolidation options for shipping liners. However, by the same mechanism the salience strategy degraded the very skeleton used to prioritise links, therefore the strategy's efficacy was short-lived.

More recently, Travieso and da Fontoura Costa (2012) argued that typical centrality approaches do not adequately reflect the connectivity of the network. *Spectral decomposition* (or eigendecomposition) presents a new range of metrics (such as eigenvector centrality) that better characterise cycles, modularity and cuts in the network. In comparison to other techniques, targeting links according to their spectral decomposition metrics had a more devastating effect on overall network clustering. Unfortunately, spectral decomposition is suited to networks in which divisible objects (such as information or infection) flow between nodes and not indivisible objects (such as commercial vehicles). Therefore it is not directly relevant to transportation studies (Zadeh and Rajabi, 2013).

Having selected an appropriate metric to prioritise critical links, it must be decided whether this metric will be recalculated between disruptions. The critical elements to be removed next are then based on this recalculated value. In node-based removal strategies the results generally show the dynamic approach decisively more effective than an approach based on initial values (Berche et al., 2012; Holme, 2002; Nie et al., 2015). Some studies have also shown the dynamic approach to be more effective in the case of link-based removal (Holme, 2002; Nie et al., 2015). There are, however, studies such as the one by Berche et al. (2012) that do not find the dynamic approach decisively more effective in the case of link-based removal. They studied 5 node-based and 5 link-based removal strategies targeting public transport systems across 14 global cities. Their results show that dynamic link-based strategies are sometimes more effective and other times not compared to the static strategies.

However, the choice between dynamic and static prioritisation does not solely depend on which approach degrades a network quickest. It depends on the question to be answered by the vulnerability study. One would use static prioritisation if one is interested in the longevity of that characteristic's ability to identify vulnerability. In other words, if the

initial prioritisation according to a characteristic continues to effectively destroy a network in successive disruptions, it would imply that the power of that characteristic is robust to the changes the network undergoes as it is destroyed. On the flip side, using an approach that is as destructive as possible allows one to explore the boundaries of a network's vulnerability.

2.4 Conclusion

CNT has been successfully applied in both the supply chain and road transport domains. In road transport (and urban transport in general) multilayered CNT is providing breakthroughs in multimodal and multi-service models such as public transit. Topological vulnerability studies have also yielded great insights in both these domains and the field has matured to offer benchmark metrics for network damage and targeted attack strategies.

In supply chain literature most studies focus on single-layer representations of supply chain relationships. The multilayered dependency on infrastructure networks has not been studied. This thesis targets a yet unexplored domain within supply chain applications of CNT.

In terms of road network applications there remains a gap in recent literature in terms of studying freight transport and its interaction with the road network. Freight transport represents a small proportion of traffic. However, when one considers congestion, emissions and infrastructure damage the contribution of commercial vehicles is disproportionately large. Equally so, the positive economic impact of this "relatively small" segment of traffic is also significant. Therefore there is justification to understand the behaviour of freight transport in its own right, not just as an inflation of passenger transport. The multilayered scenario studied in this thesis is novel. It considers freight transport activity on the urban road network instead of passenger traffic in multimodal transit systems.

Chapter 3

Multilayered network instances

This chapter describes the conceptual structure of the multilayered network by describing the *logical* and *physical* network layers and how these interrelate. It then presents the adaptation of a generic multilayered network formulation to capture this conceptual structure. Next the generation of random instances for each of the supply chain network archetypes is explained. The collection of shortest path sets that result from each randomly generated instance is unique to that instance. A formulation is proposed to describe these collections and their two key statistics namely the shortest path length and shortest path set size. Lastly, the distributions of these statistics for the initial networks were analysed to investigate the differences between the archetypes as well as between distinct instances of the same archetype.

3.1 Conceptual structure of the multilayer network

Before delving into the mathematics, we describe the two layers of the multilayered network conceptually.

3.1.1 Urban road network (*physical*) layer

The road network was presented as a regular grid. This simplification is typical in theoretical models of urban road grids as it is a good approximation of the general topology of cities around the world. Similar to the representation in Ortigosa and Menendez (2014), the road network was a directed, unweighted grid network with m rows and n columns, and the nodes were numbered sequentially from 1 to $m \times n$ as shown in Figure 3.1. Nodes were connected with two directed, opposite links, $\leftrightarrows$, instead of one undirected link, $\leftrightarrow$. The assumption was that should a road segment in one direction fail, the associated lane in the opposite direction would not necessarily be affected as well. This occurs, for example, when the two opposing lanes of a road are separated by a median strip.

3.1.2 Supply chain (*logical*) layer

Hearnshaw and Wilson (2013) proposed that efficient supply chains have a scale-free topology. This is typically referred to as a hub-and-spoke structure. In industry, hub-and-spoke supply chain designs are indeed prevalent as consolidation and economies-of-scale are tenets of an efficiency-focussed supply chain mindset (Christopher and Holweg, 2011). However, Hearnshaw and Wilson (2013) conceded that small-world characteristics

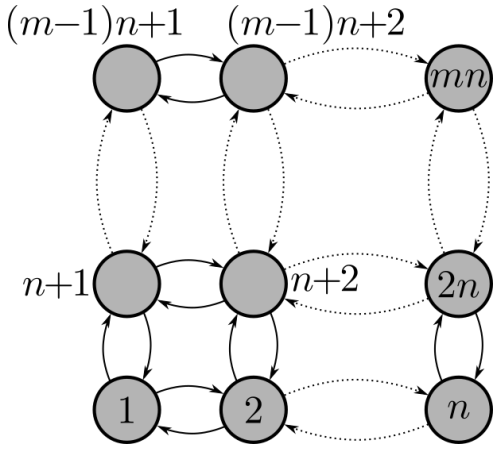


Figure 3.1: Unweighted, directed grid layouts to approximate the urban road network in an urban context.

like short average path length and high clustering could enhance efficiency. Basole and Bellamy (2014); Basole et al. (2016) and Thadakamalla et al. (2004) also promoted small-world supply chain designs but for the sake of better resilience, not efficiency.

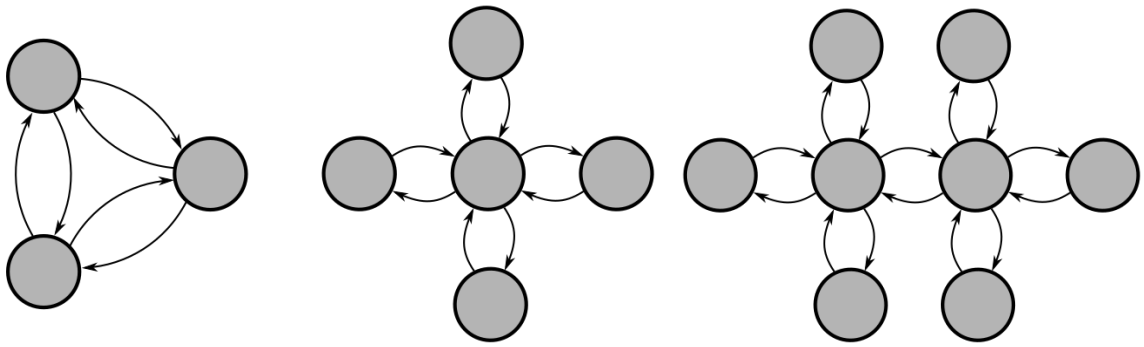
Reliable empirical data describing the true topologies of supply chain networks among facilities in an urban environment is lacking. In absence of such real-life knowledge, we postulated three theoretical supply chain network design archetypes (hereafter referred to as 'archetypes') based on the insights from literature. For the same reason we also assumed that supply chain facilities fulfil the same function. Therefore, we did not distinguish between supply and demand facilities, or between warehouses, retail stores, manufacturing facilities and so on.

The Fully Connected (FC) archetype represented a network of facilities where all facilities could ship freight to and receive freight from all other facilities. This archetype was represented as a fully connected, unweighted, directed graph (Figure 3.2a). Each node-pair was connected by bi-directional links. By definition this network had an average shortest path length of one. In addition, because each node was connected to every other node in the network the clustering coefficient of each node was one. The clustering coefficient of a node is calculated as the fraction of actual versus theoretically possible connections between the first order neighbours of a node. The average clustering coefficient of the FC archetype was thus also one, reflecting the dense connectedness of the network. The FC archetype was one extreme on the spectrum between small-world and scale-free.

The Single Hub (SH) archetype represented a network where all facilities were connected to one single hub facility, but not to each other. The archetype was a typical star-network and lay on the other extreme between small-world and scale-free (Figure 3.2b). The clustering coefficient was, by definition, zero and the average shortest path length approached the diameter of the network.

The Double Hub (DH) archetype consisted of two hub facilities, each with their own spoke facilities, connected to each other through bi-directional links. Each hub was connected to its spoke facilities by means of bi-directional links, but the spoke facilities were not connected to one another (Figure 3.2c). In essence it was two SH networks connected at the hubs. The clustering coefficient of this network was also, by definition, zero. The concentration of centrality, however, was now split between two hubs. It lay between the

FC and SH archetypes on the spectrum, albeit very close to the SH network.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 3.2: Conceptual representation of the three supply chain network archetypes.

With this collection of network archetypes, the thesis explored the performance on the boundaries of the topology spectrum. Granted there are many potential topologies between the FC and SH network and the choice of the DH may seem arbitrary. The assertion is that more real-life networks resemble hub-and-spoke networks with distributed hubs than densely connected networks. Therefore, we wanted to concentrate the study on hub-and-spoke topologies using the FC topology as a comparative archetype. The choice of a DH archetype instead of networks with three or more hubs that would be further along the spectrum towards the FC archetype was also intentional. It seemed like a small change to divide the centrality of the network from one hub into two, but it was suspected that it could cause step-changes in the vulnerability metrics related to shortest paths. This was a phenomenon worth monitoring. There is scope for further research to investigate many more network archetypes, but time and computational resources limited this thesis to three.

3.2 Multilayered network formulation

With multilayered research exploding simultaneously amongst different groups of network scientists, a diverse cloud of definitions and mathematical formulations have mushroomed around the topic. Although there has been effort to cement a common vernacular, scientists have not yet reached a point of consensus and one is faced with many different alternatives to describe and formulate multilayered networks.

3.2.1 Generic multilayered formulation

The formulation in this thesis was based on the generic formulation of Boccaletti et al. (2014) as it was intuitive to the application at hand. The *multilayer network* was a pair $\mathcal{M} = (\mathcal{G}, \mathcal{C})$ where $\mathcal{G} = \{G_b; b \in \{1, 2, \ldots, M\}\}$ was a family of M individual graphs $G_b = (X_b, E_b)$ which each represented a layer of $\mathcal{M}$.

In this generic formulation, α and β referred to layers of $\mathcal{G}$ such that $\alpha, \beta \in \{1, 2, \ldots, M\}$ and $\alpha \neq \beta$. The set of nodes in a layer G_α were denoted by $X_\alpha = \left\{x_1^\alpha, \ldots, x_{N_\alpha}^\alpha\right\}$ and $E_\alpha \subseteq X_\alpha \times X_\alpha$. The set of interconnections between nodes in G_α and G_β with $\alpha \neq \beta$

were defined by

$$\mathcal{C} = \{E_{\alpha,\beta} \subseteq X_\alpha \times X_\beta; \forall \alpha, \beta \in \{1, \ldots, M\}, \alpha \neq \beta\} \tag{3.1}$$

Therefore the elements of $E_{\alpha,\beta}, \alpha \neq \beta$ were *interlayer* connections while elements of E_α and E_β were the *intralayer* connections.

3.2.2 Customised multilayered formulation

For this thesis we adapted the generic formulation. One universal adaptation was that the indices that named the different network layers (α and β in the generic formulation) were superscripts in the customised formulation instead of subscripts. This was necessary to avoid confusion with node indices which were (as per convention) indicated as subscripts.

Let $\mathcal{M} = (\mathcal{G}, \mathcal{C})$ be the multilayered network where $\mathcal{G} = (G^{1K}, G^2)$. G^{1K} represented the *logical* layer with $K \in \{F, S, D\}$ such that F, S and D denoted the FC, SH and DH archetypes, respectively. There were 12 nodes in each of the archetype models. The nodes[1] of G^{1K} were thus defined by:

$$N^{1K} = 12 \qquad \forall K \in \{F, S, D\} \tag{3.2}$$
$$X^{1K} = \{x_1^{1K}, x_2^{1K}, \ldots, x_{N^{1K}}^{1K}\} \qquad \forall K \in \{F, S, D\} \tag{3.3}$$
$$E^{1K} = \{e_{ij}^{1K}\}\ \forall i, j \in \{1, 2, \ldots, N^{1K}\} | i \neq j, \qquad \forall K \in \{F, S, D\} \tag{3.4}$$

where

$$e_{ij}^{1K} = \begin{cases} 1 & \text{if } x_i^{1K} \text{ was connected to } x_j^{1K} \quad \forall K \in \{F, S, D\} \\ 0 & \text{otherwise.} \end{cases} \tag{3.5}$$

G^2 represented the road network with $m = n = 10$ as illustrated in Figure 3.3. The nodes of G^2 were defined by:

$$N^2 = m \times n = 100 \tag{3.6}$$
$$X^2 = \{x_1^2, x_2^2, \ldots, x_{N^2}^2\} \tag{3.7}$$
$$E^2 = \{e_{st}^2\} \qquad \forall s, t \in \{1, 2, \ldots, N^2\} | s \neq t \tag{3.8}$$

where

$$e_{st}^2 = \begin{cases} 1 & \text{if } x_s^2 \text{ was connected to } x_t^2 \\ 0 & \text{otherwise.} \end{cases} \tag{3.9}$$

[1] Throughout this thesis there is no comma between node indices in the subscript unless one or both of the indices were double digits.

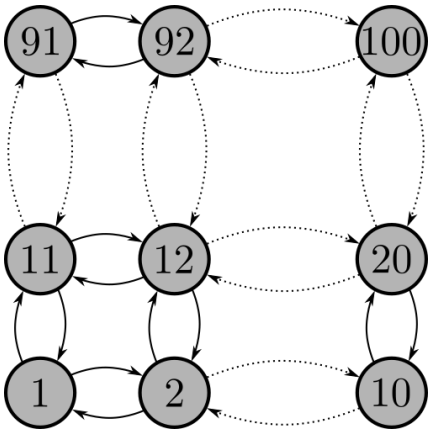


Figure 3.3: G^2 — the 10×10 directed, unweighted representation of the road network.

3.3 Sample generation of multilayered network instances

To generate a multilayered network instance $\mathcal{M}$, the *logical* layer G^{1K} had to be placed onto the *physical* layer G^2 by associating each node in X^{1K} with exactly one node in X^2. Simplifying assumptions made were that the logistics facilities correspond with their nearest intersections on the grid and that at most one logistics facility is associated with each intersection. These associations were the *interlayer* connections denoted by $E^{1K,2}$. The adjacency matrix of the *interlayer* links $E^{1K,2}$ was defined as $A^{[1K,2]} = (a_{is}^{1K,2})$, where:

$$a_{is}^{1K,2} = \begin{cases} 1, & \text{if } (x_i^{1K}, x_s^2) \in E^{1K,2} \quad \forall i \in \{1,2,\ldots,N^{1K}\}, s \in \{1,2,\ldots,N^2\} \\ 0, & \text{otherwise} \end{cases} \tag{3.10}$$

The pseudocode in Algorithm 1 shows how the associations were randomly generated to produce multilayered network instances of G^{1F} on G^2. In this manner, a sample of 500 random instances was generated for the FC archetype.

Algorithm 1: Random generation of $A^{[1F,2]}$

Input : G^{1F}, G^2
Output: $A^{[1F,2]}$

```
used ← NULL //vector that stores x_i^2 already assigned;
A^[1F,2] ← NULL //set all elements in matrix as unassigned;
for x_i^1F ∈ X^1F do
    continue =TRUE;
    while continue do
        randVertex ← randomly selected x_j^2 ∈ X^2;
        if randVertex ∉ used then
            a_ij^[1F,2] ← 1 ;
            Append used with randVertex;
            continue = FALSE;
return A^[1F,2]
```

Multilayered network instances of G^{1S} on G^2 were generated similarly as shown in Algorithm 2. A sample of 500 instances was generated for the SH archetype.

Algorithm 2: Random generation of $A^{[1S,2]}$

Input : G^{1S}, G^2
Output: $A^{[1S,2]}$

```
used = NULL //vector that stores x_i^2 already assigned;
A^[1S,2] = NULL//set all elements in matrix as unassigned;
for x_i^1S ∈ X^1S do
    continue =TRUE;
    while continue do
        randVertex ← randomly selected x_j^2 ∈ X^2;
        if randVertex ∉ used then
            a_ij^[1S,2] ← 1 ;
            Append used with randVertex;
            continue = FALSE;
return A^[1S,2]
```

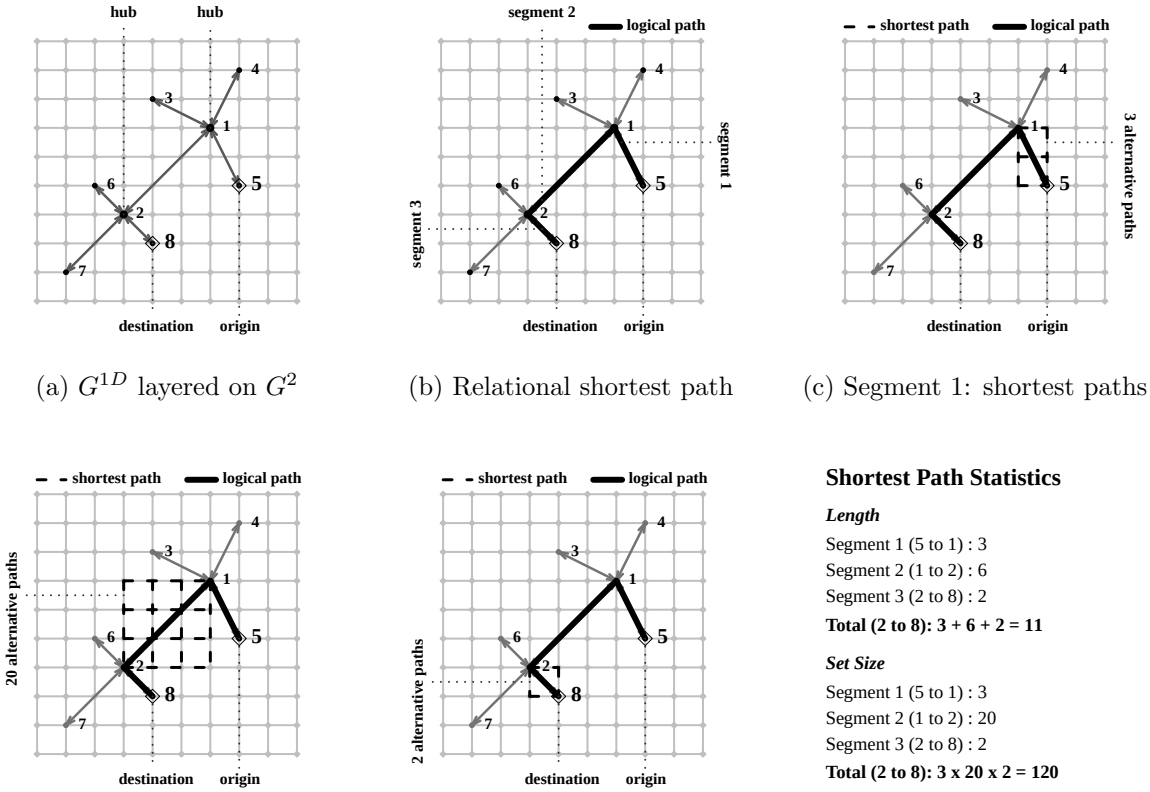
Finally, generating multilayered instances of G^{1D} on G^2 was slightly different as the assignment of nodes to hubs had to be constrained. The two hubs were first associated with grid nodes. Thereafter, the nodes assigned to a hub had to be placed on the grid such that the distance to the hub (along G^2) was shorter than or equal to the distance to the other hub. Algorithm 3 displays the pseudocode.

The probability of generating identical multilayered instances was very small. Therefore, although the algorithms did not explicitly prevent the generation of duplicate instances, no duplicates were produced for any of the three archetypes. A description of the data files used as input and produced as output is described in the working paper Viljoen and Joubert (2017). The datasets are also published on Mendeley (Joubert and Viljoen, 2017).

3.4 Formulation of the collection of shortest path sets

The shortest *relational* path between any two nodes in G^{1K} was predefined. However, freight cannot move along relational paths, it requires the road infrastructure. Thus, when calculating the shortest *freight movement* path between two nodes in G^{1K}, the positions of these nodes on G^2 had to be taken into account.

Figure 3.4 shows the example of G^{1D} layered on G^2 with nodes x_5^{1D} and x_8^{1D} identified as origin and destination, respectively. If only the relational path were considered, the shortest path between x_5^{1D} and x_8^{1D} would have consisted of three segments: $x_5^{1D}\rightarrow x_1^{1D}\rightarrow x_2^{1D}\rightarrow x_8^{1D}$ (Figure 3.4b). However, the physical constraints of the infrastructure also had to be regarded. Where there was only one shortest relational path between the node-pair G^{1K}, there could have been be multiple alternatives in the multilayered network. Figure 3.4c presents the three alternative shortest paths between x_5^{1D} and x_1^{1D}, each of length 3. Similarly, there were 20 alternative shortest paths between x_1^{1D} and x_2^{1D}, each of length 6 (Figure 3.4d) and 2 alternative shortest paths between x_2^{1D} and x_8^{1D}, each of length 2 (Figure 3.4e). The length of the shortest paths between node x_5^{1D} and x_8^{1D} was thus 11 and there were 120 alternatives.



(a) G^{1D} layered on G^2 (b) Relational shortest path (c) Segment 1: shortest paths

(d) Segment 2: shortest paths (e) Segment 3: shortest paths (f) Calculations

Figure 3.4: Calculating the length and number of shortest paths between a node-pair in $\mathcal{M}$ by adhering to both relational and physical constraints.

All metrics pertaining to shortest paths referred to a specific realisation of $\mathcal{M}$, therefore in the definitions that follow we dropped the superscripts relating to the layers and network archetype for simplicity's sake. Generally:

$$\mathcal{S}_{ij} = \{SD_{ij}, SI_{ij}\} \tag{3.11}$$

where SD_{ij} was the subset of all shortest path sets between node-pairs (x_i^{1K}, x_j^{1K}) that were directly connected in G^{1K}:

$$SD_{ij} = \{s^1, s^2, \ldots, s^{P_{ij}}\} \qquad \forall e_{ij}^{1K} \in E^{1K} \tag{3.12}$$

and SI_{ij} was the subset of all shortest path sets between node-pairs that were indirectly connected in G^{1K}:

$$SI_{ij} = \{s^1, s^2, \ldots, s^{P_{ij}}\} \qquad \forall e_{ij}^{1K} \notin E^{1K} \tag{3.13}$$

We then defined the collection of shortest paths as

$$\mathcal{C}(\mathcal{S}_{ij}) = \bigcup_{i,j} \mathcal{S}_{ij} \qquad \forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j \tag{3.14}$$

There were two statistics of interest namely the length of a shortest path and the number

of alternative shortest paths between two nodes. Therefore:

$$L_{ij} \equiv \text{length of a shortest path between node } x_i^{1K} \text{ and } x_j^{1K} \quad \forall i,j \in \{1,2,\ldots,N^{1K}\}, i \neq j \tag{3.15}$$

$$P_{ij} \equiv \text{number of alternative shortest paths between node } x_i^{1K} \text{ and } x_j^{1K} \quad \forall i,j \in \{1,2,\ldots,N^{1K}\}, i \neq j \tag{3.16}$$

The efficiency of $\mathcal{M}$, which was the average shortest path length, was:

$$\bar{L} = \frac{\Sigma_{i,j,i\neq j} L_{ij}}{N^{1K}(N^{1K}-1)} \text{ where } i,j \in \{1,2,\ldots,N^{1K}\} \tag{3.17}$$

where N^{1K} was the number of nodes in G^{1K}.

3.5 Shortest path statistics of the initial datasets

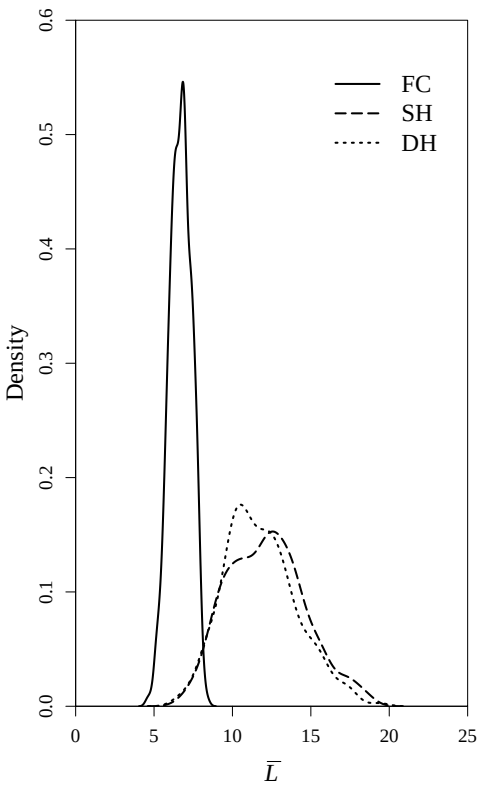
The initial characteristics of $\mathcal{C}(\mathcal{S}_{ij})$ depended greatly on the network archetype. We specifically compared the distributions of the initial shortest path lengths and the set sizes.

3.5.1 Initial distribution of the shortest path length

Figure 3.5a plots the distributions of $\bar{L}$, showing that the FC archetype had significantly shorter paths than the other two archetypes. This directly resulted from the fact that all node-pairs were directly connected, not requiring rerouting through a hub node. A Kolmogorov-Smirnov test (KS-test) comparing the distributions of $\bar{L}$ for the SH and DH archetypes rejected the null hypothesis that the distributions were similar with $p = 0.0047$. The lower mean and wider spread in both tails of the DH archetype was explained by the structure of G_1^D. Intra-hub paths (paths between nodes that shared a common hub node) in the DH archetype were generally shorter than intra-hub paths of the SH archetype. This was because in the DH archetype the placement of the two hubs on the grid effectively split the grid and therefore the intra-hub nodes were closer to their respective hubs. On the other hand, the inter-hub movements were longer as they had to be rerouted through two hubs. Intra-hub paths accounted for 45% of the network, inter-hub for 38% and the remainder of the paths were links between one hub and a node assigned to the other hub and vice versa. Paths in this last group were not distinctly longer or shorter between the two archetypes. As the majority of the links in the DH archetype had shorter lengths, $\bar{L}$ was lower.

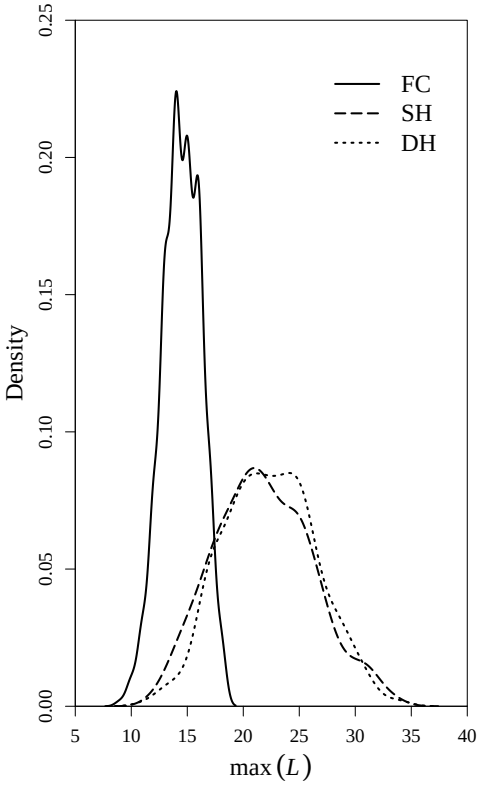
The diameter of a network is the length of its longest shortest path denoted by $\max(L)$. Figure 3.5b plots the distributions of $\max(L)$, where once again the FC archetype showed far shorter paths than the others. This time, the KS-test failed to reject the null hypothesis that the distributions of $\max(L)$ for DH and SH archetypes were similar with $p = 0.29$.

The FC archetype was thus most efficient by a large margin, whereas the clear winner between the two hub archetypes was instance-specific, with a slight prejudice in favour of the DH archetype.



Archetypes	Min	Mean	Max	Kurtosis
FC	4.6	6.8	8.4	2.6
SH	6.5	12.2	18.8	2.7
DH	6.4	11.8	19.5	3.0

(a) Average shortest path length, $\bar{L}$



Archetype	Min	Mean	Max	Kurtosis
FC	9.0	14.5	18.0	2.7
SH	12.0	21.8	34.0	2.7
DH	11.0	22.3	34.0	2.7

(b) Diameter, max (L)

Figure 3.5: Distributions of the diameter and average shortest path lengths for each of the three archetypes.

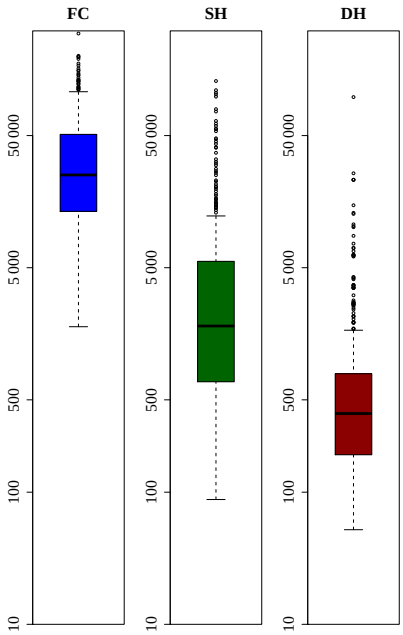
3.5.2 Initial distributions of the shortest path set sizes

The shortest path set size (P_{ij}) for two directly connected nodes was the combinatorial product of the number of columns and rows of G^2 traversed when moving from one node to the other (refer to Figure 3.4). Furthermore, P_{ij} for any indirectly connected node-pair was the combinatorial product of the shortest path set sizes of all the directly connected node-pairs that constituted the logical path in G^{1K}.

The sum of the shortest path set sizes of an instance were denoted by $\sum_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$ and $\sum_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$, when considering only the direct connections (SD_{ij}) and full network ($\mathcal{S}_{ij}$), respectively. Figure 3.6a shows the box plots of the distributions of the sums over the directly connected node-pairs (SD_{ij}) in each instance, according to the archetypes. The distributions had very long tails as confirmed by the high kurtosis values. A kurtosis value of 3 indicates that a distribution is mesokurtic, being no more likely to produce outliers than a normal distribution. A value greater than three indicates a leptokurtic distribution that has a greater degree of "tailedness" than the normal distribution and kurtosis less than 3 indicates a platykurtic distribution with a smaller likelihood than the

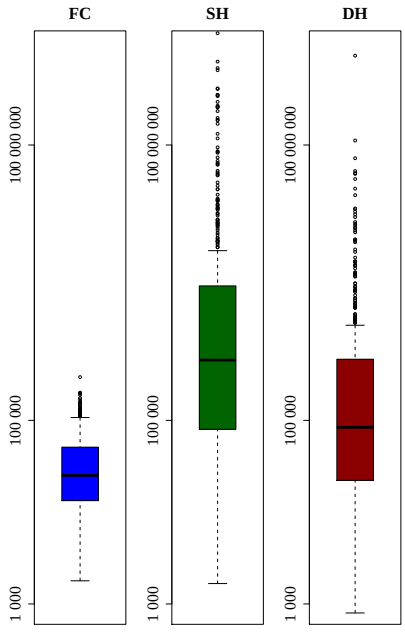
normal distribution to produce outliers.

For the FC archetype all nodes were directly connected thus the relevant distributions in Figures 3.6a and 3.6b are the same. For the hub archetypes the distributions in Figure 3.6b were disproportionately wider due to the fact that 83.3% of the node-pairs were indirectly linked in these archetypes and each of these sets SI_{ij} was the product of the set sizes of its component sets SD_{ij}. The DH archetype had the smallest set sizes for directly connected node-pairs, owing to the fact that its directly connected nodes were closer together on the grid. Interestingly, the kurtosis of SD_{ij} for the DH archetype also showed that it was far more likely to produce outliers than the other two archetypes.



Archetype	Median	Max	Kurtosis
FC	25 171	3.0×10^5	10
SH	1 810	1.3×10^5	30
DH	394	9.8×10^4	265

(a) Distributions of $\sum_{i,j \in SD_{ij}; i \neq j} P_{ij}$



Archetype	Median	Max	Kurtosis
FC	25 171	3.0×10^5	10
SH	453 765	1.7×10^9	134
DH	84 021	9.4×10^8	465

(b) Distributions of $\sum_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$

Figure 3.6: Analysis of the distributions of the sum of the shortest path set sizes for direct paths (SD_{ij}) and the full network ($\mathcal{S}_{ij}$)

When regarding only shortest path lengths one could have concluded that there wasn't great variance between the different instances of a single archetype, nor was there really a pronounced difference between the two hub archetypes. This was definitely not true when considering the sizes of the shortest path sets. One randomly generated instance of $\mathcal{M}$ could have had vastly different set sizes than the next randomly generated instance.

This chapter showcased the first artefact of the thesis: a multilayered complex network model to capture the dependence of a supply chain network on an urban road network. It also described how a sample of random instances were generated and discussed the characteristics of the initial collections of shortest paths. These initial instances were also referred to as the "undisturbed" instances and presented the baseline for the targeted

attack and random error simulations discussed in the following chapters.

Algorithm 3: Random generation of $A^{[1D,2]}$

Input : G^{1D}, G^2
Output: $A^{[1D,2]}$

$used = NULL$ //vector that stores x_j^2 already assigned;
$A^{[1D,2]} = NULL$//set all elements in matrix as unassigned;
//Assign hubs first
for $i \leftarrow 1$ **to** 2 **do**
    continue =TRUE;
    while *continue* **do**
        $randVertex \leftarrow$ randomly selected $x_j^2 \in X^2$;
        if $(randVertex \notin used)$ **then**
            $a_{ij}^{[1D,2]} \leftarrow 1$;
            Append *used* with $randVertex$;
            continue =FALSE;
//Assign remaining vertices
for $i \leftarrow 3$ **to** N_1 **do**
    continue =TRUE;
    while *continue* **do**
        $randVertex \leftarrow$ randomly selected $x_j^2 \in X^2$;
        if $randVertex \notin used$ **then**
            $Dist1 \leftarrow$ Dijkstra's shortest path on G^2 between $randVertex$ and x_1^{1D};
            $Dist2 \leftarrow$ Dijkstra's shortest path on G^2 between $randVertex$ and x_2^{1D};
            if $i \leq 2 + (N^{1S} - 2)/2$ **then** //vertices around hub 1//
                if $Dist1 \leq Dist2$ **then** //distance to hub 1 must be smaller or equal to distance to hub 2//
                    $a_{ij}^{[1D,2]} \leftarrow 1$;
                    Append *used* with $randVertex$;
                    continue =FALSE;
            else //vertices around hub 2//
                if $Dist2 \leq Dist1$ **then** //distance to hub 2 must be smaller or equal to distance to hub 1//
                    $a_{ij}^{[1D,2]} \leftarrow 1$;
                    Append *used* with $randVertex$;
                    continue = FALSE;
return $A^{[1D,2]}$

Chapter 4

Discovering vulnerability characteristics

In topology-based vulnerability studies, targeted attack simulations are frequently used to identify the most critical or most vulnerable elements in a network. A targeted attack simulation selects nodes or links to remove based on some characteristic — for example degree centrality or node betweenness. If the simulation causes significant damage, one can know that the characteristic used to prioritise elements for removal is a key determinant of the network's vulnerability.

In this thesis three link-based targeted attack strategies were used to investigate which characteristics of the collection of shortest path sets, $\mathcal{C}(\mathcal{S}_{ij})$, could be most effective in identifying critical links in G^2. The analysis of the results of the three simulations paved the way for the development of a suite of vulnerability metrics.

4.1 Link-based targeted attack simulations

An illustrative example of a simple supply chain network around the O.R. Tambo International Airport near Johannesburg, South Africa is introduced to facilitate the description of the targeted attack strategies. Imagine a supply chain network that contains three facilities namely an Airfreight Warehouse, a Distribution Centre and a Retail Outlet. The Distribution Centre is an intermediary between the Airfreight Warehouse and the Retail Outlet. Therefore, freight can be shipped from the Airfreight Warehouse to the Distribution Centre and visa versa. Similarly, freight can be shipped from the Distribution Centre to the Retail Outlet and visa versa. However, freight cannot be directly shipped between the Retail Outlet and the Airfreight Warehouse. Figure 4.1 shows these three facilities and their logical relationships. In the terminology of this thesis this supply chain network represents a Single Hub (SH) network with three nodes (G^{1S}) where the Airfreight Warehouse and the Distribution Centre as well as the Distribution Centre and the Retail Outlet are directly connected, while the Airfreight Warehouse and Distribution Centre are indirectly connected.

To keep the illustration simple we assume that there are only three shortest path alternatives between the Airfreight Warehouse and the Distribution Centre, namely *Path 1*, *Path 2* and *Path 3* as indicated in Figure 4.2. Similarly, there are also only three shortest paths in the opposite direction from the Distribution Centre to the Airfreight Warehouse (*Paths 4, 5* and *6*). We also assume that there are only two shortest path alternatives, namely *Path 7* and *Path 8* between the Distribution Centre and the Retail

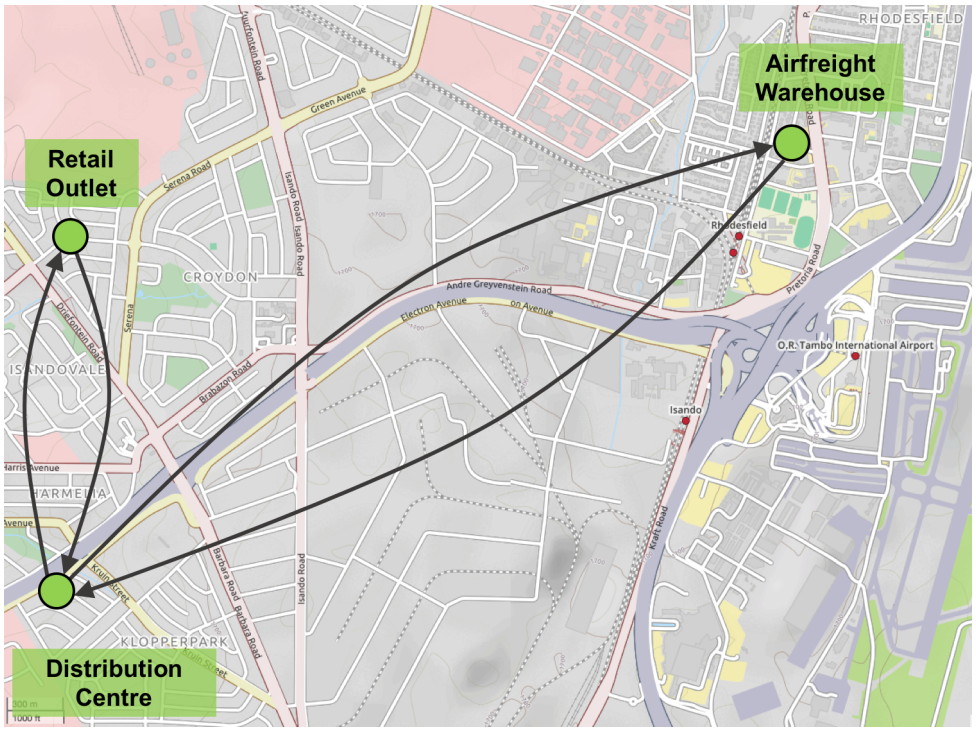


Figure 4.1: An illustrative example of a supply chain network around O.R. Tambo International Airport, South Africa. The arrows indicate the logical relationships that dictate how freight flows between facilities (Source: OpenStreetMap contributors (2017)).

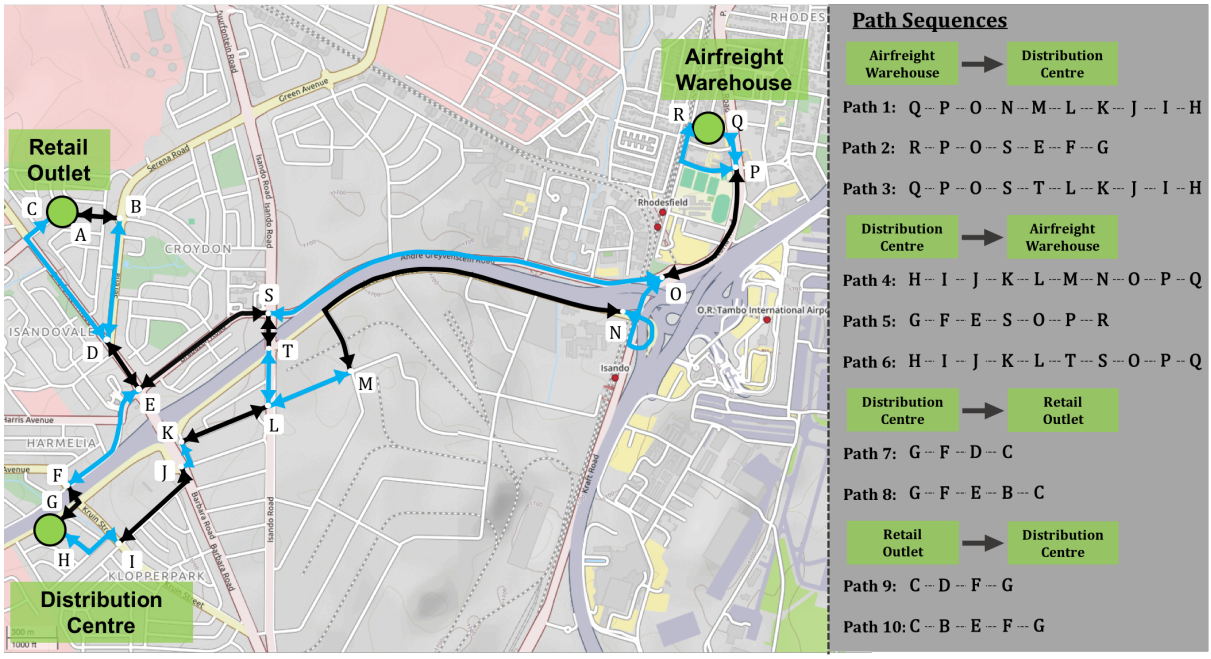


Figure 4.2: Shortest path alternatives of the illustrative supply chain network. Although more alternatives could exist in reality, only a few paths are assumed to keep the illustration simple. The difference in path lengths is also assumed to be within some allowable tolerance (Source: OpenStreetMap contributors (2017)).

Outlet and two paths (*Paths 9* and *10*) in the opposite direction. A further assumption is that the differences in these paths' lengths are within some allowable tolerance and

therefore they can be assumed to be of "equivalent" length.

In the remainder of this chapter we will refer back to this example when describing the design of the targeted attack strategies. When designing the three link-based targeted attack simulations, three questions had to be answered:

1. How was network damage to be defined and measured?
2. How were links to be prioritised for removal?
3. Would the prioritisation be revised after each disruption or not?

4.1.1 Defining network damage

In this thesis we tracked the **robustness** and **efficiency loss** of the network as indicators of damage. In keeping with contemporaries we regarded efficiency as the average shortest path length between connected node-pairs in G^{1K}. With regards to robustness we recognised that in practice facilities would be heterogenous and could not always be substituted (Zhao et al., 2011). Unfortunately, we lacked the data required to make assumptions regarding facility function in typical urban supply chains. To compensate for this we assumed that all facilities were important to the functioning of the supply chain and thus could not be substituted. Therefore the network was considered "connected" as long as *all* node-pairs in G^{1K} were still connected. Given these two dimensions, three escalating levels of network damage were defined.

Efficiency loss

Efficiency was lost when the average shortest path length increased. In the example, efficiency loss would occur when road links are removed so that all of the three paths between the Airfreight Warehouse and the Distribution Centre are destroyed. This would require that a new set of alternative shortest paths be found by looking for detours on the road network. The lengths of these new shortest path alternatives would be greater than that of the initial paths because of the detours. Therefore, the average shortest path length of the whole network would increase. Similarly, efficiency would be lost if the path sets in the opposite direction or between the Distribution Centre and Retail Outlet were destroyed as new alternative path sets would be longer.

Therefore, in the simulations, efficiency loss was regarded as an increase in the average shortest path length ($\bar{L}$).

Disconnection

A network was disconnected when one or more node-pairs in G^{1K} no longer had any path connecting them. Imagine that *Paths 1–3* were all destroyed and there were no viable detours that could be used to find new alternatives. It would then be impossible to travel from the Airfreight Warehouse to the Distribution Centre (even though it may still be possible to travel in the opposite direction). According to our definition of robustness, the network would be dysfunctional because one of the logical links are now disconnected.

In the simulations, a network was considered disconnected if one or more node-pairs in G^{1K} became disconnected. However, while still investigating the behaviour and characteristics of the collection of shortest paths we decided to continue the simulations beyond the initial disconnection.

Destruction

A network was considered destroyed when all node-pairs had become disconnected. In the example this would mean that it is not possible to travel between the Airfreight Warehouse and Distribution Centre in either direction nor is it possible to travel between the Distribution Centre and the Retail Outlet in either direction.

In the simulations destruction meant that G^{1K} had become an empty graph.

4.1.2 Prioritising links for removal

In reviewing the literature on link-based strategies, it was noted that betweenness and network skeletons were the most relevant concepts as these are defined by the shortest paths of a network. This aligned with our interest in the characteristics of the collection of shortest paths. In terms of network skeletons, the hub-and-spoke nature of the SH and Double Hub (DH) archetypes promoted the use of link salience (Grady et al., 2012). Therefore prioritisation relating to link betweenness and link salience was used.

The purpose of using the targeted attack simulations in this study was to compare the effectiveness of three prioritisation characteristics namely Overall link betweenness (Overall-B) (Section 4.1.3), Elemental link betweenness (Elemental-B) (Section 4.1.4), and Link salience (Overall-S) (Section 4.1.5). With limited computational resources and time either static or dynamic prioritisation could have been selected, but not both. A dynamic prioritisation approach, where Overall-B, Elemental-B and Overall-S were recalculated after each removal of links, held two specific benefits. Firstly, simulations were expected to be shorter as targeted attacks remained more effective. Secondly, the evolution of the characteristics could also be tracked by comparing the intermediate and final prioritisation to the initial prioritisation.

In summary, link-based targeted attack simulations with dynamic prioritisation were used to test three characteristics based on link betweenness and link salience. Each simulation started with a sample of 500 undisturbed multilayered instances for each network archetype. Iterative disruptions removed approximately 1% of links from G^2 according to the specified strategy. The performance of these simulations was compared based on the progression through three levels of escalating network damage namely efficiency loss, disconnection and destruction. The following three subsections describe the three prioritisation characteristics with reference to the illustrative example.

4.1.3 Overall link betweenness (Overall-B)

If we consider the set of alternative paths between the Airfreight Warehouse and the Distribution Centre, we notice that e_{OP}, which is the link from node O to P, features on all the paths, while e_{PO} features on all the paths in the opposite direction. Therefore, regardless of which path you choose, you will travel on e_{OP} in one direction or e_{PO} in the other. On the flip side, if either of these links are removed, it would immediately destroy all the alternatives in a particular direction. This is what it means when a link is "between" two nodes. It features heavily in the set of shortest paths that connects them.

Link betweenness is then an aggregated measure of a link's importance when considering all the shortest paths of the network. We call this Overall-B. To calculate Overall-B for e_{OP} (i.e. from the Airfreight Warehouse to the Distribution Centre), we start by counting how many times it features in the shortest paths of the entire network (see Table 4.1).

It then follows that:

$$\begin{aligned}\text{Overall-B}(e_{OP}) &= \frac{\text{Occurrences}}{\text{Total shortest paths}} \\ &= \frac{9}{22} \\ &= 0.41\end{aligned}$$

Table 4.1: Illustrative example of calculating Overall-B.

From	**To**	**Path sequence**	**Occurrences of** e_{OP}
Airfreight Warehouse	Distribution Centre	*Path 1*	–
		Path 2	–
		Path 3	–
Distribution Centre	Retail Outlet	*Path 7*	–
		Path 8	–
Airfreight Warehouse	Retail Outlet	*Path 1+Path 7*	–
		Path 1+Path 8	–
		Path 2+Path 7	–
		Path 2+Path 8	–
		Path 3+Path 7	–
		Path 3+Path 8	–
Distribution Centre	Airfreight Warehouse	*Path 4*	✓
		Path 5	✓
		Path 6	✓
Retail Outlet	Distribution Centre	*Path 9*	–
		Path 10	–
Retail Outlet	Airfreight Warehouse	*Path 4+Path 9*	✓
		Path 4+Path 10	✓
		Path 5+Path 9	✓
		Path 5+Path 10	✓
		Path 6+Path 9	✓
		Path 6+Path 10	✓
Total occurrences in all shortest paths:			**9**

To calculate Overall-B of each intralayer link in G^2 required that we first find a way to count on how many shortest paths each link featured. The collection of shortest paths, $\mathcal{C}(\mathcal{S}_{ij})$, was a collection of all of the path sequences for directly and indirectly connected node-pairs in G^{1K}. Each of these sequences consisted of a subset of the intralayer links of G^2. Therefore, if one reconstructed the grid using only the links in $\mathcal{C}(\mathcal{S}_{ij})$, the result was a partial grid of all the links that featured in the shortest paths of $\mathcal{M}$. We reconstructed such a grid using all the paths in $\mathcal{C}(\mathcal{S}_{ij})$ and called this $G^\zeta = (X^\zeta, E^\zeta)$ where $X^\zeta \subseteq X_2 \mid x_u^\zeta \in \mathcal{C}(\mathcal{S}_{ij})$ and $u \in \{1, 2, \ldots, N^\zeta\}$. The links were defined by $E^\zeta \subseteq E_2 \mid e_{uv}^\zeta \in \mathcal{C}(\mathcal{S}_{ij})$ and

$u, v \in \{1, 2, \ldots, N^{\zeta}\}$.

Unlike G^{1K} and G^2, the links of G^{ζ} were weighted. The weight of each link $e^{\zeta}_{uv} \in E^{\zeta}$ was the number of times that link featured in the path sequences of $\mathcal{C}(\mathcal{S}_{ij})$ and was denoted by w^{ζ}_{uv}. Therefore the link betweenness could be calculated from these weights[1] as follows:

$$\text{Overall-B}(e^2_{st}) = \begin{cases} \frac{w^{\zeta}_{uv}}{\sum\limits_{i,j;i\neq j} P_{ij}} & \text{if } e^2_{st} \equiv e^{\zeta}_{uv} \text{ and thus } e^2_{st} \in E^{\zeta}; \\ & \text{where } P_{ij} \text{ was the shortest path set size,} \\ & u, v \in \{1, 2, \ldots, N^{\zeta}\}, \\ & i, j \in \{1, 2, \ldots, N^{1K}\}, \text{ and} \\ & s, t \in \{1, 2, ..., N^2\} \\ 0 & \text{otherwise} \end{cases} \tag{4.1}$$

The Overall-B attack strategy prioritised e^2_{st} in descending order of Overall-B(e^2_{st}). The four top-ranking links were removed during a disruption ($\approx 1\%$ of E^2). After a disruption, $\mathcal{C}(\mathcal{S}_{ij})$ and Overall-B were recalculated before selecting the next four links to remove.

The average Overall-B score was calculated per instance and the distribution of these instance-averages is shown in Figure 4.3. The Fully Connected (FC) archetype had a relatively narrow distribution with a low mean (4.4%) indicating that across all 500 instances the grid links featuring in the shortest paths, e^{ζ}_{uv}, had very low betweenness scores as the shortest paths were widely spread across G^2. Contrast this to the SH and DH archetypes that forced shortest paths to route via the hubs. This resulted in fewer links in E^{ζ} that each carried a larger proportion of shortest paths. A Kolmogorov-Smirnov test (KS-test) rejected the null hypothesis that the hub network statistics were drawn from the same distribution and therefore it was settled that the DH archetype had a higher concentration of shortest paths which resulted in the higher Overall-B values. Intuitively this made sense as the grid links in the shortest path between the two hubs would have featured in all inter-hub shortest paths. A strategy that removed links based on Overall-B was thus expected to destroy the hub networks far quicker than the FC network.

Including **all** the shortest paths in the calculation is the classic way of calculating betweenness. However, in $\mathcal{M}$ the shortest path sets of indirectly connected node-pairs SI_{ij} were combinatorial products of the shortest path sets of the directly connected node-pairs SD_{ij} that constituted them (refer Figure 3.4). Therefore if all the shortest paths of any directly connected node-pair $(x^{1K}_i, x^{1K}_j) \in E^K_1$ were disconnected (i.e. $SD_{ij} = \emptyset$), all the indirectly connected node-pairs that contained that SD_{ij} were also disconnected. In the calculation of Overall-B there was thus a double counting of sorts. Therefore we developed another betweenness metric that focussed only on the path sets of directly connected node-pairs.

[1]For clarity's sake, it is emphasised that both the index sets s, t and u, v referred to nodes in the *physical* layer, the difference being that u, v referred to the subset of the nodes that appeared in $\mathcal{C}(\mathcal{S}_{ij})$. The distinction was necessary as an edge in G^{ζ} had a weight $w_{u,v}$ while the corresponding edge in G^2 was unweighted.

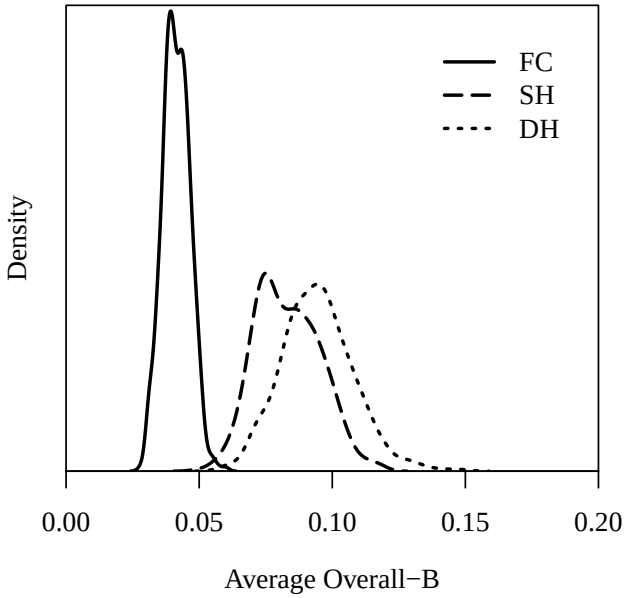


Figure 4.3: Distribution of the instance-specific averages of Overall-B values in the initial instances before disruption.

4.1.4 Elemental link betweenness (Elemental-B)

If we calculate Elemental-B for e_{OP}, we only consider the paths between the Airfreight Warehouse and the Distribution Centre as well as the Distribution Centre and the Retail Outlet. Table 4.1.4 shows the count of the number of occurrences and it follows that:

$$\begin{aligned} \text{Elemental-B}(e_{OP}) &= \frac{\text{Occurrences}}{\text{Total shortest paths (direct only)}} \\ &= \frac{3}{10} \\ &= 0.3 \end{aligned}$$

Table 4.2: Illustrative example of calculating Elemental-B.

From	**To**	**Path sequence**	**Occurrences of** e_{OP}
Airfreight Warehouse	Distribution Centre	*Path 1*	–
		Path 2	–
		Path 3	–
Distribution Centre	Retail Outlet	*Path 7*	–
		Path 8	–
Distribution Centre	Airfreight Warehouse	*Path 4*	✓
		Path 5	✓
		Path 6	✓
Retail Outlet	Distribution Centre	*Path 9*	–
		Path 10	–
Total occurrences in all shortest paths:			**3**

Once again, to calculate Elemental-B of any intralayer link in G^2, it was required to first count how many of the shortest paths between directly connected nodes in G^{1K} it featured in. We thus constructed a network of shortest paths $G^\gamma = (X^\gamma, E^\gamma)$ where $X^\gamma \subseteq X_2 \mid x_u^\gamma \in SD_{ij}$ and $u \in \{1, 2, \ldots, N^\gamma\}$. The links were defined by $E^\gamma \subseteq E^2 \mid e_{uv}^\gamma \in SD_{ij}$ and $u, v \in \{1, 2, \ldots, N^\gamma\}$.

Similar to G^ζ, the links of G^γ were weighted by the number of times a link featured in the path sequences of SD_{ij}. These link weights were denoted by w_{uv}^γ, where $u, v \in \{1, 2, .., N^\gamma\}$. Therefore we used w_{uv}^γ to calculate Elemental-B(e_{st}^2) as follows:

$$\text{Elemental-B}(e_{st}^2) = \begin{cases} \frac{w_{uv}^\gamma}{\sum\limits_{i,j;i\neq j} P_{ij}} & \text{if } e_{st}^2 \equiv e_{uv}^\gamma \text{ and thus } e_{st}^2 \in E^\gamma; \\ & u, v \in \{1, 2, \ldots, N^\gamma\}, \\ & i, j \in \{1, 2, \ldots, N^{1K}\}, \text{ and} \\ & s, t \in \{1, 2, ..., N^2\} \\ 0 & \text{otherwise} \end{cases} \tag{4.2}$$

The Elemental-B attack strategy prioritised e_{st}^2 in descending order of Elemental-B(e_{st}^2). The four top-ranking links were removed during a disruption ($\approx 1\%$ of E^2). After a disruption SD_{ij} and Elemental-B were recalculated before selecting the next four links to remove.

In the FC archetype all supply chain nodes were directly connected to all others and therefore there were no indirectly connected sets ($SI_{ij} = \emptyset \forall i, j$). The Overall-B and Elemental-B scores for the FC instances were thus identical.

In Figure 4.4 we see that the distributions of the average values of Elemental-B for the hub archetypes were very similar to that of the FC archetypes. (However, a KS-test still rejected the null hypothesis that these sample distributions were drawn from the same continuous distribution.) Suddenly importance (in terms of betweenness) was far more dispersed among the links of G^2. This showed the significant effect that the combinatorial

nature of SI_{ij} had. Based on the initial distributions it seemed likely that the SH archetype would have been most adversely impacted by a strategy based on Elemental-B. On average the links in E^{γ} had slightly higher betweenness in the SH instances and would thus affect more shortest paths if removed.

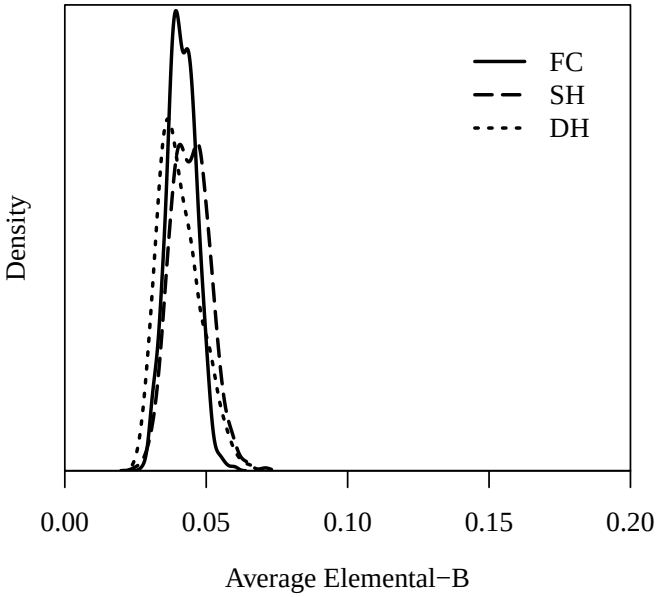


Figure 4.4: Distribution of the instance-specific averages of Elemental-B values in the initial instances before disruption.

4.1.5 Overall link salience (Overall-S)

Where link betweenness calculates the absolute dependence of a network's shortest paths on one link, link salience is a consensus-based metric that quantifies to what fraction of the shortest path sets a certain link is important. If a link features in one, more than one or all of the shortest paths between two nodes, it is important to that node-pair, if it does not feature on any of the shortest paths, it is not important to that node-pair. We denoted link salience by Overall-S.

To calculate Overall-S for e_{OP} we first determine whether it played a role in each of the shortest paths of each node-pair to determine a consensus score (see Table 4.3). It then follows that:

$$\begin{aligned}\text{Overall-S}(e_{OP}) &= \frac{\text{Consensus score}}{\text{Number of node-pairs}} \\ &= \frac{2}{6} \\ &= 0.33\end{aligned}$$

The most salient links tend to be spread out throughout the network while the most between links are concentrated in the barycenter (Grady et al., 2012; Viljoen and Joubert, 2016). One of the powerful advantages of salience as a prioritising metric is that most empirical networks display a bathtub distribution of salience scores which means that links can be clearly defined as salient or not (Grady et al., 2012). For our purposes, we

Table 4.3: Illustrative example of calculating Overall-S.

From	**To**	**Path sequence**	**Occurrence of e_{OP} in path set** [✓Yes/– No]
Airfreight Warehouse	Distribution Centre	*Path 1* *Path 2* *Path 3*	–
Distribution Centre	Retail Outlet	*Path 7* *Path 8*	–
Airfreight Warehouse	Retail Outlet	*Path 1+Path 7* *Path 1+Path 8* *Path 2+Path 7* *Path 2+Path 8* *Path 3+Path 7* *Path 3+Path 8*	–
Distribution Centre	Airfreight Warehouse	*Path 4* *Path 5* *Path 6*	✓
Retail Outlet	Distribution Centre	*Path 9* *Path 10*	–
Retail Outlet	Airfreight Warehouse	*Path 4+Path 9* *Path 4+Path 10* *Path 5+Path 9* *Path 5+Path 10* *Path 6+Path 9* *Path 6+Path 10*	✓
Total consensus score:			**2**

regarded salience as a better indicator of how widely spread the impact of a disruption would be than betweenness.

To calculate Overall-S for links in G^2 we first had to calculate the consensus scores for each link:

$$c_{st} = \sum_{i,j;i\neq j} c_{st}(i,j) | s,t \in \{1,2,...N^2\}, i,j \in \{1,2,...,N^{1K}\} \tag{4.3}$$

where

$$c_{st}(i,j) = \begin{cases} 1 & \text{if } e^2_{st} \in \mathcal{S}_{ij} \\ 0 & \text{otherwise} \end{cases} \tag{4.4}$$

In a network where all node-pairs are still connected, the number of shortest path sets $\|\mathcal{C}(\mathcal{S}_{ij})\|$ is equal to the number of node-pairs. However, we continued the targeted attack simulation beyond disconnection, therefore we divided the consensus score by the number

of shortest path sets and not the number of node-pairs:

$$\text{Overall-S}(e_{st}^2) = \frac{c_{st}}{\|\mathcal{C}(\mathcal{S}_{ij})\|} \text{ where } i, j \in \{1, 2, ..., N^{1K}\} \tag{4.5}$$

Overall-S prioritised e_{st}^2 in descending order of Overall-S(e_{st}^2). The four top-ranking links were removed during a disruption ($\approx$ 1% of E^2). After a disruption $\mathcal{C}(\mathcal{S}_{ij})$ and Overall-S were recalculated before selecting the next four links to remove.

We expected to find a clear distinction between salient and non-salient links and thus calculated two statistics per instance. The first was the average Overall-S score of all non-salient links of an instance (Overall-S(e_{st}^2) < 0.5). The second was the average Overall-S score of all salient links of an instance (Overall-S(e_{st}^2) ≥ 0.5). Figure 4.5 shows the distribution of these two statistics side-by-side.

The first observation was that there were no salient links for the FC archetype. That meant that no e_{st}^2 was included in more than half of the shortest path sets. In fact, links were quite conclusively non-salient with averages below 0.15. The lack of restrictive business rules to govern the *logical* paths in G^{1F} resulted in a somewhat random dispersion of shortest paths. There was no salience skeleton to extract.

While the hub archetypes showed a clear concentration of salient and non-salient links, the average for the salient links was far lower than observed in other studies which typically noted scores $\approxeq 1$. Similarly, the values for the non-salient links were far higher than in other studies which typically saw scores $\approxeq 0$. The skeleton structure in these instances was thus less distinct than in other real-life networks (Grady et al., 2012; Shekhtman et al., 2014; Viljoen and Joubert, 2016). These real-life networks typically had strong hub-spoke topologies (approximating scale-free networks). The FC archetype had no hub-and-spoke characteristics at all, explaining the difference. One might've expected the SH and DH archetypes to have a skeleton due to their explicit hub-and-spoke topologies. However, these hub-and-spoke archetypes were layered on a regular grid, which had a completely different skeleton structure. Therefore the multilayered nature diluted the skeleton structure. It was presumed that the lack of skeleton structure would weaken the effectiveness of Overall-S as prioritisation strategy.

4.2 Results of link-based targeted attack simulations

In analysing the results of the targeted simulations, we first determined the effectiveness of each of the disruption strategies. This was based on the three defined levels of damage. Secondly, we investigated how the metrics used to prioritise disruptions (Overall-B, Elemental-B and Overall-S) changed with progressive recalculation.

4.2.1 Effectiveness of simulation strategies

To compare effectiveness, we first considered the efficiency loss in the instances up until the point of disconnection. Thereafter, the number of disruptions required before instances were disconnected was tracked. The simulations were continued until all instances were completely destroyed and so the number of further disruptions required from disconnection to destruction was also analysed.

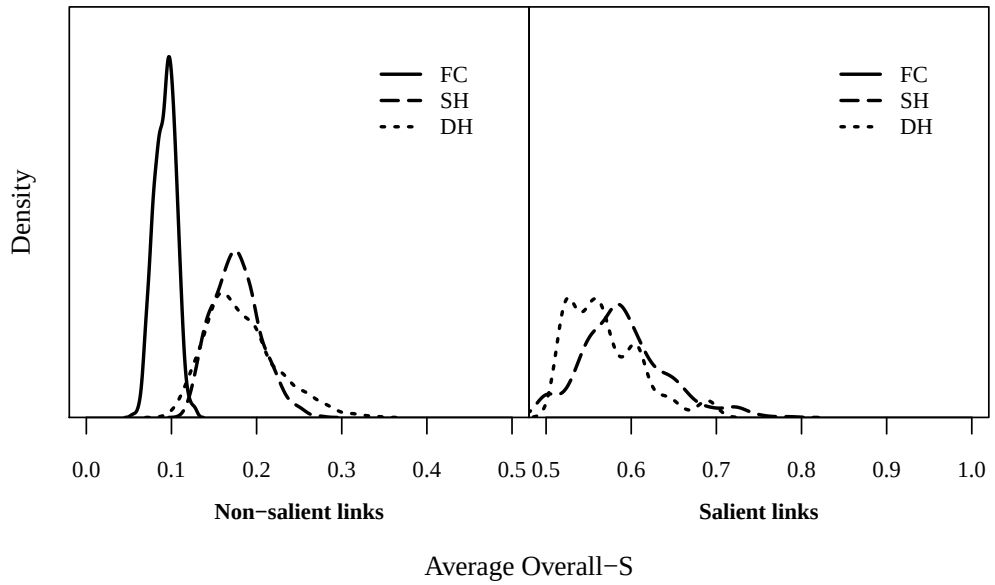


Figure 4.5: Distribution of the instance-specific averages of Overall-S values for salient ($0.5 \leq$ Overall-S$(e_{st}^2) \leq 1$) and non-salient ($0 \leq$ Overall-S$(e_{st}^2) < 0.5$) links in the initial instances before disruption. There were no links in the initial FC instances that had a salience value greater than 0.5.

Efficiency loss

Efficiency is lost in a network when it takes longer, on average, to travel from one node to another. In $\mathcal{M}$, a loss in efficiency was interpreted as an increase in the average shortest path $\bar{L}$. To measure the efficiency loss of an instance, the % difference between $\bar{L}$ in the initial undisturbed network and $\bar{L}$ in the network right before it became disconnected was determined.

The way in which efficiency and efficiency loss was measured in this thesis prevented one from tracking efficiency after disconnection for two specific reasons. Firstly, it could report "better" efficiency after a node-pair had become disconnected. If a node-pair was disconnected L_{ij} became 0. If this L_{ij} were still included in the calculation of $\bar{L}$ it would have lowered the average making the network "more" efficient, which would obviously have been untrue. Alternatively, if this node-pair were completely removed from the calculation of $\bar{L}$ it could still have resulted in a "higher" efficiency, depending on the changes in the remaining node-pairs. Secondly, if one were to disregard disconnected node-pairs in the calculation of $\bar{L}$ different instances would have had different numbers of node-pairs across which efficiency and efficiency loss were measured, making them incomparable. A number of researchers have encountered a similar problem in the measurement of efficiency in vulnerability studies and have suggested alternative measures to overcome it (Costa et al., 2007). For the purposes of this thesis the average shortest path remained the most intuitive. Thus, efficiency was not tracked beyond disconnection.

Instances were grouped into samples based on the number of progressive disruptions endured before disconnection. In other words, the % change in $\bar{L}$ of the 165 FC instances that were disconnected after 4% of the grid links were removed by the Overall-B simulation all constituted independent observations within one sample. These observations were all made when 3% of the grid links had been removed as the next disruption disconnected the instances. Figure 4.6 plots the observations in the samples corresponding to the

point in the simulation at which they were made. The figure distinguishes between the three simulations for each of the network archetypes[2]. In addition to the observations of each sample, the sample means are also shown for all samples that had more than one observation.

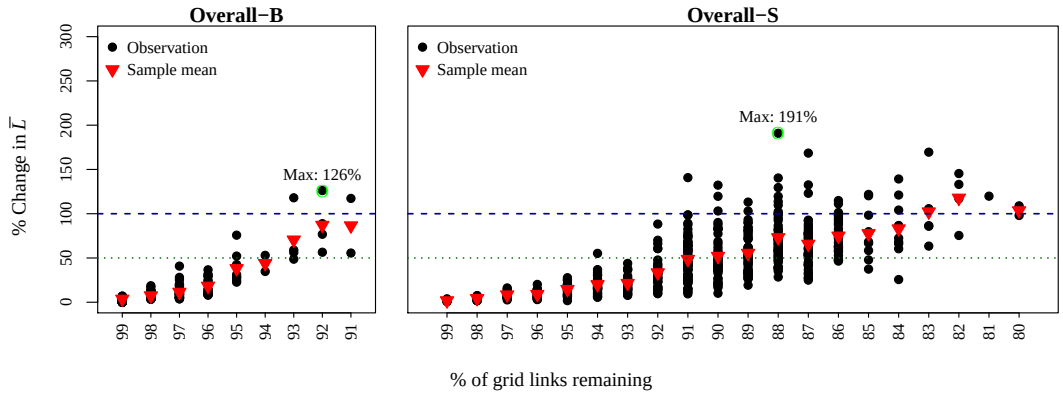
Across the board, the longer an instance survived, the greater its efficiency loss was. This was evident from the upward trend in the sample means of the % change in $\bar{L}$. With regards to the effectiveness of the prioritisation strategy, this could have been interpreted in two ways. Firstly, a strategy could have been considered *more* effective because it resulted in greater efficiency loss, which was the first level of damage. However, instances only incurred greater efficiency loss because they did not become disconnected, which was the second level of damage. The second interpretation could have been that a strategy was actually less effective in causing damage if it caused greater efficiency loss simply because that meant that it could not disconnect the network.

There were also two practical interpretations to consider. The first was that a supply chain would rather endure efficiency loss than disconnection and therefore a strategy that disconnects sooner is more lethal. The second perspective was that a strategy that doesn't disconnect facilities would not necessarily raise a red flag with management. Unbeknownst it would quietly whittle away efficiency until the cumulative damage turns out to be a silent killer for the supply chain. Cognisant of both perspectives, we chose to regard a strategy that disconnected sooner (i.e. lower efficiency loss) to be more effective.

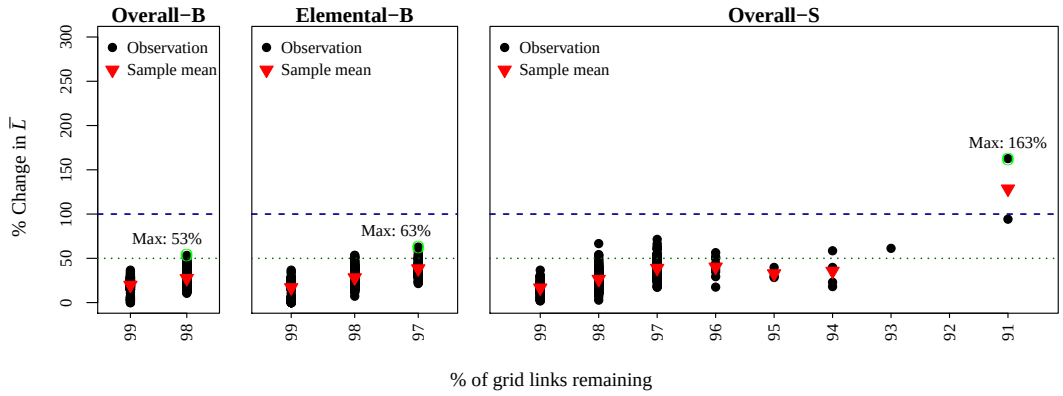
The FC archetype sustained far greater efficiency loss before disconnection with $\bar{L}$ increasing by more than half in 2.4% and 41.8% of the instances under the Overall-B and Overall-S strategies, respectively (Figure 4.6a). In fact, the Overall-S strategy doubled $\bar{L}$ in 6.6% of the instances. The first reason for this was that FC instances survived longer and thus endured more disruptions. The second reason was that shortest paths could explore the entire grid and were not constrained by hubs, allowing very inefficient, round-about paths to keep two nodes connected. The hub archetypes sustained, overall, lower efficiency losses than the FC archetype (see Figures 4.6b and 4.6c). This was because they became disconnected quicker and the constraint of the hubs limited the opportunity for greatly inefficient shortest paths to develop.

The Overall-B strategy resulted in more rapid efficiency losses for both the FC and DH archetypes before instances became disconnected (Figures 4.6a and 4.6c, respectively). However, efficiency losses in both these archetypes reached a peak under the Overall-S strategy as more disruptions were endured before disconnection. Interestingly, differences in efficiency losses were not notifiable across the three strategies for the SH archetype (Figure 4.6b), presumably because instances became disconnected so quickly that there was no opportunity for differentiation.

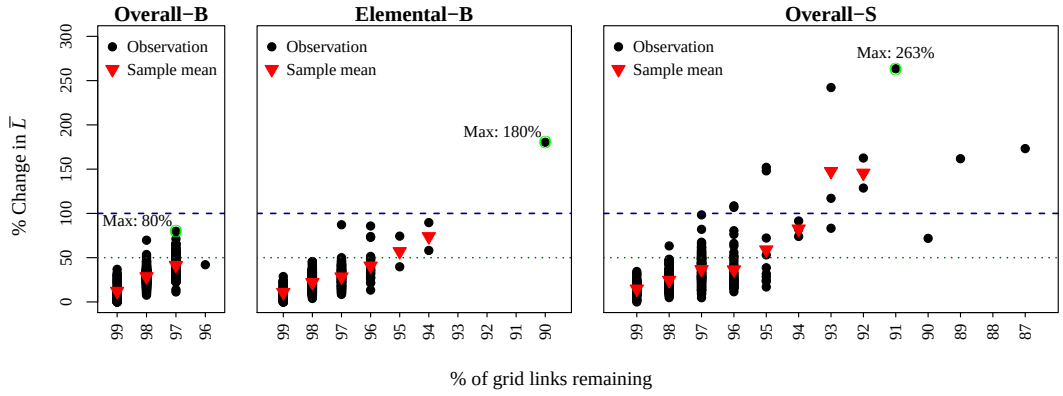
[2]For instances that became disconnected immediately during the first disruption, no % change in $\bar{L}$ could be measured.



(a) FC archetype



(b) SH archetype



(c) DH archetype

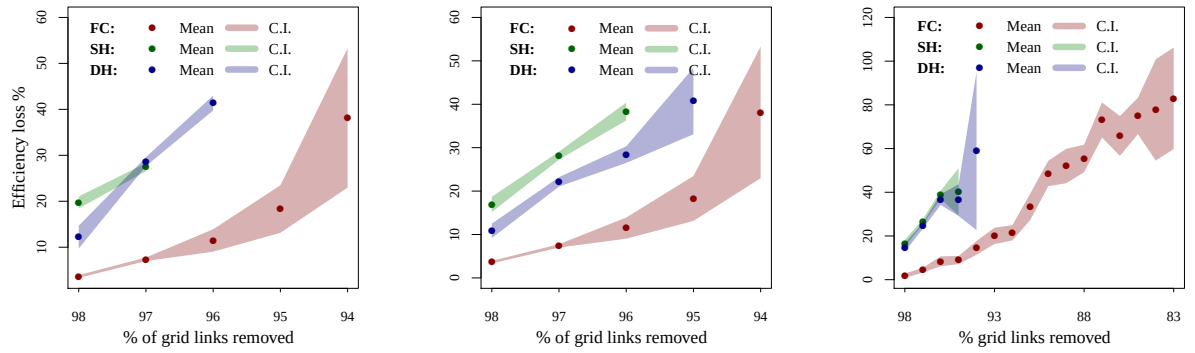
Figure 4.6: Samples of efficiency loss observations corresponding to the % of grid links remaining right before instances became disconnected.

Sample sizes varied quite dramatically in the investigation of efficiency loss. Smaller sample sizes affected the confidence intervals of the sample means. The insights and generalisations made with reference to the sample means therefore had to be tempered by an appreciation of these confidence intervals. With sample sizes varying from 1 to 353 the Student's t-distribution seemed most appropriate for calculating the confidence intervals. However, the t-distribution required an assumption regarding the normality of

the distribution of the % Change in $\bar{L}$. The distributions of the initial values of $\bar{L}$ discussed in Section 3.5 did not show evidence of heavy tails. This was because the length of shortest paths was not prone to combinatorial explosion, unlike the set sizes. The % Change was calculated from these 'contained' distributions and therefore it was assumed that the resulting distributions would also have central tendencies. The confidence intervals were calculated with $\alpha = 95\%$ for all samples that had five or more observations.

The confidence intervals were verified by a bootstrap analysis that performed 3 000 replications for each sample using the same parameters as the t-distribution calculations. In comparing these intervals with those obtained by the t-distribution, it was found that the upper bounds differed by a maximum of 7.6% while the lower bounds differed by less than 5%. The exceptions were the FC archetype and the Overall-S simulation on the DH archetype where the maximum difference was between 20% and 36%. This level of consensus between the two techniques was regarded as satisfactory.

Figures 4.7 plots the confidence intervals along with the sample means to compare efficiency losses for each strategy. It was clear that the impact of Overall-B was more pronounced for the hub archetypes. In the case of the Overall-S strategy, greater efficiency losses were initially observed for the hub archetypes. However, the efficiency losses for the FC archetype were greater after many more disruptions were endured. Similarly, the efficiency losses caused by the Elemental-B strategy were greater for the SH instances initially but eventually DH instances suffered greater damage.



(a) Overall-B (b) Elemental-B (c) Overall-S

Figure 4.7: Sample means and confidence intervals for efficiency losses inflicted by each targeted attack strategy across the archetypes.

In summary, the Overall-B strategy was the quickest to cause efficiency loss across the archetypes. However, because the Overall-S strategy took longer to disconnect the instances, it resulted in greater efficiency loss in those instances that survived longest. Similarly, the efficiency losses sustained by the FC archetype eventually outstripped that of the hub archetypes.

Next we investigated the efficacy of the three strategies in disconnecting and destroying the network instances.

Disconnection and Destruction

The cumulative plots in Figures 4.8–4.10 show what percentage of the grid links in G^2 had to be removed through progressive disruption before instances became disconnected and later destroyed for each of the three simulation strategies. The plots show that Overall-B

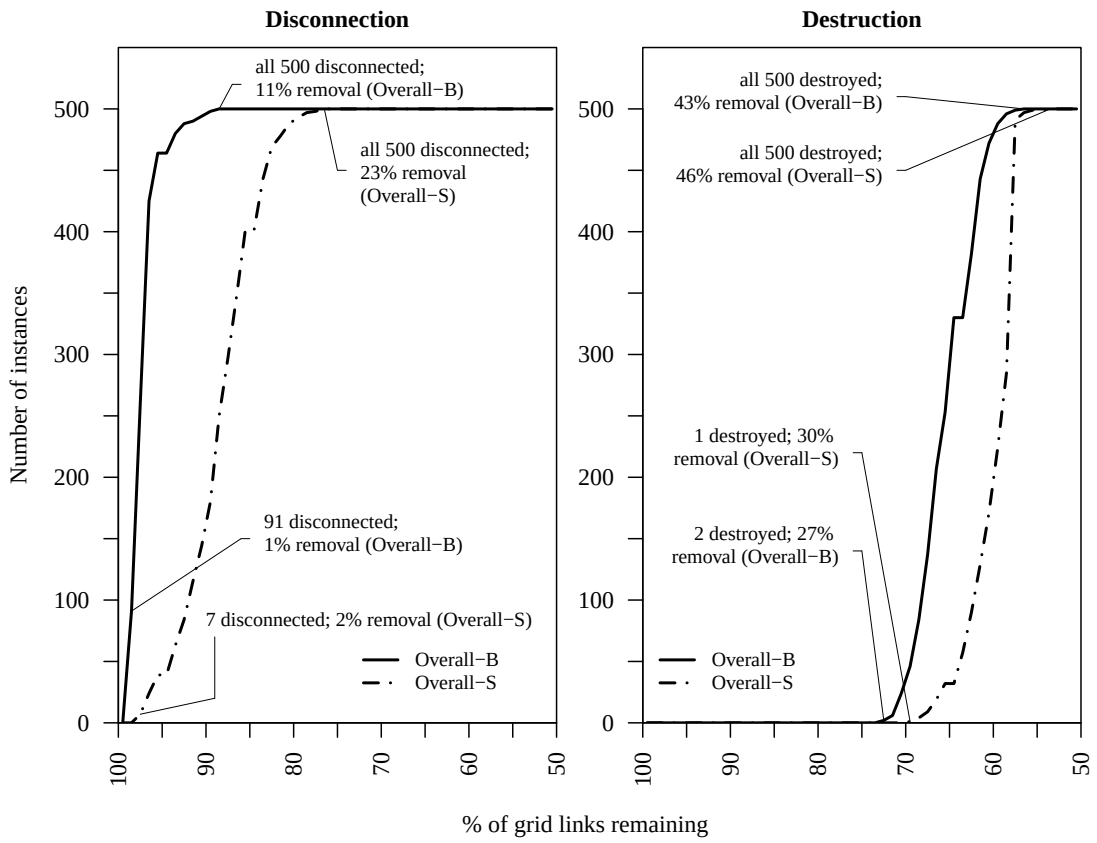


Figure 4.8: Comparison of the % of grid links removed by each simulation strategy before instances became disconnected or destroyed in the FC archetype.

was the most effective disruption strategy for all three archetypes. Its curves on the cumulative plots were consistently higher than that of the Elemental-B and Overall-S strategies. This meant that at any point in the simulation (i.t.o. grid links removed) the Overall-B strategy had broken more instances than either of the other strategies. Overall-S, on the other hand, was the least effective across all three archetypes as illustrated by its cumulative curve that is consistently lowest of the other strategies. Elemental-B was only applied to the hub networks as Overall-B $\equiv$ Elemental-B for the FC archetype (refer to Section 4.1.4). The performance of Elemental-B was better than that of Overall-S and in some cases closely mimicked that of Overall-B.

The instances of the FC archetype deteriorated very differently compared to those of the hub archetypes (Figure 4.8). An impressive 91 instances were disconnected by the Overall-B strategy after the very first disruption but only after a further 26 disruptions was the first instance completely destroyed. The Overall-S strategy was also quick to disconnect the first 7 instances but it took another 23 disruptions to completely destroy an instance. In the case of the FC archetype, instances' shortest paths need not have been routed through hubs. Once the initial shortest paths were destroyed the full extent of G^2 could be explored to find alternative routes. Therefore it is intuitive that these instances survived longest.

Contrastingly, in both hub archetypes instances were completely destroyed soon after they were disconnected. This was most evident in the SH archetype where the Overall-B and Elemental-B strategies almost simultaneously disconnected and destroyed instances. Each hub node had only four incoming and four outgoing links which, probabilistically,

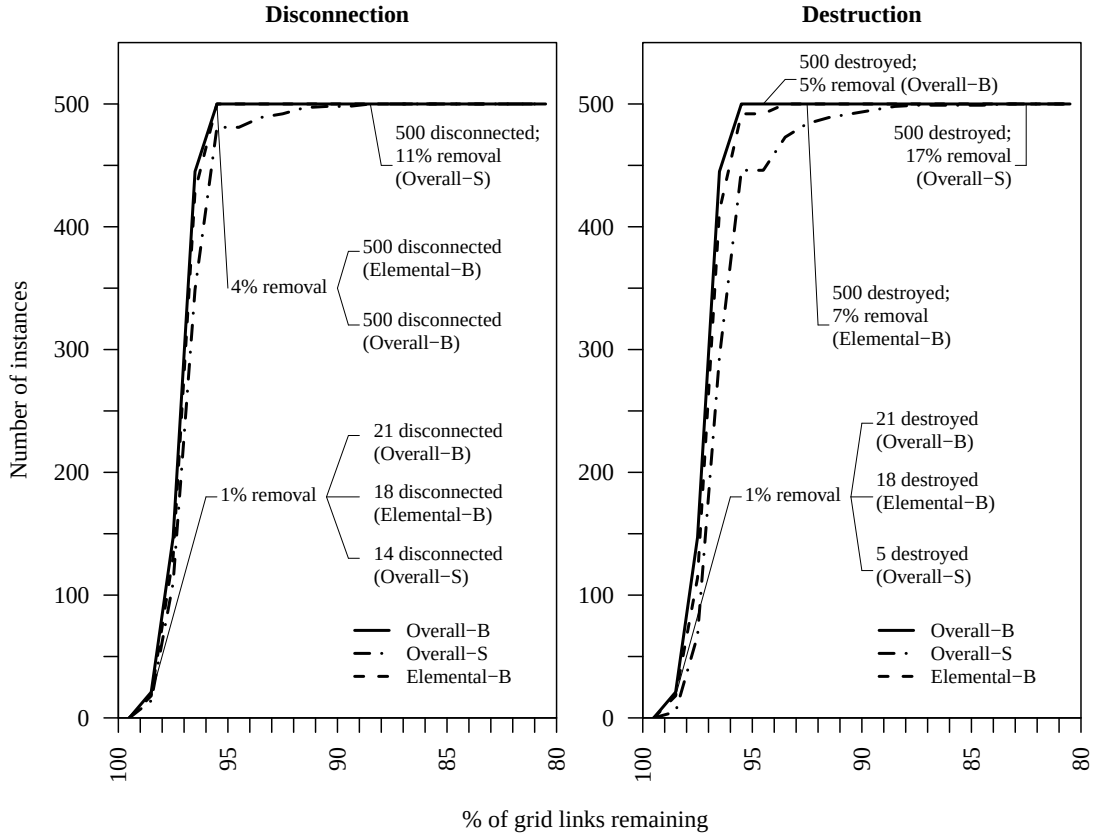


Figure 4.9: Comparison of the % of grid links removed by each simulation strategy before instances became disconnected or destroyed in the SH archetype.

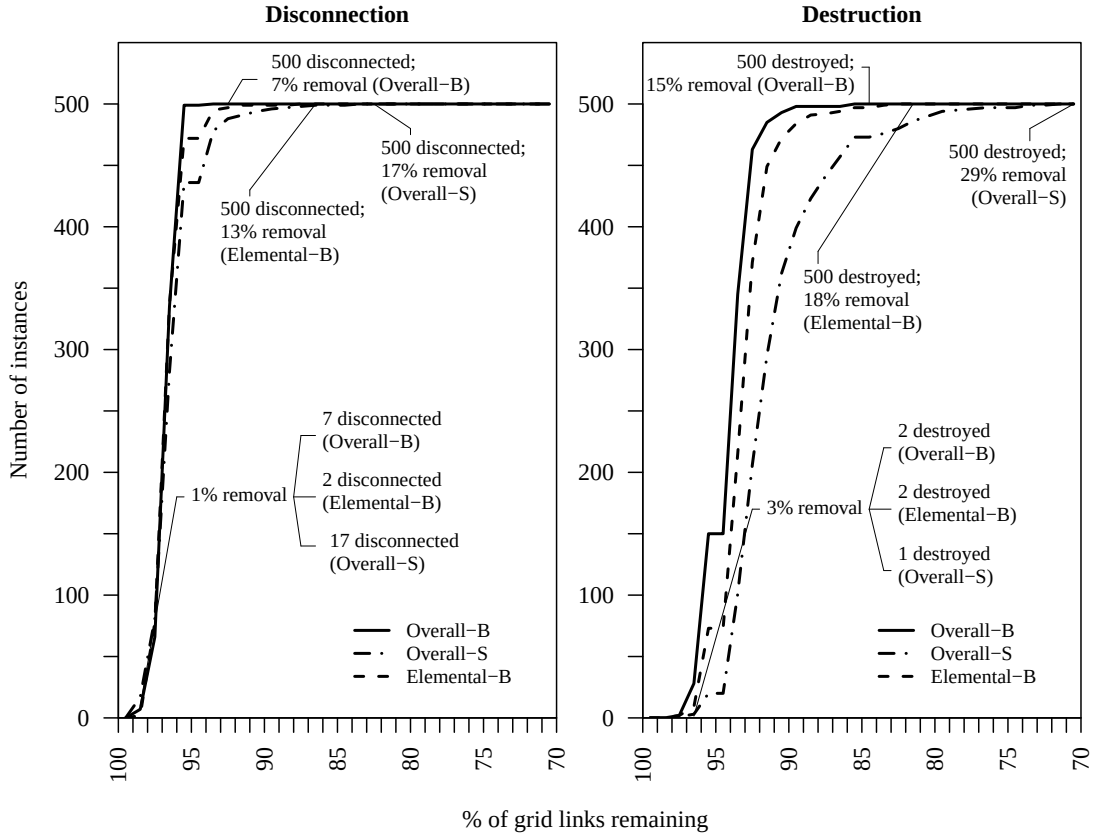


Figure 4.10: Comparison of the % of grid links removed by each simulation strategy before instances became disconnected or destroyed in the DH archetype.

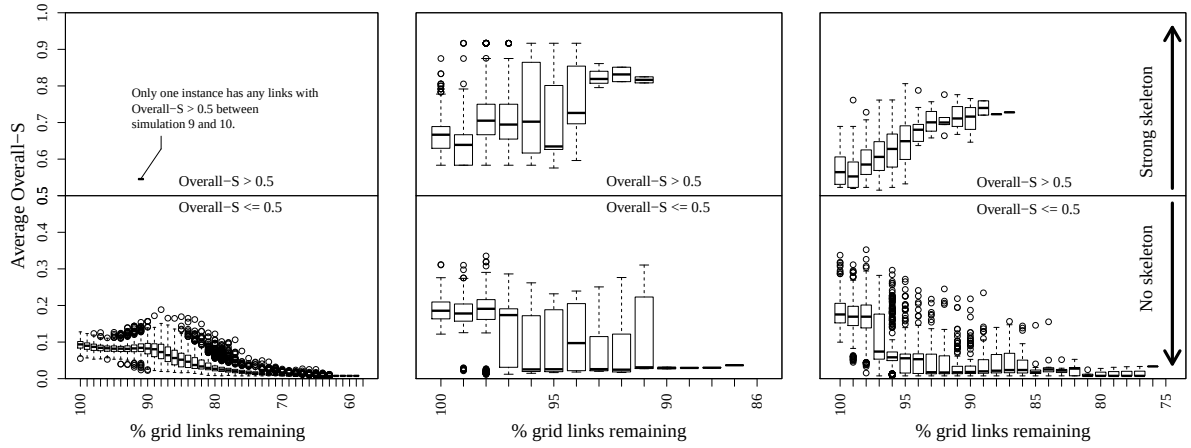
all had high Overall-B, Elemental-B and Overall-S scores. Therefore the links around the nodes became targets very soon, disconnecting the instances. Furthermore, the logical routing of G^{1K} limited the options for alternative shortest paths, which led to quicker destruction.

Reflecting on the initial distributions of Overall-B, Elemental-B and Overall-S the dominance of the betweenness-based strategies was unsurprising. The next topic to investigate was how these prioritisation metrics changed as the networks were degraded.

4.2.2 Evolution of the prioritisation metrics

A dynamic prioritisation approach was chosen specifically so that the evolution of the prioritisation metrics (Overall-B, Overall-S and Elemental-B) could be tracked. Observing the change in these metrics gave further insight to the change in the shortest path sets as G^2 lost connectivity[3].

The Overall-S strategy performed poorly in disconnecting the networks, especially the FC archetype. In Figure 4.5 it was already identified that the initial instances showed weak skeleton structure for the hub archetypes and zero skeleton structure for the FC archetype. Figure 4.11 tracks the change in average Overall-S for salient and non-salient links. In the FC archetype even the little consensus that existed quickly eroded (Figure 4.11a). This meant that after the first few disruptions the Overall-S strategy had been reduced to a quasi-random strategy. For the hub archetypes the scores of the non-salient links decreased while those of the salient links increased (Figure 4.11b and 4.11c). This suggested that the weak skeleton structure actually became more prominent in the hub archetypes as the destruction of alternatives forced more and more node-pairs to use the same remaining grid links. However, the skeletons were still not prominent enough to make it an effective strategy.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 4.11: Change in the average Overall-S for salient and non-salient links as disruptions progressed.

In both the betweenness-based strategies the effectiveness depended on whether there

[3]Two caveats to the observations in this section were that the sample sizes of surviving instances decreased rapidly in the last few disruptions and that the number of remaining shortest path sets also decreased as an instance went from being disconnected to being destroyed.

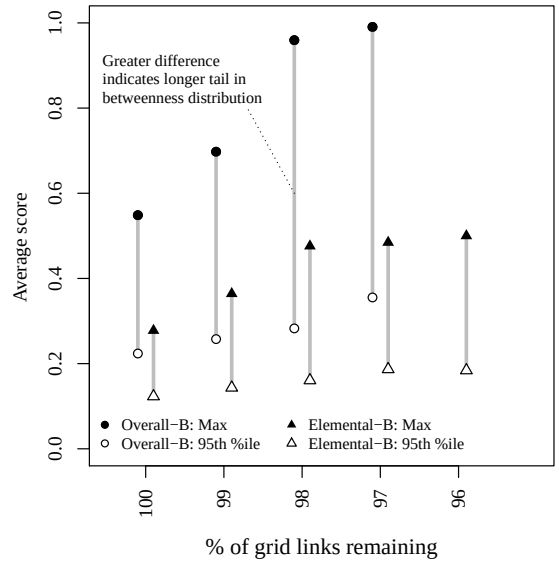
were links that were far more important than others based on the frequency of their appearance in shortest paths. Thus, distributions of Overall-B and Elemental-B with heavy right-tails would have shown that a specific instance was extremely vulnerable to the removal of its highest ranking links.

In Figures 4.3 and 4.4 the initial distributions of these two metrics were not heavily skewed. Zooming in on the right tail of these distributions we investigated the range between the 95^{th} percentile and the maximum score value. The broader this range, the longer the right tail, meaning that the highest ranking links were *far more* between than the others. Figure 4.12 displays these ranges for both Overall-B and Elemental-B in the case of the hub archetypes and Overall-B in the case of the FC archetype. After each progressive disruption the average of the 95^{th} percentile cut-offs and average of the maximum values were determined across all instances that had not yet been disconnected. Results are only shown until the number of surviving instances was smaller than 1% of the original set of 500 instances.

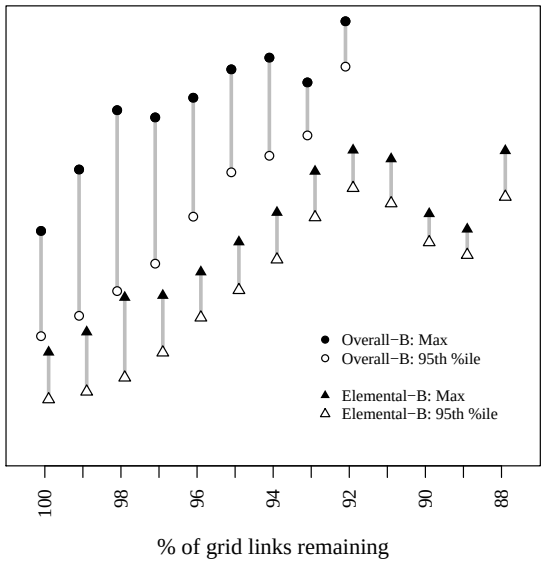
In the case of the SH archetype (Figure 4.12a) we noticed that the 95^{th} percentile did not shift that far upward. That implied that the bulk of the betweenness values didn't increase much. However, for both Overall-B and Elemental-B the distance between the maximum values and the rest of the distribution increased. This implied that the most between links became even more crucial to the survival of the instances as disruptions progressed, justifying the decisive effectiveness of the betweenness strategies on the SH archetype.

In the DH archetype, the range increased slightly in the first few iterations but then actually reduced again until the final ranges were even narrower than the initial ones (Figure 4.12b). This implied that while initially there had been a concentration of shortest paths — making some links crucially important — the importance of links later became more evenly spread. Do note, however, that the 95^{th} percentile increased by quite a margin indicating an upward shift in the betweenness of many links. As disruptions removed more and more routing options, all remaining links in G^2 played more significant roles.

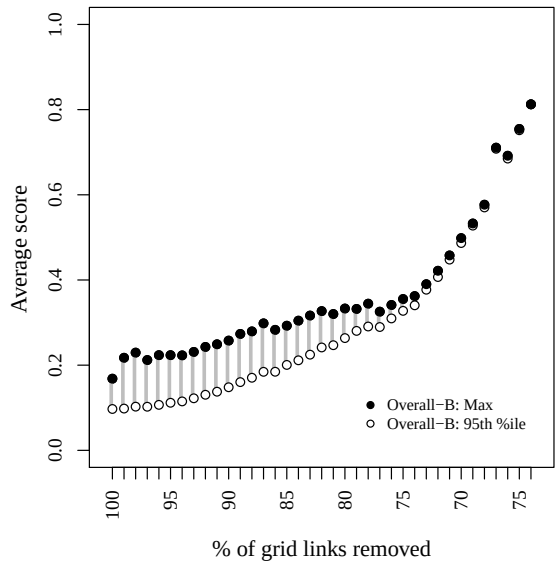
The FC archetype shows a similar pattern for Overall-B in that the range first increased and then decreased while the 95^{th} percentile shifted upward.



(a) SH archetype



(b) DH archetype



(c) FC archetype

Figure 4.12: Change in the right tail of the Overall-B and Elemental-B distributions as disruptions progressed.

Initially the betweenness-based metrics were better at prioritising crucial links than the salience-based metric. As disruptions progressed, the discriminatory ability of the betweenness metrics actually increased before eventually tapering off while the limited power of the salience metric just eroded further. The performance of the three metrics constituted two elements: the ability to reduce efficiency and disconnect networks and their sustained prioritising power. This was taken into consideration in defining the characteristics that could identify a certain network $\mathcal{M}$ as more vulnerable than another.

4.2.3 Characteristics that made $\mathcal{M}$ vulnerable

The second objective of this thesis was:

> 2. Identification of the characteristics of $\mathcal{M}$ that describes the nature of the supply chain's vulnerability to the integrity of the urban road network.

Three targeted attack simulations were designed based on characteristics of $\mathcal{C}(\mathcal{S}_{ij})$. The Overall-S strategy tested the importance of links that were salient — in other words links that featured in many different shortest path sets. Firstly, it was found that in the multilayered formulation of $\mathcal{M}$ there were no clearly salient links in G^2 as observed in other systems (Grady et al., 2012; Shekhtman et al., 2014; Viljoen and Joubert, 2016). The networks in these other studies were not multilayered. The regular grid structure of G^2 and its effect on constraining shortest paths influenced the salience of links. As was expected, the salience strategy was least effective in disconnecting the networks. We acknowledged that a redefined formulation of link salience better suited to the multilayered context could potentially have performed better. This is regarded as worthwhile future work. Our conclusion was that link salience as defined in this thesis was not a significant indicator of network vulnerability.

On the other hand, the two metrics that measured the frequency with which links appeared in $\mathcal{C}(\mathcal{S}_{ij})$ were far better in identifying the links on which the connectivity of $\mathcal{M}$ hinged. Higher betweenness scores implied that there were few alternative shortest paths, or if there were alternatives, that there was a lack of diversity (i.e. the alternatives included nearly all of the same links on G^2). Overall-B also consistently outperformed Elemental-B, therefore it was acknowledged that taking the full ambit of shortest paths into account was more telling than just focussing on the shortest paths between directly connected node-pairs.

Stepping away momentarily from micro-level statistics that pertain to individual links, we focussed on $\bar{L}$. When $\bar{L}$ changed, it implied that a shortest path set had been emptied and replaced by another with longer shortest path length. As noted in the case of the FC archetype, this could happen many times before an instance is actually in danger of disconnection. So the % change in $\bar{L}$ was not necessarily an indicator of the increasing vulnerability of $\mathcal{M}$. However, when the difference in L_{ij} for a specific node-pair before and after a disruption was relatively large, it implied that G^2 was becoming increasingly sparse and it was becoming difficult to find alternative paths. Therefore, the % change of L_{ij} between disruptions was also deemed a characteristic worth further investigation.

With these characteristics in mind, the following two chapters detail the definition and testing of a suite of vulnerability metrics that were proposed as indicators of the inherent vulnerability of $\mathcal{M}$.

Chapter 5

Development and analysis of vulnerability metrics

The way in which an urban road network is disturbed can be described as a combination of *random errors* and *targeted attacks*. Non-random, targeted link-disruption strategies were used to identify the characteristics of the shortest path sets that could be most indicative of the vulnerability of $\mathcal{M}$. The third objective of this thesis was then:

> 3. Development of metrics that could quantify a supply chain's inherent vulnerability based on its internal configuration and the underlying road network.

The internal configuration referred to the logical design of G^{1K}, where $K \in \{F, S, D\}$, while G^2 was the underlying road network.

This chapter presents the suite of vulnerability metrics developed and their empirical validation. Given that these metrics were deduced from the outcomes of targeted strategies, it was expected that they would be good quantifiers and predictors of vulnerability under similar targeted strategies. Confirming this intuition would not have validated the metrics. Instead the power of these metrics had to be tested using a completely randomised link-based disruption strategy. If a metric was a good quantifier and predictor of vulnerability under completely random link disruption, it was expected that its power would have been even greater in circumstances where disruptions had both random and targeted elements. Testing the metrics using random link-based disruptions was thus the most conservative method of evaluation.

5.1 Link-based random error simulation

The link-based random error simulation started again with the set of undisturbed network instances described in Chapter 3. Each progressive disruption randomly removed 18 ($\approx 5\%$) of the links from G^2. A greater percentage of links was removed per iteration in the random error simulation compared to the targeted simulations. In the targeted simulations it was guaranteed that each link removed played a role in $\mathcal{C}(\mathcal{S}_{ij})$. This was not the case with the random error simulation where, depending on the spread of G^{1K} on G^2, the likelihood of removing links from G^2 that played no role in $\mathcal{C}(\mathcal{S}_{ij})$ was considerable. The random error simulation degraded $\mathcal{C}(\mathcal{S}_{ij})$ far slower than the targeted simulations.

The simulation was continued only until an instance was disconnected and not until it was completely destroyed (as with the targeted simulations). The assumption was that in a practice, corrective action would be taken once two facilities become disconnected.

Therefore, the primary interest was to test the ability of the metrics to quantify the likelihood of efficiency loss and disconnection. The performance of the metrics between disconnection and destruction was not monitored.

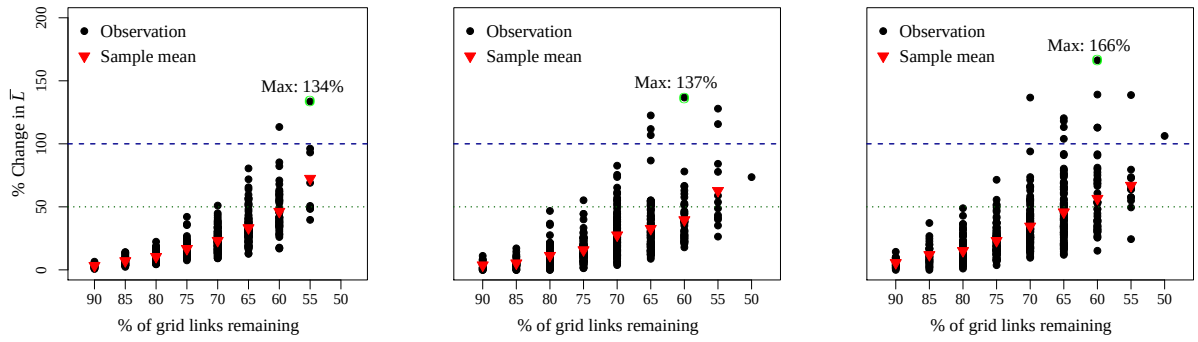
After each progressive disruption $\mathcal{C}(\mathcal{S}_{ij})$ was recalculated and the vulnerability metrics (presented hereafter) were measured. At the end of the simulation each instance had a time-series of values for each vulnerability metric throughout the progressive disruptions.

5.2 Results of the link-based random error simulation

First we investigated the effectiveness of the random error strategy in terms of the efficiency loss caused before disconnection. Thereafter the rate at which instances became disconnected was analysed.

5.2.1 Efficiency loss before disconnection

The efficiency loss per instance was once again described by the percentage change in the average shortest path and was denoted by % change in $\bar{L}$ (refer to Section 4.1.1). Similar to the investigation of efficiency loss in Section 4.2.1, instances were categorised into samples according to the number of progressive disruptions endured before disconnection. The % change in $\bar{L}$ reflected the difference between the initial, undisturbed network and the last state of the network before disconnection. Figure 5.1 plots the observations per sample as well as the sample means for efficiency loss across the three archetypes. The average efficiency loss increased monotonically the longer instances survived. This was similar to the trend observed in Figure 4.6 and occurred for the same reason: the more disruptions an instance endured before becoming disconnected, the sparser G^2 became resulting in more of a roundabout route connecting a node-pair.



(a) FC archetype (b) SH archetype (c) DH archetype

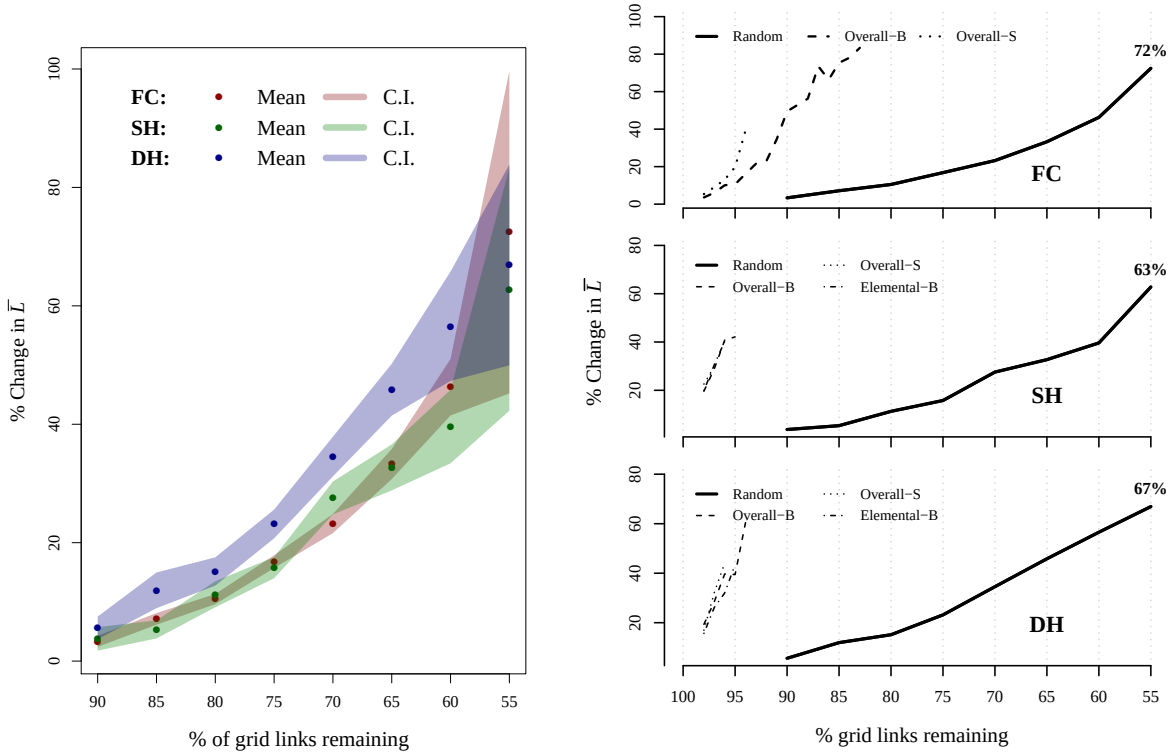
Figure 5.1: Samples of efficiency loss observations corresponding to the % of grid links remaining right before instances became disconnected under the random error strategy.

Once again the sample sizes varied greatly and it was decided to calculate the confidence intervals for the sample means using the Student's t-distribution. For samples with five or more observations the confidence intervals were calculated with $\alpha = 95\%$. These intervals were again compared to intervals calculated by a bootstrap analysis that performed 3 000 replications at $\alpha = 95\%$ for each sample. It was found that the upper

bound of the intervals had a maximum deviation less than 5.3% across all the networks. Meanwhile, the lower bound had maximum deviations of 10.9%, 18.3% and 22.5% for the FC, SH and DH archetypes, respectively. This level of consensus between the two techniques was regarded as satisfactory.

The sample means and confidence intervals were remarkably similar across the three archetypes under the random error simulation (Figure 5.2a). Despite the similarity, it could still be noticed that the DH archetype lost its efficiency more rapidly than the other archetypes. However, in the end it was the FC archetype that had degraded the most overall while the SH archetype had degraded the least. This relative similarity was in contrast to the efficiency loss under the targeted attack simulations as shown in Figure 4.7.

Figure 5.2b plots the efficiency loss trajectories of the targeted simulations along with that of the random error simulation for each archetype. Notice that for all three archetypes efficiency loss was far more rapid (i.e. a steeper slope) under the targeted simulations than under the random simulation. The only reason the overall efficiency loss was higher under the random simulation was because the instances survived so much longer. The targeted simulations thus caused far greater damage than the random error simulation from an efficiency loss point of view. Next the rate at which the instances became disconnected under the random error simulation was investigated.



(a) Comparison of efficiency loss across the three archetypes under the random error simulation.

(b) Comparison of efficiency loss per archetype under different simulations.

Figure 5.2: Comparison of efficiency loss across networks and under different simulations.

5.2.2 Disconnection of networks

The cumulative percentage of instances that had become disconnected after each disruption is shown in Figure 5.3. The distribution was very similar across the three archetypes with the FC archetype being disconnected faster in the mid-range of the disruptions. All instances of the FC archetype were disconnected after 50% of the links were removed with the final surviving instances of the hub archetypes disconnected after 55% link removal.

Both in terms of efficiency loss and disconnection, it was notable that the effectiveness of the random error was indifferent to the network archetype. This implied that it was possible that a supply chain's vulnerability to random road network disruptions were more dependent on the external circumstances, rather than the internal configuration. This was in stark contrast to the targeted simulation results in Chapter 4 where the archetypes showed different levels of resilience.

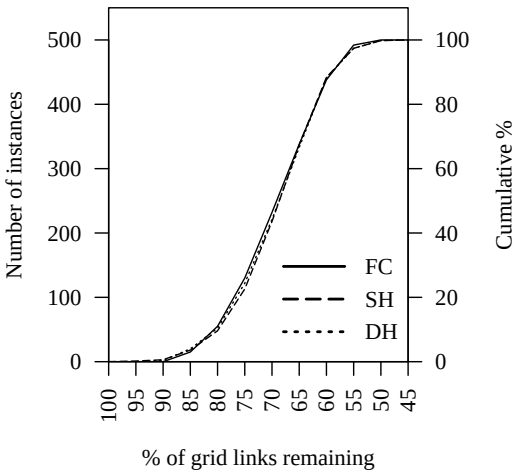


Figure 5.3: Cumulative percentage of instances that became disconnected with each progressive disruption.

The remainder of this chapter presents the vulnerability metrics that were developed based on the results of the targeted attack simulations. These metrics were grouped under three categories: Redundancy, Overlap and Efficiency step-change. First each category is described conceptually. To facilitate this description we refer back to the illustrative example used in Chapter 4. For convenience we repeat the image of the shortest path sets (Figure 5.4) below. Thereafter, the mathematical formulation of the metrics and the different measurements applied to each are presented. Finally, the ability of each metric to quantify vulnerability is assessed by means of statistically testing the measurements obtained from the random error simulation.

5.3 Vulnerability category 1: Redundancy

5.3.1 Conceptual description

In the illustrative example, there are three alternative paths when travelling from the Airfreight Warehouse to the Distribution Centre namely Paths 1–3. It is recognised that

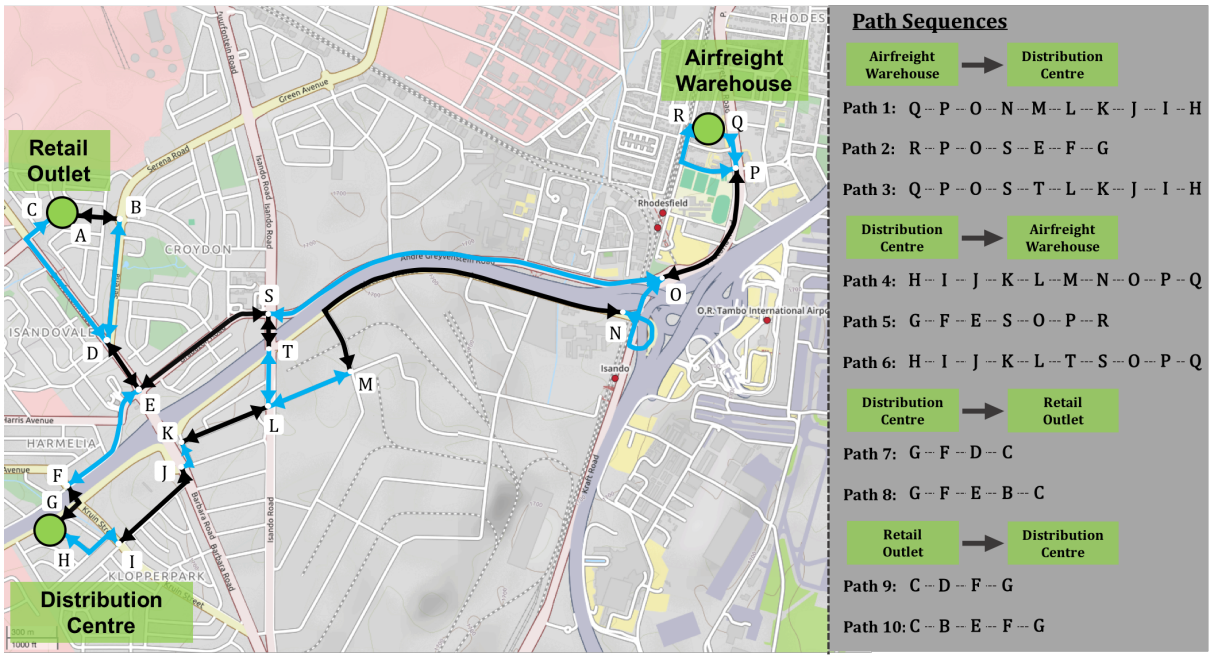


Figure 5.4: Shortest path alternatives of the illustrative supply chain network. Although more alternatives could exist in reality, only a few paths are assumed to keep the illustration simple. The difference in path lengths is also assumed to be within some allowable tolerance (Source: OpenStreetMap contributors (2017)).

the actual difference in length of these paths is not negligible, but for the purposes of this example we regarded the lengths as relatively equivalent. Suppose that there were never any disruptions of the road network — no traffic jams or road closures. Then having more than one alternative would be unnecessary, or redundant. However, disruptions are a frequent reality and therefore having more alternative options available to the truck driver or fleet manager is always better.

The availability of more alternative paths defines the concept of redundancy. The rationale was that a network instance with more redundancy in its shortest path sets could be more immune to efficiency loss and disconnection.

5.3.2 Formulation of metrics

The size of the shortest path set between any two nodes x_i^{1K} and x_j^{1k} in G^{1K} was captured by P_{ij}. Therefore the distribution of P_{ij} was a reflection of the redundancy of a network instance. The distribution of P_{ij} was greatly skewed with a long right tail (see Section 3.5.2). The underlying reason for this skew was that P_{ij} increased exponentially with the diagonal distance between two nodes on G_2. Furthermore, for indirectly connected node-pairs P_{ij} was the product of the set sizes of its constituent, directly connected node-pairs.

Both the centrality of the entire distribution and the average of its left tail was proposed to be of importance when assessing redundancy. Centrality considered the redundancy of all shortest path sets while the average of the left tail focussed on the 'weakest links' — those shortest path sets that had the least alternatives.

Given the skewness of the distribution, the median of P_{ij} was a better estimator of centrality than the mean. The ordered set of P_{ij} was defined by $\boldsymbol{P}$. In the case of an even

number of sets the median was calculated by:

$$\tilde{P} = (P_{\text{mid}} + P_{\text{mid}+1})/2 \tag{5.1}$$

where $P_{\text{mid}} \in \boldsymbol{P}$, and

$$\text{mid} = (N^{1K}(N^{1K} - 1) + 1)/2 - 0.5 \tag{5.2}$$

with N^{1K} the number of nodes in G^{1K}. The number of sets in the theoretical instances were always even. Later in the case study instances the number of shortest path sets were not always even (refer to Chapter 7). In that case the median was calculated by:

$$\tilde{P} = P_{\text{mid}} \tag{5.3}$$

where $P_{\text{mid}} \in \boldsymbol{P}$, and

$$\text{mid} = (N^{1K}(N^{1K} - 1) + 1)/2 \tag{5.4}$$

Conservatively, the left tail was interpreted as the first quartile (i.e. the 25^{th} percentile) of the distribution of P_{ij}. The set of values that fell within the 25^{th} percentile was denoted by:

$$\boldsymbol{P}^{25\%} \subset \boldsymbol{P} \text{ such that } \boldsymbol{P}^{25\%} = \left\{P_1, P_2, \ldots, P_{\lceil \|\boldsymbol{P}\|/4 \rceil}\right\} \tag{5.5}$$

The mean of the elements of $\boldsymbol{P}^{25\%}$ was defined as:

$$\tilde{P}^{25\%} = \frac{\sum\limits_{P_n \in \boldsymbol{P}^{25\%}} P_n}{\|\boldsymbol{P}^{25^{\text{th}}}\|} \tag{5.6}$$

Although the primary interest was in the redundancy across *all* shortest path sets in $\mathcal{C}(\mathcal{S}_{ij})$, comparing these statistics to those of only the directly connected node-pairs (SD_{ij}) was considered prudent. Therefore, $\tilde{P}$ and $\tilde{P}^{25\%}$ were also calculated for the distribution of P_{ij} that only considered SD_{ij}. Table 5.1 summarises the four metrics defined for redundancy.

Table 5.1: Summary of redundancy metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\tilde{P}$(All)	(5.1)(5.3) $(P_{ij} \in \mathcal{C}(\mathcal{S}_{ij}))$
	SD_{ij}	$\tilde{P}$(Dir)	(5.1)(5.3) $(P_{ij} \in SD_{ij})$
Left-tail centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\tilde{P}^{25\%}$(All)	(5.6) $(P_{ij} \in \mathcal{C}(\mathcal{S}_{ij}))$
	SD_{ij}	$\tilde{P}^{25\%}$(Dir)	(5.6) $(P_{ij} \in SD_{ij})$

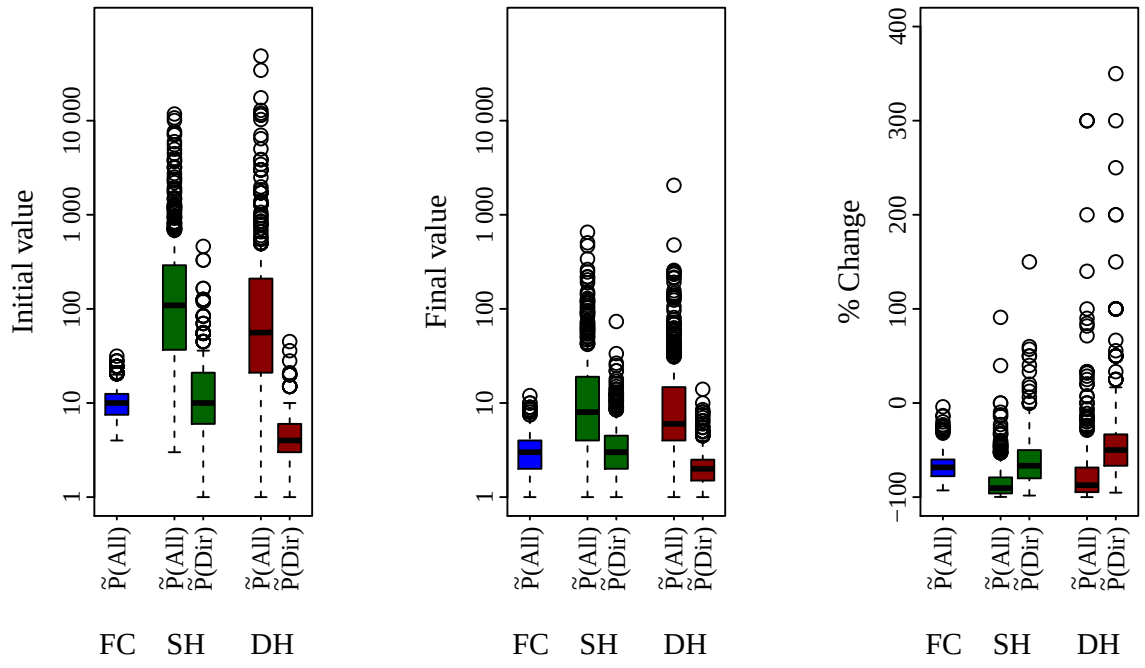
5.3.3 Results

After each disruption by the random error simulation, the four redundancy metrics were calculated. Thus, for each instance there existed a time-series of observations for each of the metrics from the initial, undisturbed network until the last state of the network before it became disconnected. Three measurements were extracted from these time-series. The *initial value* was the first observation of the undisturbed network. The *final value* was the last observation of the time-series before the instance became disconnected. The *% Change* measured the relative difference between the initial and final values.

The distributions of the initial values of $\tilde{P}$ (All) and $\tilde{P}$ (Dir) per network archetype are displayed in box plots in Figure 5.5a. Initially the distributions of $\tilde{P}$ (All) were very broad in the hub archetypes, spanning orders of magnitude. On the other hand, the FC instances were far more similar with the range of $\tilde{P}$ (Dir) between 4 and 31.5.

The distributions of the % Change in Figure 5.5c show that redundancy decreased in almost all of the instances (% change < 0). However, in the hub archetypes a number of exceptional instances saw an increase in redundancy before breaking (% change > 0). This could have happened when all the alternatives in a shortest path set were broken and a new set with longer shortest paths was formed. If this new set offered more alternatives than the initial shortest path set, it would have increased the redundancy of the instance.

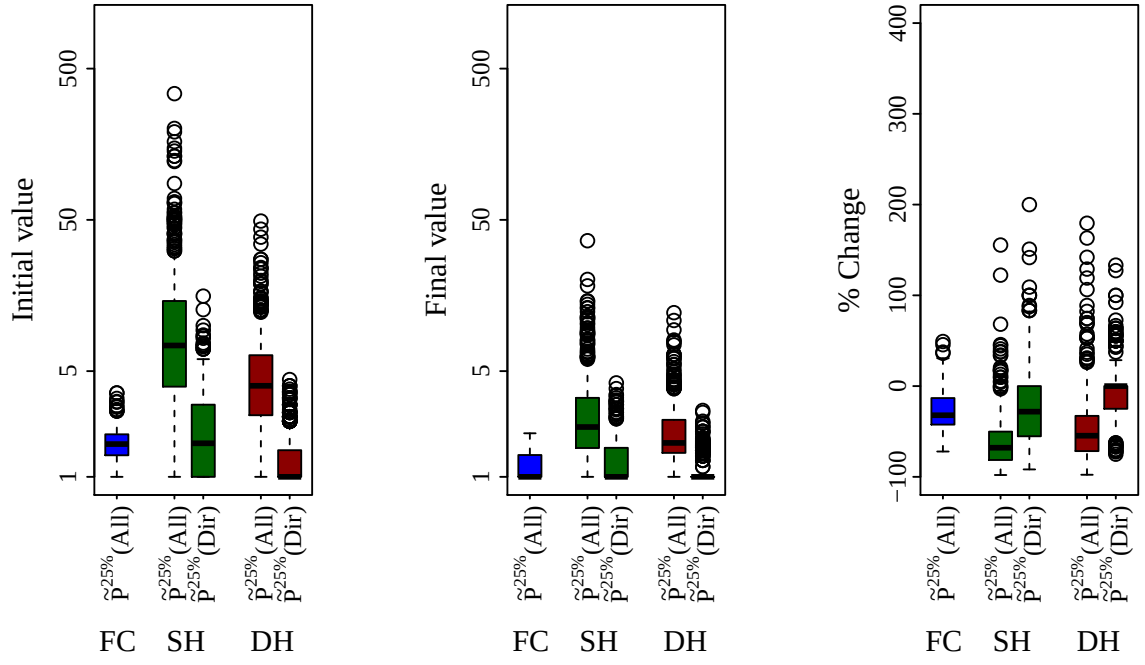
While redundancy decreased significantly before disconnection, the distributions of the final values in Figure 5.5b were still broad for the hub archetypes.



(a) Initial value (b) Final value (c) % Change

Figure 5.5: Distributions of the three measurements of $\tilde{P}$(All) and $\tilde{P}$(Dir). (Three outliers of DH: $\tilde{P}$(All) not visible on graph (b): 500, 500 and 3 900.)

The distributions of the initial values (Figure 5.6a), final values (Figure 5.6b) and % Change (Figure 5.6c) of $\tilde{P}^{25\%}$ showed the same trends as the distributions of $\tilde{P}$.



(a) Initial value distributions (b) Final value distributions (c) % Change distributions

Figure 5.6: Distributions of the three measurements of $\tilde{P}^{25\%}$(All) and $\tilde{P}^{25\%}$(Dir).

The first insight gained by examining the distributions was that redundancy, expressed by the four metrics, does indeed decrease the closer instances come to disconnection. The second insight, however, was that significant levels of redundancy were still present in networks right before disconnection (Figure 5.5b). Even the least redundant shortest path sets still had a number of alternatives available right before the final blow disconnected the instance (Figure 5.6b). Finally we noticed that redundancy was the most variable in the hub archetypes. These archetypes had additional constraints that routed shortest paths via hubs. These constraints induced a greater overlap of shorter paths in the hub archetypes compared to the FC archetype (Figure 4.12). The conclusion was that redundancy alone was not the only determinant of vulnerability. If there existed 20 alternative paths but all shared one specific link in G^2, that network could still have been very vulnerable. The following category of vulnerability metrics focussed on quantifying this overlap.

5.4 Vulnerability category 2: Overlap

5.4.1 Conceptual description

Reverting back to the illustrative example, we can observe two types of overlap. The first is where one road segment features on multiple path sequences between two facilities. An example is e_{OP} that features in every path between the Airfreight Warehouse and the Distribution Centre. The second type of overlap is where a road segment features in more that one set of paths. For example e_{EF} is part of the path set between the Airfreight Warehouse and the Distribution Centre, but one would also travel on e_{EF} when going

from the Retail Outlet back to the Distribution Centre. So in this case two separate path sets overlap in that they utilise the same link.

5.4.2 Formulation of metrics

The best way to account for both types of overlap was to calculate the link betweenness of the road segments as described in the Overall link betweenness (Overall-B) and Elemental link betweenness (Elemental-B) simulation strategies. Using (4.1) and (4.2) the relative link betweenness for each link in G^2 could be calculated considering $\mathcal{C}(\mathcal{S}_{ij})$ and SD_{ij}, respectively.

Similar to redundancy, we were interested in two characteristics of the distributions of relative link betweenness: the centrality and the span of the right tail. The higher the central value of the link betweenness distribution, the less diversity there was in alternative shortest paths. This implied a high level of overlap across many shortest path sets. The span of the right tail gave an indication of how pivotal those links with highest betweenness scores were compared to the rest of the links.

Although the Overall-B strategy proved more effective than the Elemental-B strategy on all fronts, both types of link betweenness were proposed as vulnerability metrics. Overall-B accentuates the relative importance of the most between links while elemental provides a more even scoring as it doesn't double-count shortest paths. Therefore, because the focus was not on identifying one pivotal link, but rather a group of pivotal links, there was merit in using Elemental-B as well.

We started by defining $\boldsymbol{B}_{overall}$ as the set of Overall-B (4.1) scores for links in G^2 in descending order. Similarly, $\boldsymbol{B}_{elemental}$ was the set of Elemental-B (4.2) scores for links in G^2 in descending order. The centrality of the link betweenness distributions was then defined as:

$$\overline{B}_{\text{overall}} = \frac{\sum\limits_{B_n \in \boldsymbol{B}_{overall}} B_n}{\|\boldsymbol{B}_{overall}\|} \tag{5.7}$$

and

$$\overline{B}_{\text{elemental}} = \frac{\sum\limits_{B_n \in \boldsymbol{B}_{elemental}} B_n}{\|\boldsymbol{B}_{elemental}\|} \tag{5.8}$$

To investigate the span of the right tail the range between the 75$^{\text{th}}$ percentile and the maximum value was calculated. The range of the tail gave a better idea of *how much more* important the links in the right tail were compared to the rest of the links. The value of the 75$^{\text{th}}$ percentile was denoted by:

$$B^{75\%}_{overall} = B_{\lfloor \|\boldsymbol{B}_{overall}\|/4 \rfloor} \tag{5.9}$$

and

$$B^{75\%}_{elemental} = B_{\lfloor \|\boldsymbol{B}_{elemental}\|/4 \rfloor}. \tag{5.10}$$

The range was then defined by:

$$R(B_{\text{overall}})^{75\%} = \max(\boldsymbol{B}_{overall}) - B^{75\%}_{overall} \tag{5.11}$$

and

$$R(B_{\text{elemental}})^{75\%} = \max(\boldsymbol{B}_{elemental}) - B^{75\%}_{elemental}. \tag{5.12}$$

Table 5.2 summarises the four metrics developed to quantify overlap.

Table 5.2: Summary of overlap metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\overline{B}_{\text{overall}}$	(5.7)
	SD_{ij}	$\overline{B}_{\text{elemental}}$	(5.8)
Right-tail range	$\mathcal{C}(\mathcal{S}_{ij})$	$R(B_{\text{overall}})^{75\%}$	(5.11)
	SD_{ij}	$R(B_{\text{elemental}})^{75\%}$	(5.12)

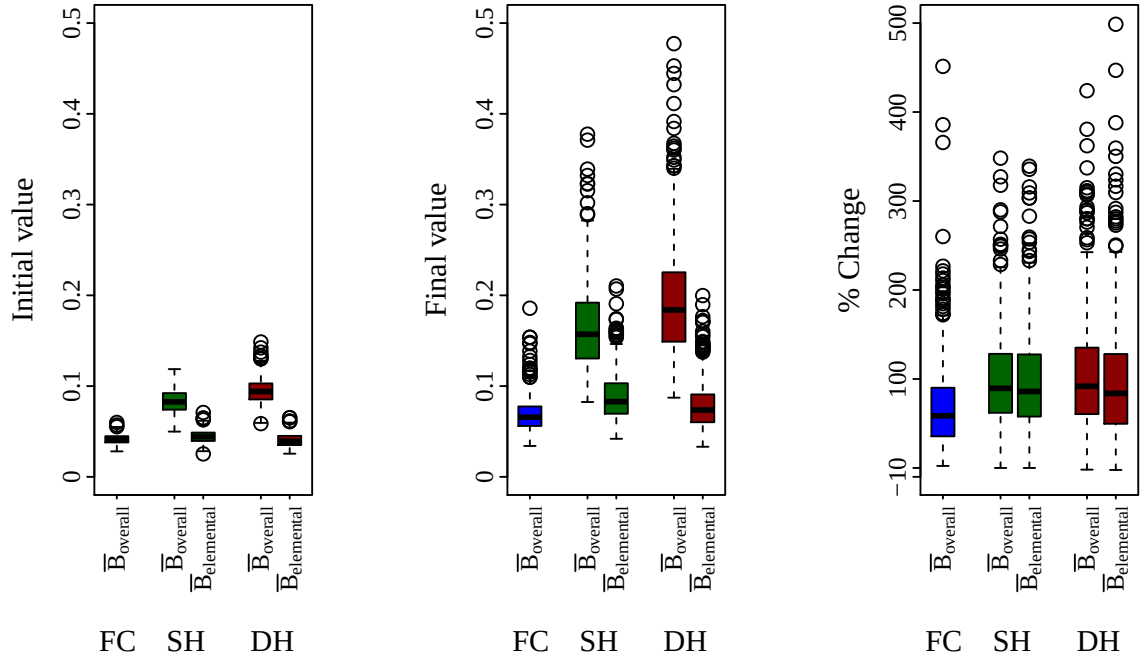
5.4.3 Results

Following the same procedure as for the redundancy metrics, time-series of overlap metrics were calculated for each instance from the initial, undisturbed network until the last state of the network before it became disconnected. The same three measurements, namely *initial value*, *final value* and *% Change* were extracted from these time-series for each instance.

The distributions of the initial values of $\overline{B}_{\text{overall}}$ and $\overline{B}_{\text{elemental}}$ (Figure 5.7a) show that there wasn't a great level of overlap present in any of the archetypes before the simulation started. This could have seemed surprising as the targeted attack strategies based on betweenness metrics had been most effective. The explanation lay in the initial values of $R(B_{\text{overall}})^{75\%}$ and $R(B_{\text{elemental}})^{75\%}$ as shown in Figure 5.8a. For instance, in the SH archetype the bulk of $R(B_{\text{overall}})^{75\%}$ values lay between 0.3 and 0.5. This meant that in most instances there had been at least one link that had a betweenness score more than 30% higher than three quarters of the other links. This link or multiple links were the anchors of a great deal of overlap in those instances and were prioritised for removal during the targeted attack simulations.

The distributions of the final values of $\overline{B}_{\text{overall}}$ and $\overline{B}_{\text{elemental}}$ (Figure 5.7b) showed that the overlap increased across all archetypes. This occurred because the reduced number of links in G^2 concentrated the shortest paths on the remaining links. It was also evident that the distributions had become broader, meaning that the impact of the random error simulation on the overlap was greatly instance-specific. Interestingly, the rate of change from the initial to final values was similar regardless of whether $\overline{B}_{\text{overall}}$ or $\overline{B}_{\text{elemental}}$ was considered (Figure 5.7c).

The distributions of the initial values of $R(B_{\text{overall}})^{75\%}$ and $R(B_{\text{elemental}})^{75\%}$ in Figure 5.8a showed that the FC archetype had the shortest right tail of all. It illustrated that when networks are not constrained by hubs the importance of the links in G^2 are far more homogenous. However, the right tail did become more pronounced as more



(a) Initial value (b) Final value (c) % Change

Figure 5.7: Distributions of the three measurements of $\overline{B}_{\text{overall}}$ and $\overline{B}_{\text{elemental}}$.

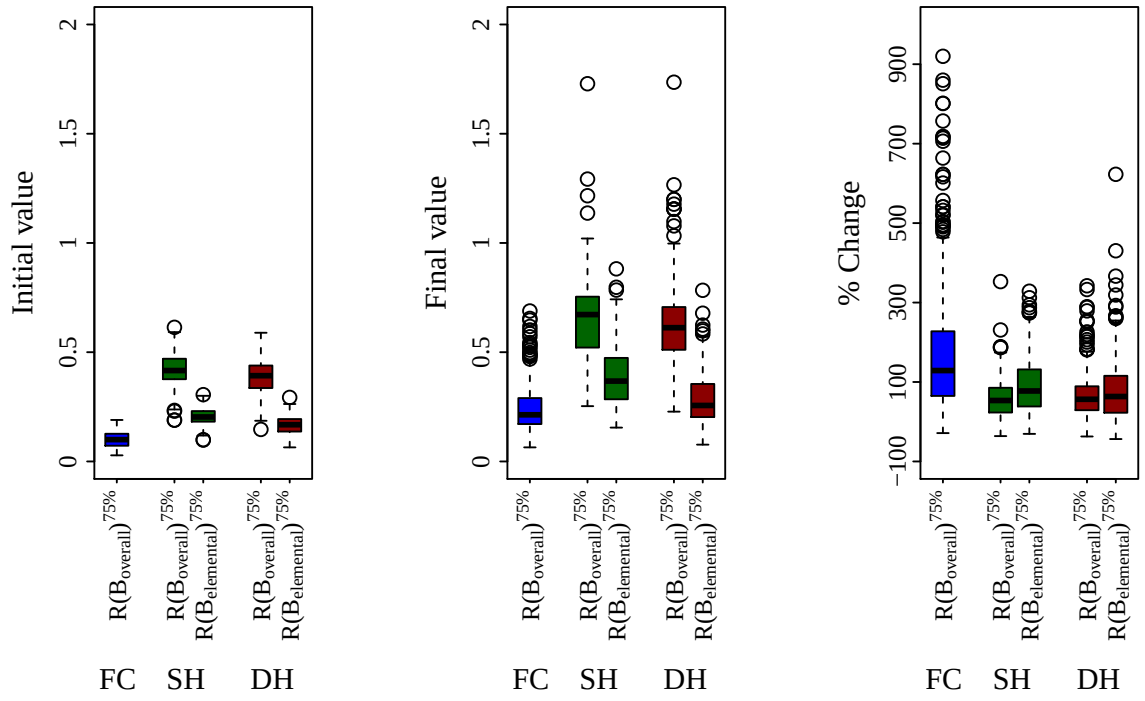
links were removed from G^2 as evidenced in the distribution of the final values in Figure 5.8b. The % Change for the FC archetype was remarkably high with most instances doubling the range of the right-tail (% change ¿ 100% in Figure 5.8c). Furthermore, the broadness of the distribution of the % Change implied that the degree to which the right tail increased was instance-specific. The distributions of the $R(B_{\text{elemental}})^{75\%}$ and $R(B_{\text{overall}})^{75\%}$ showed that the right tails of the hub archetypes also increased, showing a more pronounced dependence on a few critical links.

A curious observation in Figure 5.8b was that in both hub archetypes there were instances where $R(B_{\text{overall}})^{75\%} > 1$, meaning that one or more links in those instances had a relative betweenness score higher than 1. This seemed counterintuitive at first and is definitely not a phenomena frequently noted in vulnerability studies. This phenomena was attributed to the multilayered nature of $\mathcal{M}$ and will be best explained referring back to the illustrative example.

Imagine that pervasive strike action by unionised airport staff have closed off various road segments around O.R. Tambo International Airport. This disruption completely disabled *Paths 5 & 6* from the Distribution Centre to the Airfreight Warehouse, leaving only *Path 4* (Figure 5.9). Furthermore, all the original paths between the Airfreight Warehouse and Distribution Centre were disabled. Because e_{JI} was disabled and only e_{IJ} (indicated in red) remained, a new *Path 11* had to be created that didn't require e_{JI}. This path deviated via $e_{JE} \rightarrow e_{EF} \rightarrow e_{FG}$. In a similar vein all the paths from the Distribution Centre to the Retail outlet were also disabled and new paths had to be created. Because link e_{FE} had been disabled and only e_{EF} remained (indicated in red) *Paths 12 & 13* both deviated via $e_{HI} \rightarrow e_{IJ} \rightarrow e_{JE}$. Therefore, in any shortest path combination from the Airfreight Warehouse to the Retail Outlet e_{JE} will feature twice. This phenomena was

called "doubling back".

If there were many paths in a network that doubled-back, it resulted in one or more links in G^2 having a relative betweenness > 1 as the numerator in (4.1), which was the count of occurrences, became larger than the denominator, which was the total number of shortest paths. This phenomena occurred specifically because shortest paths were constrained by both logical relationships in G^{1K} and physical infrastructure in G^2. It was evident from the distributions in Figure 5.8b that many paths doubled-back in the hub archetypes before disconnection.



(a) Initial value (b) Final value (c) % Change

Figure 5.8: Distributions of the three measurements of $R(B_{\text{overall}})^{75\%}$ and $R(B_{\text{elemental}})^{75\%}$. (One outlier of FC not visible on graph: % Change in $R(B_{\text{overall}})^{75\%} = 1\,498$.)

In summary, overlap increased across the board with instances becoming increasingly more dependent on a few critical links. However, the magnitude of the change seemed to be very instance-specific.

The first two vulnerability categories considered the size of the shortest path sets and the overlap. The final category investigated changes in the length of the shortest paths for possible hints of increasing vulnerability.

5.5 Vulnerability category 3: Efficiency step-change

5.5.1 Conceptual design

The average shortest path of a network increased when all the paths in one or more shortest path sets were broken and new shortest path sets were routed. Figure 5.10 illustrates this in a simple grid example. Two nodes, x_i and x_j, have a set of 20 shortest paths connecting them (Figure 5.10a). Each path has a length of 6. A disruption removes

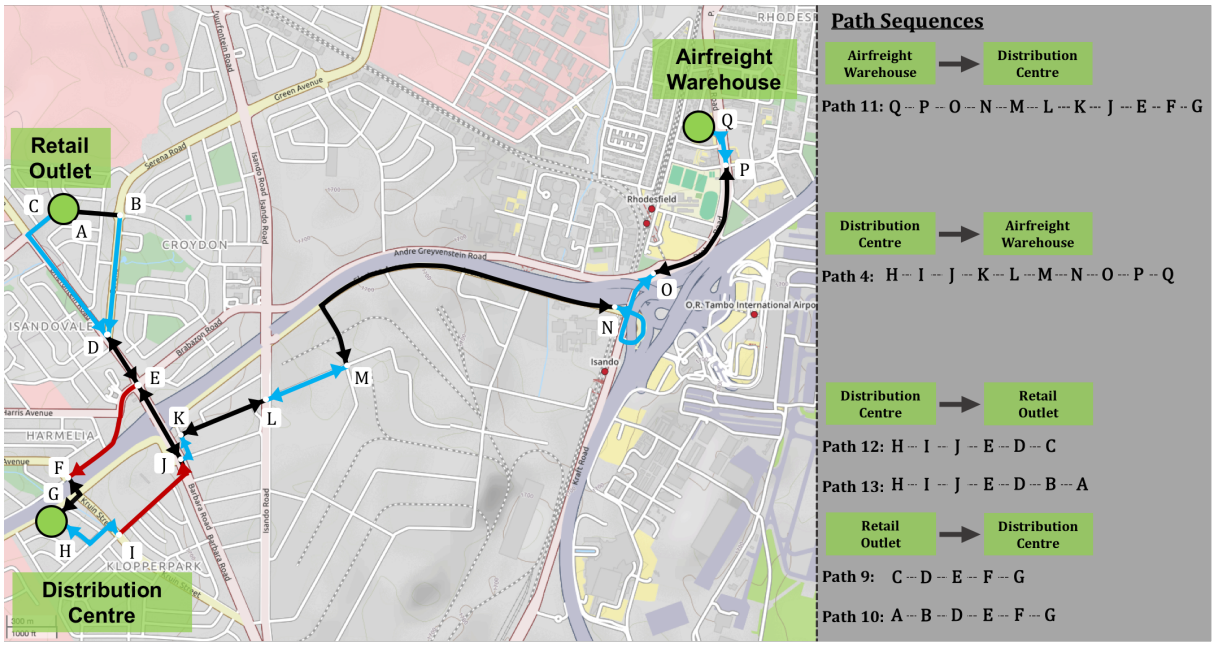
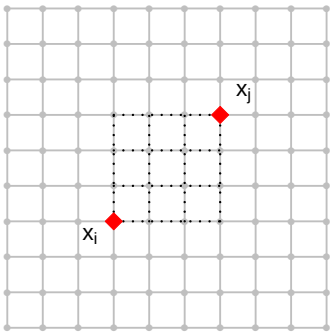


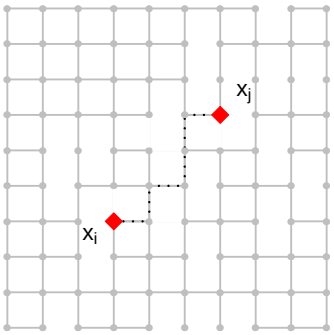
Figure 5.9: Shortest path alternatives of the illustrative supply chain network after significant disruptions in the road network (Source: OpenStreetMap contributors (2017)).

26 of the grid links so that all but one of the original shortest paths remain (Figure 5.10b). Although P_{ij} has decreased, L_{ij} has remained the same. A next progressive disruption removes a further 11 grid links so that the last remaining shortest path is also broken. A new set of 4 shortest paths of length 10 is established (Figure 5.10c). If these two nodes were part of a larger network the $\bar{L}$ would have increased after the last disruption. This is Scenario A.



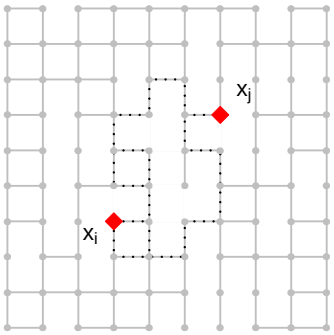
P_{ij}	20
L_{ij}	6

(a) Initial, undisturbed grid



P_{ij}	1
L_{ij}	6
Relative change in L_{ij}	$\frac{6-6}{6} = 0$

(b) First disruption



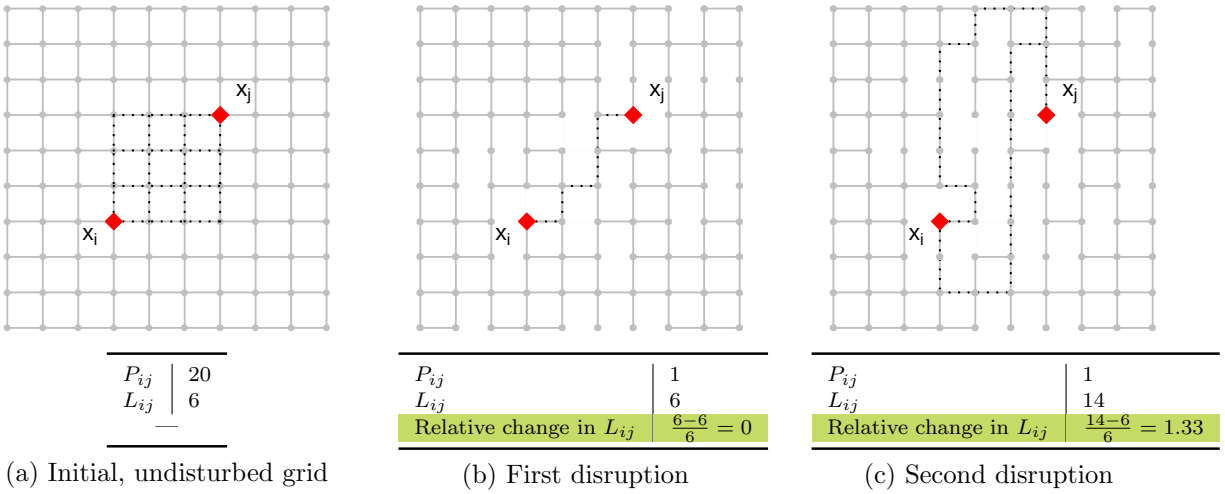
P_{ij}	4
L_{ij}	10
Relative change in L_{ij}	$\frac{10-6}{6} = 0.67$

(c) Second disruption

Figure 5.10: Step-change example: Scenario A. Two nodes on a grid are affected by two successive disruptions of the underlying grid.

Consider now a Scenario B where the same two nodes (Figure 5.11a) undergo an identical first disruption (Figure 5.11b). This time the second disruption removes a *different* set of 11 grid links (Figure 5.11c). As a result only one shortest path of length 14 could

be found.



P_{ij}	20
L_{ij}	6
—	

(a) Initial, undisturbed grid

P_{ij}	1
L_{ij}	6
Relative change in L_{ij}	$\frac{6-6}{6} = 0$

(b) First disruption

P_{ij}	1
L_{ij}	14
Relative change in L_{ij}	$\frac{14-6}{6} = 1.33$

(c) Second disruption

Figure 5.11: Step-change example: Scenario B. Two nodes on a grid are affected by two successive disruptions of the underlying grid.

The second disruption in Scenario B many of the grid nodes between x_i and x_j whereas Scenario A's second disruption removed grid links at the edge of the grid — less important to x_i and x_j. Scenario B thus leaves the node-pair in much greater risk of disconnection than Scenario A. Simultaneously, the step-change in L_{ij} from 6 to 14 in Scenario B is greater than the step-change from 6 to 10 in Scenario A.

While the results from the targeted attack and random error simulations showed that the % Change in $\bar{L}$ was not necessarily a good indicator of imminent disconnection, the idea of step-change held promise. A large step-change was considered an indication of the level of sparseness in G^2. The efficiency step-change was defined as the relative change in the shortest path length of a node-pair. The next section proposes a metric to capture this concept on an aggregate network level.

5.5.2 Formulation of metrics

To define the relative step-change we started by defining the difference between the shortest path length of a specific node-pair (x_i^{1K}, x_j^{1K}) measured before and after a disruption:

$$\Delta L_{ij}(t; t+z) = L_{ij}(t) - L_{ij}(t+z) \tag{5.13}$$

where t is some defined point in time before a disruption occurred and z is some time after a disruption of G^2 occurred. The relative change is then:

$$Rel\Delta L_{ij}(t; t+z) = \frac{\Delta L_{ij}(t; t+z)}{L_{ij}(t)} \tag{5.14}$$

So in Scenario A $Rel\Delta L_{ij}(1; 1{+}2) = \frac{10-6}{6} = 0.67$ after two disruptions while in Scenario B $Rel\Delta L_{ij}(1; 1+2) = \frac{14-6}{6} = 1.33$. This reflected the step-change for an individual node-pair. In the interest of overall vulnerability, we aggregated the step-changes across all

node-pairs:

$$Rel\overline{\Delta L}(t;t+z) = \frac{\sum\limits_{i,j;i\neq j} Rel\Delta L_{ij}(t;t+z)}{N^{1K}(N^{1K}-1)} \text{ where } i,j \in \{1,2,\dots N^{1K}\} \tag{5.15}$$

Only one metric was developed in this vulnerability category, as shown in Table 5.3.

Table 5.3: Summary of efficiency step-change metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$Rel\overline{\Delta L}(t;t+z)$	(5.15)

5.5.3 Results

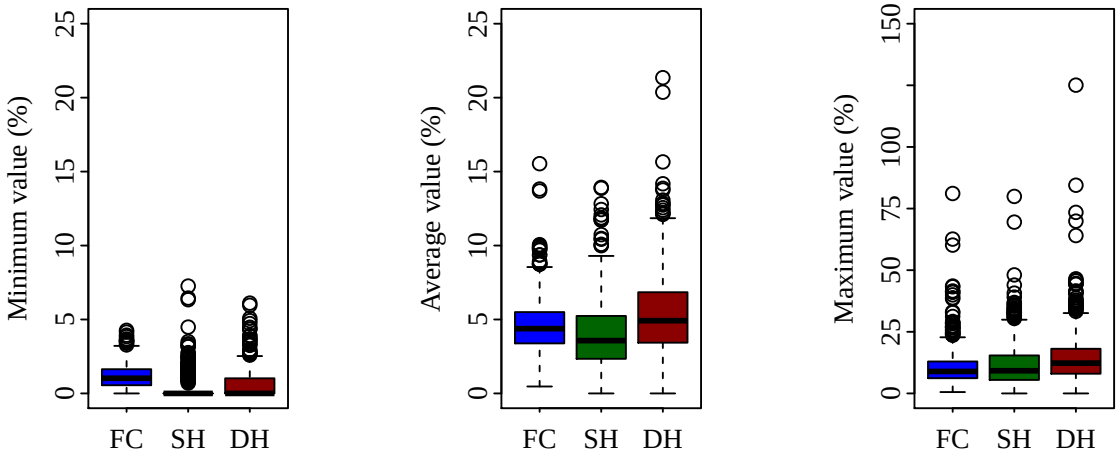
The efficiency step-change was measured after each disruption until disconnection, providing a time-series of values for each instance. In this thesis, step-change was measured over only one disruption, not multiple disruptions (i.e. $z = 1$). To investigate this metric, it made more sense to extract the *minimum*, *maximum* and *average* values from the time series instead of the initial values, % Change and final values as was done for the other two categories.

It was notable that the distributions of the minimum, average and maximum values of the efficiency step-change were very similar across the three archetypes (Figure 5.12). This was in contrast to the other vulnerability metrics where there were definite differences between the archetypes. The majority of instances in all the archetypes displayed minimum step-changes under 5 (Figure 5.12a) and maximum step-changes under 50 (Figure 5.12c) with an average around 5 (Figure 5.12a). The distributions were also not as broad as for the other metrics. One reason for this was that L_{ij}, upon which this metric was based, was a summative function and not a combinatorial product like P_{ij}. Another reason could have been that this metric was more generalisable and not as instance-specific as the others.

This chapter described the random error simulation that was used to validate three categories of vulnerability metrics and presented the results of this simulation in terms of efficiency loss and disconnection. The three vulnerability categories namely redundancy, overlap and efficiency step-change were discussed individually. For each category a conceptual description was given followed by the mathematical formulation of the metrics. Time-series of values for each metric were calculated during the random error simulation. The behaviour of each metric was analysed by extracting three measurements from these time-series and plotting their distributions.

The responses of the vulnerability metrics could be intuitively explained with regards to the real-life problem addressed by this thesis. That in itself was positive feedback regarding the usefulness of the first artefact of the thesis — the multilayered network formulation of $\mathcal{M}$.

While the behaviour of the metrics followed expectation, that was not enough to assert their validity in quantifying the vulnerability of $\mathcal{M}$. The next chapter continues the evaluation of the metrics by testing a number of hypotheses.



(a) Minimum value (b) Average value (c) Maximum value

Figure 5.12: Distributions of the three measurements of $Rel\overline{\Delta L}(t; t+1)$.

Chapter 6

Statistical validation of vulnerability metrics

The hypothesis tests in this chapter address the first aspect of the fourth objective of the thesis:

> 4. Evaluation of the validity of the suite of vulnerability metrics through **statistical analysis** and a real-life case study.

Chapters 7 and 8 will present how the validity of the metrics was tested using a real-life case study.

6.1 Statistical tests

In evaluating the statistical relation between the vulnerability metrics and the actual vulnerability of $\mathcal{M}$ it was necessary to differentiate between *correlation* and *causation*.

Correlation: A statistical measure that expresses the strength and directionality of the relationship between two or more variables.

Causation: An indication that one variable is causal in the behaviour of another variable.

Correlation between two variables does not imply causation. Variables may be correlated due to external factors or mutual relation to a third variable. However, if there is a causal relationship between two variables then they are most definitely also correlated. While we could test for correlation using established statistical methods, establishing causality required experiments with control groups. This was left for future work. Although we could not test for causality, we did test whether certain variables could discriminate between likely outcomes in a next disruption.

6.1.1 Correlation of efficiency loss and robustness to the vulnerability metrics

The two levels of damage assessed were efficiency loss and how quickly an instance became disconnected. The longer it took to disconnect an instance, the more robust it was considered to be. The interest was thus in identifying whether any of the vulnerability metrics had a strong correlation with the efficiency loss, the robustness or both. In this thesis we limited the evaluation to single variate correlation.

Spearman and Kendall pairwise correlation tests

Pearson's r correlation is the most widely used bivariate correlation test. Unfortunately, it requires three assumptions that all proved troublesome:

- *The variables must be normally distributed.* The redundancy and overlap metrics were not normally distributed and had heavy right tails (Figures 5.5–5.8).
- *The variables must be linearly related.* The combinatorial nature of P_{ij} made linear correlations including redundancy and overlap metrics unlikely.
- *Variables must be homoscedastic — i.e. have the same finite variance.* Finite variance could not be comfortably assumed.

These assumptions could not be made with much confidence regarding all the vulnerability metrics and therefore less restrictive tests had to be used.

The Spearman and Kendall rank correlations are both non-parametric tests that do not make any assumptions regarding the underlying distributions. The Spearman test determines the strength and direction of the *monotonic* relationship between two variables. It requires that variables be at least ordinal. The Kendall test determines whether variables are similarly ordered. Both Spearman's ρ and Kendall's τ are real numbers on the interval $(-1; 1)$ where -1 indicates the strongest possible negative correlation and $+1$ the strongest possible positive correlation. In both tests the significance of the result is expressed by means of the p-value. The p-value is the probability of observing the calculated correlation if, in reality, no correlation exists. A p-value < 0.05 is considered a significant result.

The correlations of the vulnerability metrics to efficiency loss and robustness were tested using both the Spearman and Kendall tests. The two tests yielded similar results. Both identified the same bivariate correlations as significant. In each of these cases both tests coincided in terms of the direction of the correlation. Kendall's test, however, consistently calculated a weaker correlation than Spearman's.

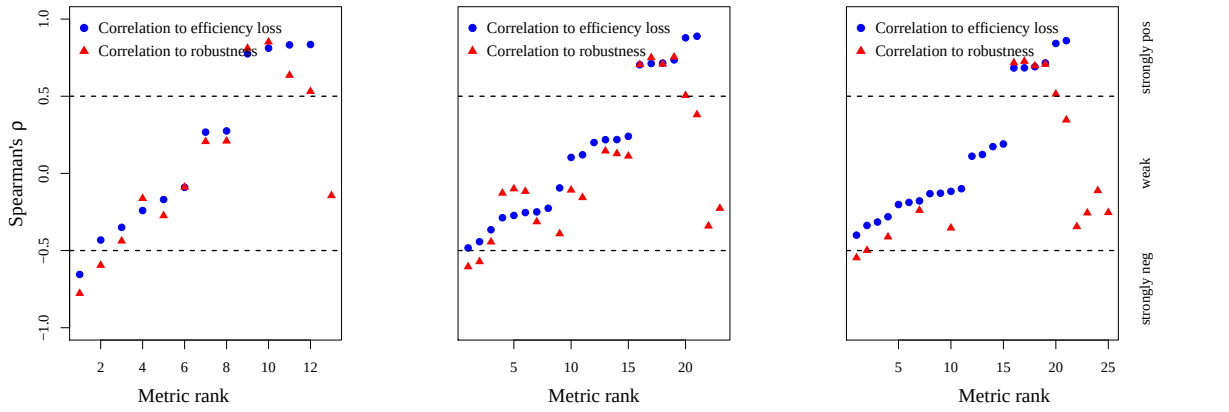
The Spearman correlations were used to test hypotheses regarding the relationship between the vulnerability metrics and efficiency loss or robustness. Six sets of hypotheses were formulated. In these formulations v_i represents the vulnerability metrics where $i \in \{\tilde{P}(\text{All}), \tilde{P}(\text{Dir}), \tilde{P}^{25\%}(\text{All}), \tilde{P}^{25\%}(\text{Dir}), \overline{B}_{\text{overall}}, \overline{B}_{\text{elemental}}, R(B_{\text{overall}})^{75\%}, R(B_{\text{elemental}})^{75\%}, Rel\overline{\Delta L}(t; t+z)\}$ and $K \in \{F, S, D\}$ represents the different archetypes.

1. Correlation between the *initial value* of v_i and *efficiency loss.*
   - H_0: There is a correlation between the initial value of v_i and the efficiency loss in the instances of archetype K under a random error strategy.
   - H_A: There is no correlation between the initial value of v_i and the efficiency loss in the instances of archetype K under a random error strategy.
2. Correlation between the *% change* of v_i and *efficiency loss.*
   - H_0: There is a correlation between the % change of v_i and the efficiency loss in the instances of archetype K under a random error strategy.
   - H_A: There is no correlation between the % change of v_i and the efficiency loss in the instances of archetype K under a random error strategy.

3. Correlation between the *final value* of v_i and *efficiency loss*.
    - H_0: There is a correlation between the final value of v_i and the efficiency loss in the instances of archetype K under a random error strategy.
    - H_A: There is no correlation between the final value of v_i and the efficiency loss in the instances of archetype K under a random error strategy.
4. Correlation between the *initial value* of v_i and *robustness*.
    - H_0: There is a correlation between the initial value of v_i and the robustness in the instances of archetype K under a random error strategy.
    - H_A: There is no correlation between the initial value of v_i and the robustness in the instances of archetype K under a random error strategy.
5. Correlation between the *% change* of v_i and *robustness*.
    - H_0: There is a correlation between the % change of v_i and the robustness in the instances of archetype K under a random error strategy.
    - H_A: There is no correlation between the % change of v_i and the robustness in the instances of archetype K under a random error strategy.
6. Correlation between the *final value* of v_i and *robustness*.
    - H_0: There is a correlation between the final value of v_i and the robustness in the instances of archetype K under a random error strategy.
    - H_A: There is no correlation between the final value of v_i and the robustness in the instances of archetype K under a random error strategy.

Using the Spearman correlation test, if two variables had a significant correlation (p-value < 0.05) then we failed to reject the H_0 that those variables were correlated. If, on the other hand, the p-value ≥ 0.05 we rejected the H_0 and accepted H_A. Figure 6.1 plots the correlation values of those vulnerability metrics that had a significant correlation to efficiency loss, robustness or both. The metrics are ranked according to increasingly positive correlation to efficiency loss. In the case of the Single Hub (SH) and Double Hub (DH) archetypes at least 70% of the significant relationships had a relatively weak correlation ($-0.5 < \rho < 0.5$), while in the Fully Connected (FC) archetype 55% of the significant correlations were relatively weak. The correlation results are tabulated in Tables 6.1–6.3. All the cases where H_0 was rejected are identified by the label *insig*. Every other value tabulated indicates a significant correlation, meaning that H_0 could not be rejected for that pair of variables. Variables with a strongly positive or strongly negative correlation are highlighted in the tables.

The significant correlations that were observed seemed, for lack of a better word, vexing.



	Metric	Measure
1	$\tilde{P}$(All)	Final
2	$\tilde{P}$(All)	% Change
3	$\tilde{P}^{25\%}$(All)	Final
4	$\tilde{P}$(All)	Initial
5	$\tilde{P}^{25\%}$(All)	% Change
6	$\overline{B}_{\text{overall}}$	Initial
7	$Rel\overline{\Delta L}(t;t+1)$	Minimum
8	$R(B_{\text{overall}})^{75\%}$	% Change
9	$R(B_{\text{overall}})^{75\%}$	Final
10	$\overline{B}_{\text{overall}}$	% Change
11	$Rel\overline{\Delta L}(t;t+1)$	Maximum
12	$Rel\overline{\Delta L}(t;t+1)$	Average

(a) FC archetype

	Metric	Measure
1	$\tilde{P}$(All)	Final
2	$\tilde{P}$(Dir)	Final
3	$\tilde{P}^{25\%}$(All)	Final
4	$\tilde{P}$(All)	Initial
5	$\tilde{P}^{25\%}$(All)	Initial
6	$\tilde{P}$(Dir)	Initial
7	$\tilde{P}^{25\%}$(Dir)	Final
8	$\tilde{P}^{25\%}$(Dir)	Initial
9	$\tilde{P}$(All)	% Change
10	$\tilde{P}$(Dir)	% Change
11	$\tilde{P}^{25\%}$(All)	% Change
12	$\tilde{P}^{25\%}$(Dir)	% Change
13	$Rel\overline{\Delta L}(t;t+1)$	Minimum
14	$R(B_{\text{overall}})^{75\%}$	Final
15	$R(B_{\text{elemental}})^{75\%}$	% Change
16	$R(B_{\text{overall}})^{75\%}$	% Change
17	$R(B_{\text{elemental}})^{75\%}$	Final
18	$\overline{B}_{\text{elemental}}$	Final
19	$\overline{B}_{\text{elemental}}$	% Change
20	$\overline{B}_{\text{overall}}$	Final
21	$\overline{B}_{\text{overall}}$	% Change
22	$Rel\overline{\Delta L}(t;t+1)$	Maximum
23	$Rel\overline{\Delta L}(t;t+1)$	Average

(b) SH archetype

	Metric	Measure
1	$\tilde{P}$(All)	Final
2	$\tilde{P}$(Dir)	Final
3	$\tilde{P}$(All)	Initial
4	$\tilde{P}^{25\%}$(All)	Final
5	$\tilde{P}^{25\%}$(All)	Initial
6	$\tilde{P}$(Dir)	Initial
7	$\tilde{P}^{25\%}$(Dir)	Final
8	$R(B_{\text{overall}})^{75\%}$	Initial
9	$\tilde{P}^{25\%}$(Dir)	Initial
10	$\tilde{P}$(Dir)	% Change
11	$R(B_{\text{elemental}})^{75\%}$	Initial
12	$\tilde{P}$(All)	% Change
13	$\tilde{P}^{25\%}$(All)	% Change
14	$\tilde{P}^{25\%}$(Dir)	% Change
15	$Rel\overline{\Delta L}(t;t+1)$	Minimum
16	$R(B_{\text{overall}})^{75\%}$	Final
17	$R(B_{\text{elemental}})^{75\%}$	Final
18	$R(B_{\text{elemental}})^{75\%}$	% Change
19	$R(B_{\text{overall}})^{75\%}$	% Change
20	$\overline{B}_{\text{overall}}$	Final
21	$\overline{B}_{\text{overall}}$	% Change
22	$\overline{B}_{\text{elemental}}$	% Change
23	$\overline{B}_{\text{elemental}}$	Final
24	$Rel\overline{\Delta L}(t;t+1)$	Maximum
25	$Rel\overline{\Delta L}(t;t+1)$	Average

(c) DH archetype

Figure 6.1: Correlations of all significant relationships between the vulnerability metrics and efficiency loss or robustness.

Table 6.1: Spearman's correlation (ρ) between the vulnerability metrics and efficiency loss and the vulnerability metrics and robustness for the FC archetype. Insignificant correlations (p-value < 0.05) are indicated by *insig.* Strongly positive and strongly negative correlations are indicated by pink and green highlighting, respectively.

	Metric	Measurement	% Efficiency loss	Robustness
Redundancy	$\tilde{P}$(All)	Initial value	-0.24	-0.16
		% Change	-0.43	-0.60
		Final value	-0.65	-0.78
	$\tilde{P}^{25\%}$(All)	Initial value	*insig.*	*insig.*
		% Change	-0.17	-0.27
		Final value	-0.35	-0.44
Overlap	$\overline{B}_{\text{overall}}$	Initial value	-0.09	-0.09
		% Change	0.77	0.81
		Final value	0.81	0.85
	$R(B_{\text{overall}})^{75\%}$	Initial value	*insig.*	*insig.*
		% Change	0.27	0.21
		Final value	0.28	0.21
Efficiency step-change	$Rel\overline{\Delta L}(t;t+1)$	Minimum	*insig.*	-0.14
		Average	0.84	0.53
		Maximum	0.83	0.64

Table 6.2: Spearman's correlation (ρ) between the vulnerability metrics and efficiency loss and the vulnerability metrics and robustness for the SH archetype. Insignificant correlations (p-value < 0.05) are indicated by *insig.* Strongly positive and strongly negative correlations are indicated by pink and green highlighting, respectively.

	Metric	Measurement	% Efficiency loss	Robustness
Redundancy	$\tilde{P}$(All)	Initial value	-0.29	-0.13
		% Change	-0.09	-0.39
		Final value	-0.48	-0.60
	$\tilde{P}^{25\%}$(All)	Initial value	-0.27	-0.10
		% Change	*insig.*	-0.23
		Final value	-0.37	-0.44
	$\tilde{P}$(Dir)	Initial value	-0.25	-0.12
		% Change	*insig.*	-0.34
		Final value	-0.44	-0.57
	$\tilde{P}^{25\%}$(Dir)	Initial value	-0.23	*insig.*
		% Change	0.10	-0.11
		Final value	-0.25	-0.31
Overlap	$\overline{B}_{\text{overall}}$	Initial value	*insig.*	*insig.*
		% Change	0.74	0.75
		Final value	0.71	0.71
	$R(B_{\text{overall}})^{75\%}$	Initial value	*insig.*	*insig.*
		% Change	0.22	0.13
		Final value	0.20	*insig.*
	$\overline{B}_{\text{elemental}}$	Initial value	*insig.*	*insig.*
		% Change	0.71	0.75
		Final value	0.70	0.70
	$R(B_{\text{elemental}})^{75\%}$	Initial value	*insig.*	*insig.*
		% Change	0.22	0.15
		Final value	0.24	0.11
Efficiency step-change	$Rel\overline{\Delta L}(t;t+1)$	Minimum	0.12	-0.16
		Average	0.89	0.38
		Maximum	0.88	0.50

Table 6.3: Spearman's correlation (ρ) between the vulnerability metrics and efficiency loss and the vulnerability metrics and robustness for the DH archetype. Insignificant correlations (p-value < 0.05) are indicated by *insig.* Strongly positive and strongly negative correlations are indicated by pink and green highlighting, respectively.

	Metric	Measurement	% Efficiency loss	Robustness
Redundancy	$\tilde{P}$(All)	Initial value	-0.32	*insig.*
		% Change	*insig.*	-0.34
		Final value	-0.40	-0.55
	$\tilde{P}^{25\%}$(All)	Initial value	-0.20	*insig.*
		% Change	*insig.*	-0.26
		Final value	-0.28	-0.41
	$\tilde{P}$(Dir)	Initial value	-0.19	*insig.*
		% Change	-0.12	-0.35
		Final value	-0.34	-0.50
	$\tilde{P}^{25\%}$(Dir)	Initial value	-0.13	*insig.*
		% Change	*insig.*	-0.11
		Final value	-0.18	-0.24
Overlap	$\overline{B}_{\text{overall}}$	Initial value	*insig.*	*insig.*
		% Change	0.68	0.73
		Final value	0.68	0.72
	$R(B_{\text{overall}})^{75\%}$	Initial value	-0.13	*insig.*
		% Change	0.19	*insig.*
		Final value	0.11	*insig.*
	$\overline{B}_{\text{elemental}}$	Initial value	*insig.*	*insig.*
		% Change	0.69	0.70
		Final value	0.72	0.71
	$R(B_{\text{elemental}})^{75\%}$	Initial value	-0.10	*insig.*
		% Change	0.17	*insig.*
		Final value	0.12	*insig.*
Efficiency step-change	$Rel\overline{\Delta L}(t;t+1)$	Minimum	*insig.*	-0.25
		Average	0.86	0.35
		Maximum	0.84	0.51

Redundancy

All the significant correlations of redundancy metrics to efficiency loss and robustness were negative. In the case of the initial and final values this implied that the greater the redundancy (i.e. the more shortest paths were available), the less the efficiency loss would

be. This seemed reasonable, until one recalled that smaller efficiency loss usually indicated that an instance became disconnected sooner rather than later. The negative correlations between the redundancy metrics and robustness confirmed that the more shortest paths were available to an instance, the sooner it would become disconnected (i.e. the weaker its robustness).

Even though it was only the relationship between the final values of $\tilde{P}(\text{All})$ and robustness that were *strongly* negative, the consistency of the negative correlations made this a perplexing result. It was expected that correlations to robustness would be positive — the more alternatives are available to a network, the longer it would survive random disruptions. To understand these observations better, we considered the correlations of the overlap metrics.

Overlap

In all three archetypes, the final values of and % change in $\overline{B}_{\text{overall}}$ showed a very strong correlation to both efficiency loss and robustness. Astonishingly, the correlation was positive, indicating that the more overlap developed in an instance, the more likely it was to survive longer. The positive correlation to the % change also implied that the more pronounced the change in overlap, the longer an instance survived. So the greater the overlap right before disconnection, the longer an instance survived. Or was it the other way around? Could it be that for the instances that survived longest there was more opportunity for shortest paths to be forced to overlap? Unfortunately, the correlation tests could not determine the direction of the causality. What was clear, was that there were diehard instances that had very high levels of overlap. This lead us to believe that greater overlap does not, in fact, suggest greater vulnerability! The expectation was for a negative correlation of overlap with efficiency loss and robustness, especially in terms of the initial values. The more overlap was prevalent in an instance, the quicker it was expected to become disconnected. As a result of quick disconnection it would then also suffer less efficiency loss.

The only plausible explanation for both the redundancy and overlap results lay in the fact that the removal of links were random. A perspective that had not been considered up until this point in the thesis was that the fewer grid links were included in the shortest paths of an instance, the less likely it was that a random selection would affect the shortest paths at all.

To investigate this further, the relationship between the grid coverage, redundancy and overlap was investigated. The grid coverage was measured as the number of grid links in $\mathcal{C}(\mathcal{S}_{ij})$ divided by 360 which was the total number of grid links.

There were significant strongly positive correlations between redundancy and grid coverage in each of the archetypes. Spearman's ρ was 0.54, 0.68 and 0.64 for the FC, SH and DH archetypes, respectively. This implied that the more alternative paths an instance had, the greater the proportion of all grid links covered by the shortest path sets. This also meant that the greater the redundancy, the more probable it was for a random disruption to damage the instance. This was confirmed by the significant correlation between the grid coverage and robustness which was -0.18, -0.24 and -0.15 for the FC, SH and DH archetypes, respectively. These negative correlations implied that the more of the grid was included in the shortest path sets, the quicker it would become disconnected. The quicker it became disconnected, the lower its overall efficiency loss. This was confirmed by the significant correlation of grid coverage with efficiency loss which was -0.18, -0.30 and -0.30 for the FC, SH and DH archetypes, respectively.

The take-away was that redundancy had a more feeble influence on robustness than first imagined. Rather, the probability of being affected at all by a random disruption held prominent bearing. This result corroborated the work of Dehghani et al. (2014) who also found that under random disruption the parameters of the disruption probability distribution were more indicative of road network vulnerability than topological characteristics.

The correlation between overlap and grid coverage was also significant, but negative. The degree of correlation was weaker than between redundancy and grid coverage. Spearman's ρ was -0.34, -0.21 and -0.42 for the FC, SH and DH archetypes, respectively. This relation illustrated that the more overlap there was within $\mathcal{C}(\mathcal{S}_{ij})$, the less of the grid was covered as the shortest paths were more concentrated. This result was congruent with the correlation of grid coverage to efficiency loss and disconnection, as mentioned earlier.

Certainly, as Chapter 4 showed, the degree of overlap is a critical vulnerability indicator under targeted attack. Under random link-disruption this theory frayed somewhat. Could it be that the probability of being affected by the upcoming disruption had more bearing on vulnerability than the level of overlap?

From these results one could have been tempted to assert that redundancy and overlap did not have bearing on vulnerability, but we cautioned against such statements. As mentioned before, it was recognised that road network failures are neither completely targeted, nor are they completely random. Rather the "disruption" mechanism is somewhere on the continuum between targeted and random. Although redundancy and overlap do not seem to capture vulnerability in its essence, neither does road utilisation — the correlations are too weak and causality was not established. What we could conclude was that vulnerability was a multi-dimensional concept.

Another pivotal observation was the weak (if not insignificant) relation of the initial values of redundancy and overlap to robustness. The initial state of the shortest paths was a very shaky foundation for quantifying vulnerability.

Moving away from the number of shortest paths and their degree of overlap, we evaluated the correlation of the changes in shortest path length to efficiency loss and robustness.

Efficiency Step-Change

The average and maximum efficiency step-change had significant, strongly positive correlations to both efficiency loss and robustness. The longer an instance survived, the sparser the road network became. A sparser road network foreshadowed larger step-changes in the length of shortest paths. Larger step-changes increased the average values and also made it more likely for a new maximum step-change to be observed.

The correlations of the efficiency step-change metrics made much more intuitive sense. However, the correlation itself did not prove that by monitoring these metrics one could discriminate between instances that were about to become disconnected and those that would survive.

Despite the controversial results observed for redundancy and overlap, we evaluated whether those metrics that were strongly correlated (either negatively or positively) could in some way have discriminated which instances were more likely to experience efficiency loss or disconnection during a subsequent disruption.

6.1.2 Discriminatory ability of correlated metrics

To investigate whether a certain metric had the ability to discriminate between instances that were about to experience efficiency loss or disconnection in a subsequent disruption, we had to ascertain whether there were significant differences between the values of the vulnerability metrics for those instances and the values for the rest. The vulnerability metrics that were tested were those that were proven to have strong correlations to efficiency loss or robustness. Table 6.4 lists these metrics for each archetype.

Table 6.4: Metrics tested for discriminatory power in each network archetype.

			Efficiency Loss			**Robustness**		
			FC	**SH**	**DH**	**FC**	**SH**	**DH**
Redundancy	$\tilde{P}$(All)	% Change			✓			
		Final value	✓			✓	✓	✓
	$\tilde{P}$(Dir)	Final value		✓	✓		✓	✓
Overlap	$\overline{B}_{\text{overall}}$	% Change	✓	✓	✓	✓	✓	✓
		Final value	✓	✓	✓	✓	✓	✓
	$\overline{B}_{\text{elemental}}$	% Change		✓	✓		✓	✓
		Final value		✓	✓		✓	✓
Efficiency step-change	$Rel\overline{\Delta L}(t;t+1)$	Average	✓	✓	✓	✓		
		Maximum	✓	✓	✓	✓	✓	✓

After each disruption, the instances were split into two samples: the sample of instances that were known to have become disconnected or to have suffered efficiency loss during the next disruption (sample X) and the sample of instances that were known to have survived or not to have suffered efficiency loss (sample Y). A Kolmogorov-Smirnov test (KS-test) was used to test the following hypothesis:

- H_0: The sample distributions of metric $v_i(X)$ and $v_i(Y)$ are drawn from the same theoretical distribution.
- H_A: The sample distributions of metric $v_i(X)$ and $v_i(Y)$ are not drawn from the same theoretical distribution.

(Where v_i is a vulnerability metric from Table 6.4.)

The test statistic (D) quantifies the distance between the Empirical Distribution Function (EDF) of the vulnerability metric as measured from the sample of surviving instances X and the EDF of the vulnerability metric as measured from the sample of non-surviving instances Y. The null hypothesis that the sample distributions were drawn from the same theoretical distribution was rejected when the p-value was lower than the chosen significance level, which in this case was 0.05.

The KS-test was conducted for each vulnerability metric after each disruption, provided that the samples X and Y had more than 15 observations each. The p-values of the

tests are tabulated in Tables B.1–B.5 in Appendix B. The results were overwhelmingly negative, failing to identify any generalisable discriminatory power in any of the metrics.

With regards to efficiency loss, the sample sizes of X in the FC archetype were less than 15 after each disruption, therefore no test could be performed. The DH archetype only had one out of 46 p-values smaller than 0.05. In the SH archetype the overlap metrics could discriminate instances about to suffer efficiency loss, but only at one time-point in the whole simulation, which made the result un-generalisable.

With regards to the robustness, there were many more occasions when the test rejected the null hypothesis and confirmed a level of discriminatory power. The step-change metrics showed discriminatory power in the earlier disruptions of the FC archetype while the overlap metrics were effective at unique time points for both the SH and DH archetypes. Nonetheless, these results were not consistent across the entire simulation for any one archetype with p-values measured at other time points variable and often greater than 0.5. This made it impossible to conclude that these metrics do really hold discriminatory power.

It is possible that the relatively small sample sizes and the inherent limitations of the KS-test affected the outcome. The KS-test asymptotically approximates p-values in the presence of ties, an approximation which could have led to inaccuracies. Furthermore, it is known to perform better with much larger sample sizes. Despite the effect of these shortcomings, it could be said that the lack of discriminatory power was congruent with the correlation results described in the previous section.

6.2 Conclusion of statistical validation

At this stage of the thesis we reverted to the design research methodology which stated:

> "Design itself is not considered research but it is through the insights derived during analysis of the designed artefact's performance that the body of knowledge in a field grows." (Section 1.6)

The hypothesis tests led to three insights. Firstly, although more than one metric was strongly correlated to efficiency loss and/or robustness for redundancy and overlap, the direction of the correlations were unexpected. This implied that vulnerability under random link disturbances was not a straight-forward product of redundancy and overlap. Instead it was multi-faceted and the probability of removing a link that featured in shortest path held pertinent influence. Secondly, the initial values of the vulnerability metrics were surprisingly uncorrelated. Therefore, it was impossible to gauge the inherent vulnerability of a network looking only at the initial, undisturbed network. Finally, by using the KS-test we confirmed that even those metrics that were strongly correlated to either efficiency loss or robustness unfortunately showed no discriminatory power. Overall we were satisfied that the metrics developed quantified the concepts of *redundancy*, *overlap* and *efficiency step-change*, however under random disruptions these were not convincing stand-alone indicators of vulnerability.

So far the thesis had presented a useful multilayered network representation and a thoughtfully crafted set of vulnerability metrics. Unfortunately, this suite of metrics did not prove to be the silver bullet hoped for in quantifying vulnerability under random disruptions. However, much was learnt from the performance of these metrics thus far. These artefacts were both novel when compared to related studies. Therefore, it was not possible to find comparable results from literature against which to assess the findings.

The final phase of this thesis sought to determine the validity of the formulation of $\mathcal{M}$ and the vulnerability metrics by applying them to real-life data from three urban areas in South Africa.

Chapter 7

Case study: Real-life networks from South Africa

Two artefacts have been developed in this thesis, a multilayered network formulation and a suite of vulnerability metrics. The theoretical validity of these two artefacts have been evaluated using targeted attack and random error simulations combined with hypothesis testing. The purpose of applying these artefacts to real-life networks was to evaluate whether the theoretical findings could be corroborated when real-life data is used. The following two chapters address the second part of the fourth thesis objective:

> 4. Evaluation of the validity of the suite of vulnerability metrics through statistical analysis and a **real-life case study**.

The case study explored real-life data from three urban areas in South Africa namely the City of Cape Town (CoCT) metropolitan municipality in the Western Cape province, the eThekwini Metropolitan Municipality (ET) in the KwaZulu-Natal province and the entire Gauteng Province (GT). The locations of the three areas, in the context of South Africa, are shown in Figure 7.1. Using road network and freight movement data from these three *areas* we were able to extract unique multilayered network instances (hereafter *case study instances*) that approximated the three theoretical network archetypes, Fully Connected (FC), Single Hub (SH) and Double Hub (DH).

This chapter starts by contextualising the three chosen areas. The methodology followed in extracting the case study instances is then presented and the characteristics of these instances are compared to the characteristics of the theoretical instances used in the thesis up until now. Finally, the customised algorithm developed to construct the shortest path sets is explained before concluding the chapter with an analysis of the initial shortest path sets of the case study instances.

7.1 Three urban areas

The CoCT in the Western Cape Province and ET in the KwaZulu-Natal Province are two metropolitan municipalities anchored by two of Southern Africa's most prominent seaports. The Port of Durban in ET is one of the busiest container ports in Africa and also handles significant volumes of dry bulk and liquid bulk cargo. The Port of Cape Town in CoCT handles mainly fruit & agricultural dry bulk trade and containerised cargo. Over decades these ports have been pivotal in cultivating urban and industrial development in their immediate hinterland.

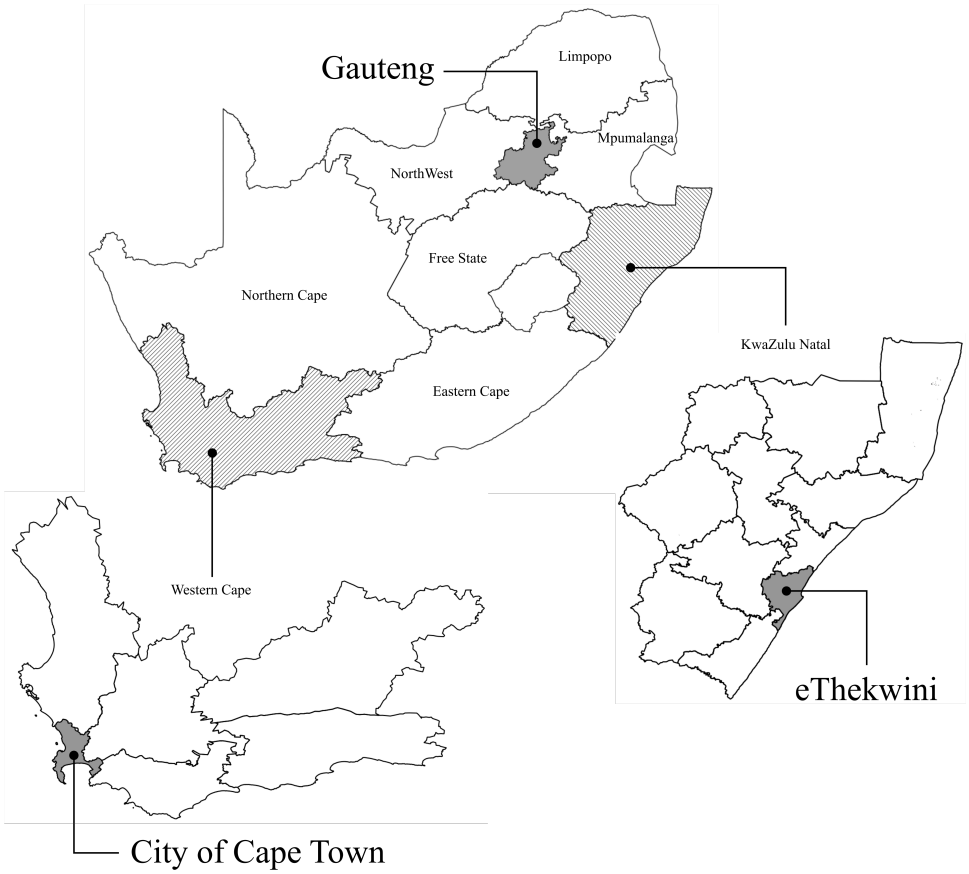


Figure 7.1: The CoCT and ET metropolitan municipalities and GT province in context of the rest of South Africa.

Figure 7.1 shows GT, Western Cape and KwaZulu-Natal provinces in context of the rest of South Africa. CoCT lies on the south-western edge of of the Western Cape province while ET lies on the eastern coast of the country towards the southern edge of KwaZulu-Natal. Geographic constraints, industrial activity and the legacy of urban development under the Apartheid regime resulted in distinct urban road network topologies for each of these areas.

The GT province is the nexus of Southern Africa's industrial and economic activity. Situated hundreds of kilometres inland from any port, its industrial heritage dates back more than a century to the discovery of gold. Today it is still the economic powerhouse of the country and even the Southern African region with industries diversified across the primary (mining & agriculture), secondary (manufacturing), and tertiary (services) sectors. It contributes a third of the country's GDP and 10% of the GDP of the entire African continent (Gauteng Online, 2017). Geographically it is the smallest province, yet it is also the most densely populated in the country with a population exceeding 13 million (Statistics South Africa, 2016). Therefore, although GT consists of five separate municipalities, it is considered a *megacity* according to international classifications (United Nations, Department of Economic and Social Affairs, Population Division, 2015). Three of its five municipalities are classified as metropolitan. The province is a dense web of interconnected urban centres, as such it did not make sense to isolate one of the five metropolitan municipalities for the case study. Instead the whole province was included.

To extract case study instances from any one of these areas required that we first

extract the logical (G^{1K}) layers of instances from freight movement data and the physical (G^2) layers from road network data. Thereafter each instance of G^{1K} had to be layered on G^2 to create a multilayered network instance $\mathcal{M}$ with its associated collection of shortest paths $\mathcal{C}(\mathcal{S}_{ij})$.

7.2 Constructing the logical layer

Joubert and Axhausen (2011) were the first to consider commercial vehicle movement as a proxy for supply chain activity. Using the Global Positioning System (GPS) vehicle logs of 40 000+ commercial vehicles over six months the authors developed a methodology to extract commercial vehicle activity chains from the data. These chains indicate when and where a vehicle performed logistics activities. The authors realised that the concentration of logistics activities in a certain location is a clue in identifying logistics facilities. In subsequent work, they used clustering algorithms to infer the positions of logistics facilities based on the activity chains extracted from the commercial vehicle movement database (Joubert and Axhausen, 2013; Joubert and Meintjes, 2015a,b). For almost a decade Joubert and collaborators have been refining the methodology of extracting activity chains and using these to identify logistics facilities in South Africa. The logical layers of the case study instances were extracted from this database of commercial vehicle activity chains.

7.2.1 Creating supply chain networks for the three areas

All activity chains with one or more activities executed inside the areas during February 2014 were extracted from the database. These chains were used to construct one large logical network layer for each area. These were called the *area networks*, each being a combination of hundreds of small supply chain neighbourhoods or "building blocks" that made up the freight economies in the areas.

These area networks were both directed and weighted. Nodes represented logistics facilities while links indicated that there had been commercial vehicle activity, i.e. direct trips between two facilities in the direction of the links. The weight of the links represented the number of times a commercial vehicle had travelled between two nodes during February 2014. It was assumed that for supply chain interactions to be frequent and ongoing, there had to be a minimum of four vehicle trips between facilities in a month — equating roughly to one vehicle trip per week. Therefore, the area networks were filtered to only include links with a weight of at least four trips per month. The dimensions of these area networks are summarised in Table 7.1.

Table 7.1: Dimensions of the supply chain area networks.

Area	**Nodes**	**Links**	**Links:Node**
GT	3 424	6 222	1.82
CoCT	1 440	3 308	2.30
ET	1 060	1 965	1.85

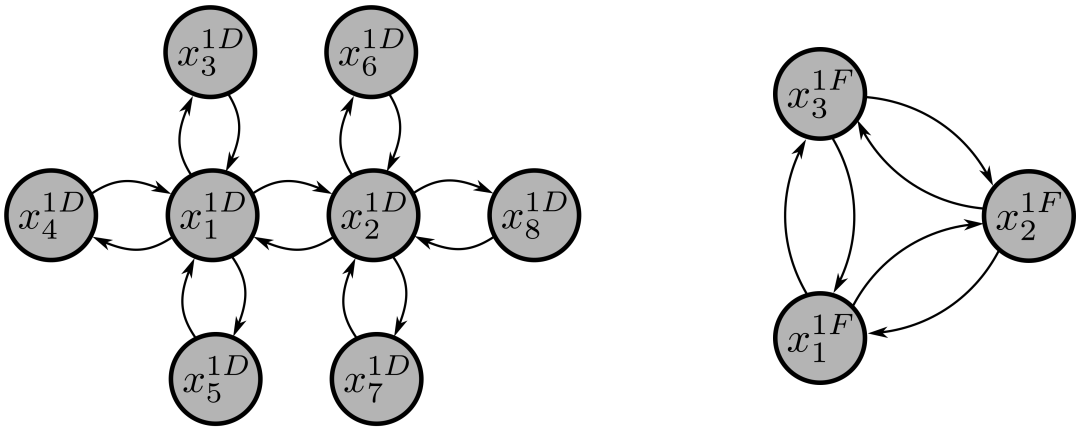
Although the area networks were weighted initially, the formulation of $\mathcal{M}$ was unweighted. Furthermore, the vulnerability metrics were based on unweighted networks. Therefore, the link weights were only used to filter the noise resulting from the comprehensive activity chain database, thereafter the area networks were modelled as unweighted.

Next, individual case study instances that mimicked the structure of the FC, SH and DH archetypes had to be extracted from these directed, unweighted area networks.

7.2.2 Extracting potential case study instances

Node degree, the First Order Neighbourhood (FON) and triangles were three Complex Network Theory (CNT) concepts used to identify and extract the logical layers of potential case study instances. These concepts are explained here with reference to the DH and FC archetype examples illustrated in Figure 7.2.

Node degree, in its simplest form, is the sum of the number of incoming and outgoing links that connect a node to other nodes in the network. For example, from Figure 7.2a we see the node degree of $x_1^{1D} = 8$, which is the sum of its four incoming and four outgoing nodes. Meanwhile the node degree of $x_6^{1D} = 2$. In the FC example of Figure 7.2b, the node degrees are all equal with $x_1^{1F} = x_2^{1F} = x_3^{1F} = 4$.



(a) Theoretical DH archetype with eight nodes, denoted by $N_1^D = 8$.

(b) Theoretical FC archetype with $N_1^F = 3$.

Figure 7.2: Illustrative DH and FC archetypes used to explain node degree, the FON and triangles.

The FON of a node includes itself and all the nodes that are directly connected to that node. The FON of x_1^{1D} is thus the set of nodes $\{x_1^{1D}, x_2^{1D}, x_3^{1D}, x_4^{1D}, x_5^{1D}\}$ while the FON of x_6^{1D} includes only two nodes $\{x_2^{1D}, x_6^{1D}\}$. While the FONs of x_1^{1D} and x_6^{1D} have a node in common (x_2^{1D}), they are not identical. The FONs of two nodes are only considered identical if they are comprised of exactly the same set of nodes. By contrast, the FONs of all the nodes in the FC archetype are identical — i.e. FON(x_1^{1F})= $\{x_1^{1F}, x_2^{1F}, x_3^{1F}\}$, FON($x_2^{1F}$)= $\{x_1^{1F}, x_2^{1F}, x_3^{1F}\}$, and FON($x_3^{1F}$)= $\{x_1^{1F}, x_2^{1F}, x_3^{1F}\}$.

A triangle occurs when two first order neighbours of a certain node are also connected to each other. In the FC archetype x_2^{1F} and x_3^{1F} are first order neighbours of x_1^{1F}. They are also directly connected to each other. Thus, the three nodes form a triangle. By contrast, there are no triangles in the DH network.

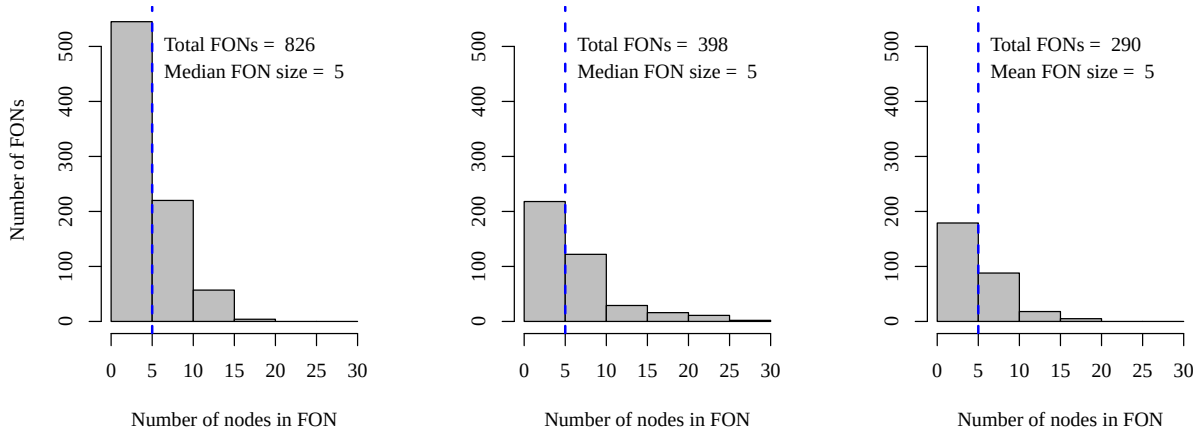
Table 7.2 summarises distinguishing characteristics of the theoretical FC, SH and DH archetypes in terms of node degree, FON and triangles.

A census of all FONs was conducted for each area network. A decision was made to filter out all FONs containing three or less nodes as a supply chain with three or less nodes was not considered a true reflection of supply chain complexity in practice. The

Table 7.2: Distinguishing characteristics of the FC, SH and DH archetypes.

	Archetype		
Characteristic	**FC**	**SH**	**DH**
Node degree	Node degree is equal to $2(n-1)$ for all nodes.	One hub node with relatively high node degree. All other nodes have a degree of 2.	Two hub nodes with relatively high node degree. All other nodes have a degree of 2.
FONs	All nodes have identical FONs, i.e. they are within the same FON.	All nodes are included in the FON of the hub node, while the FONs of the other nodes include only itself and the hub node.	All nodes are included in the FONs of the two hub nodes, but the FONs of the two hubs are not identical. The FON of every other node includes only itself and one of the hub nodes.
Triangles	All the possible triangles are complete.	No triangles present.	No triangles present.

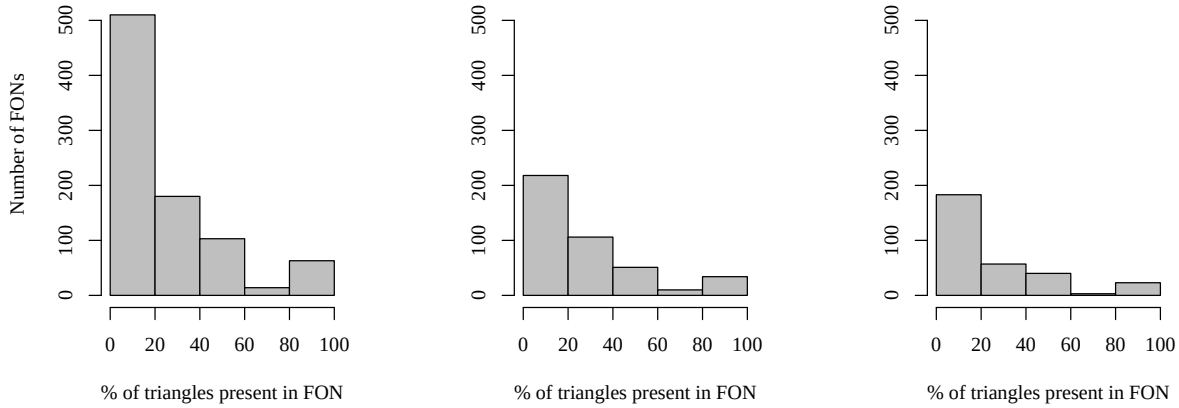
number of nodes in a FON is its size. The distributions of FON sizes were heavily skewed with long right tails for each of the area networks. For legibility's sake, the distributions of FON sizes are only shown up until the 95^{th} percentile in Figure 7.3. Although there were many more FONs in GT than in the other two areas, the median of the FON size was five for each of the areas. This implied that despite their differences, the building blocks of supply chain networks were similar across these three areas — at least in terms of the number of participating facilities.



(a) GT network (b) CoCT network (c) ET network

Figure 7.3: Distribution of FON sizes for the three area networks. Only FONs with more than three nodes that fell within the 95$^{\text{th}}$ percentile cut-off are shown here.

Next it was calculated what percentage of the possible triangles[1] was present in each FON[2] (Figure 7.4). FONs that contained 100% of all possible triangles were potential FC instances. On the other hand, FONs that contained zero triangles were potential SH and DH instances.



Triangles	No. of FONs
0%	163
100%	63

(a) GT network

Triangles	No. of FONs
0%	51
100%	34

(b) CoCT network

Triangles	No. of FONs
0%	61
100%	23

(c) ET network

Figure 7.4: Distribution of possible triangles present in FONs.

Across all three areas there were a total of 120 FONs containing 100% of possible triangles and 275 FONs containing zero triangles, yielding nearly 400 potential case study

[1]The *possible triangles* of a network with n nodes refers to the maximum number of triangles that can be formed if that network is fully connected.

[2]The function in R's `igraph` package (Csardi and Nepusz, 2006) that identifies triangles regards all links as undirected, regardless of whether the network is specified as directed or not.

instances. However, there was considerable similarity between many of these FONs. Similarity is when two FONs have a significant proportion of nodes in common. Comparisons were conducted and if two FONs had more than 50% of their node sets in common, the smaller of the two FONs was discarded. Table 7.3 shows the samples of FONs that remained in each of the areas after filtering according to triangles and removing similar FONs. All the FONs with 100% triangles represented FC instances while all FONs with

Table 7.3: FONs remaining per area after filtering according to triangles and similarity.

	GT	**CoCT**	**ET**	**Total**
0% triangles	136	41	48	225
100% triangles	36	19	15	70

0% triangles represented SH instances. Identifying DH instances was slightly more tricky.

The DH archetype is essentially a combination of two adjacent SH instances. Consider again the example network in Figure 7.2a. The FON of node x_1^{1D} includes node x_2^{1D} but not nodes x_6^{1D}, x_7^{1D} or x_8^{1D}. Conversely, the FON of node x_2^{1D} contains node x_1^{1D} but not nodes x_3^{1D}, x_4^{1D} or x_5^{1D}. The first clue to identifying adjacent SH instances was thus to search for FONs that had more than one high degree node. We did this by searching through all the 0% triangle FONs for those that contained a secondary node with degree 80% or higher than that of the focal node. Once such FONs had been identified, it had to be determined whether the secondary node was, in fact, the hub of another SH instance. If that was the case, the two SH instances were combined to form one DH instance. Altogether 20 DH instances could be identified.

The case study instances had to be kept as independent as possible. Therefore the 36 SH instances that were combined in different configurations to form DH instances were removed from the original pool of 225 SH instances. Table 7.4 lists the sample sizes of each archetype in each area. This represented the initial sample of case study instances.

Table 7.4: FC, SH and DH instances per area

	GT	**CoCT**	**ET**	**Total**
FC	36	19	15	70
	52%	27%	21%	100%
SH	110	37	42	189
	58%	19%	22%	100%
DH	15	2	3	20
	75%	10%	15%	100%

The question arose, how well does this sample represent the population of supply chain neighbourhoods in the three areas?

7.2.3 Representativity of the case study instances

By applying the aforementioned criteria to extract case study instances we had whittled away substantially at the area networks. What percentage of the original area networks

were represented by this sample of case study instances? An average of 24% of the nodes and 17% of the links of each area were included in the case study instances (Figure 7.5).

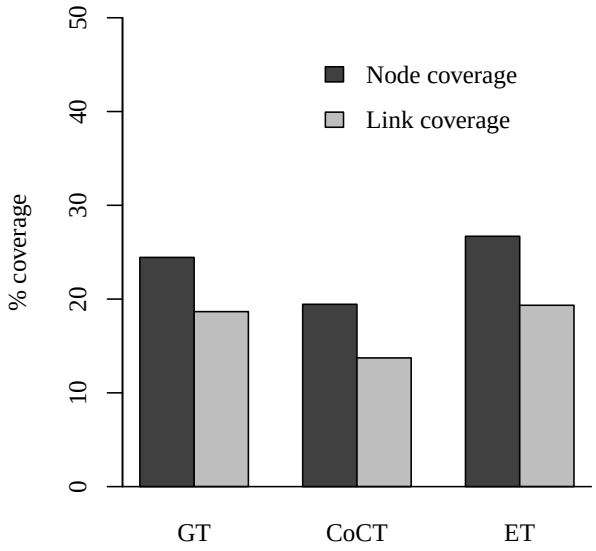


Figure 7.5: Percentage of the nodes and links of the original area networks included in the case study instances.

It was tempting to conclude then that the case study instances were significantly representative. However, the sample had a very distinct bias towards instances with specific characteristics — namely those that had either 100% or 0% of the theoretical triangles present (refer back to Figure 7.4). These FONs were regarded as closely resembling the FC, SH or DH archetypes. But what about all the other FONs? These were supply chain neighbourhoods that didn't have a clearly predefined structure. If we considered the FONs shown in Figure 7.4, then these networks of *mixed type* constituted 73%, 79% and 71% of the FONs in GT, CoCT and ET, respectively. It was notable that real-life supply chains did not obey the neat confines of theoretical archetypes. It was also notable that the percentage of supply chains that were of mixed type were similar across the three areas despite the perceived differences in the economic activity of these areas.

We could thus conclude that the case study instances represented between a fifth and a quarter of the FONs in the area networks whether considering only the number of nodes/links included or the FONs themselves. However, it completely disregarded FONs of mixed type, which constituted about three quarters of each area. This bias had to be kept in mind when interpreting findings later in the study.

The criteria used to extract case study instances ensured that these closely resembled the three archetypes. There remained, however, scope for the instances to deviate from the assumptions that had been applied when generating the 1 500 theoretical instances.

7.2.4 Deviations from theoretical assumptions

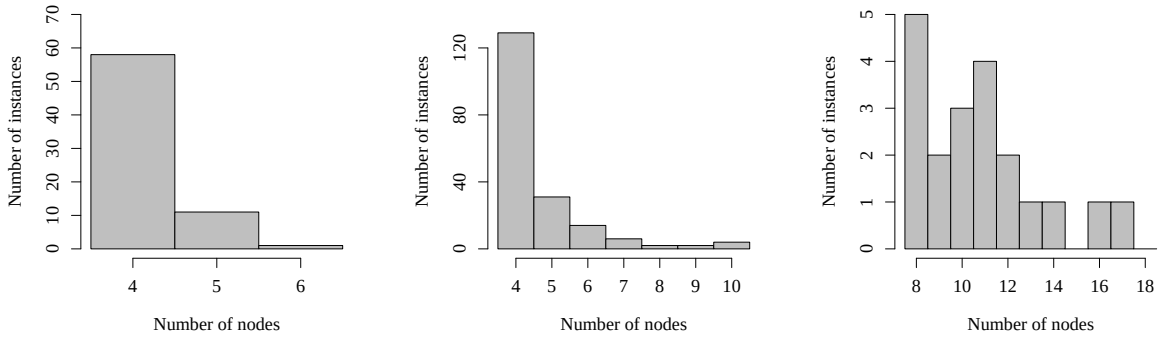
As could be expected, not all of the assumptions made when generating G^{1K} in the theoretical instances held true for the case study instances. Three deviations were identified:

- The number of nodes (N^{1K}) very seldom equalled 12 in the case study instances.

- Not all directly connected nodes were connected bi-directionally (i.e. one link in each direction).
- Not all DH instances strictly adhered to the two-hub topology.

Number of nodes (N^{1K})

Figure 7.6 shows the distribution of N^{1K} for each of the archetypes. The FC and SH instances were all smaller than their theoretical counterparts and ranged between 4 and 10 nodes. The DH instances were closer to $N^{1K} = 12$ with sizes varying between 8 and 18 nodes.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 7.6: Number of nodes in G^{1K} for the case study instances.

Smaller networks meant fewer links that had to be maintained during successive disruptions. Surely it was plausible then that smaller networks would be more robust under a random error simulation as there were fewer possible "points of failure"? Here the multilayered nature of the case study instances came into play. Although a supply chain may have contained only a few facilities, these may have been many kilometres apart, traversing a significant portion of the underlying road network. This distance introduced many more possible points of failure in the road network. Thus, one could not have concluded that instances would be less vulnerable to disruption based on the size of G^{1K} alone.

What remained true, was that fewer nodes in G^{1K} implied fewer logical links. The next two deviations also affected the number of logical links in the case study instances.

Bi-directionality

The second deviation from the theoretical assumptions was that in the logical network layers, not all node-pairs were connected bi-directionally. In the theoretical instances, every link e_{ij}^{1K} from node x_i^{1K} to node x_j^{1K} had a reciprocal link e_{ji}^{1K} from node x_j^{1K} to node x_i^{1K}. In the case study instances facilities were often *not* connected bi-directionally. It made sense that in practice freight may typically be shipped only in one direction between facilities — for example from a manufacturing facility to a warehouse. Therefore, it was not surprising that many node-pairs only had a one-directional connection.

The link structures of the logical layers were left unaltered. Where nodes were only connected in one direction, a reciprocal link was not inserted artificially. In the SH and DH archetypes this simply meant that some nodes could either only receive incoming

freight or send outgoing freight. The impact was a reduction in the total number of links, both direct and indirect. The effect in the FC archetype was somewhat different. In the theoretical instances, all nodes in G^{1F} were directly connected. The number of direct links was always $N^{1F}(N^{1F}-1)$ and the number of indirect links always zero. The absence of bi-directional links induced indirect connections between nodes that would not have been necessary before. Figure 7.7 illustrates this phenomenon. The network on the left has a direct link e_{31}^{1F} from x_3^{1F} to x_1^{1F}. When e_{31}^{1F} is not present, as in the network on the right, the path from x_3^{1F} to x_1^{1F} is indirect $x_3^{1F} \rightarrow x_2^{1F} \rightarrow x_1^{1F}$. So while this deviation also reduced the number of direct links in an FC instance, it induced indirect connections that were not present before.

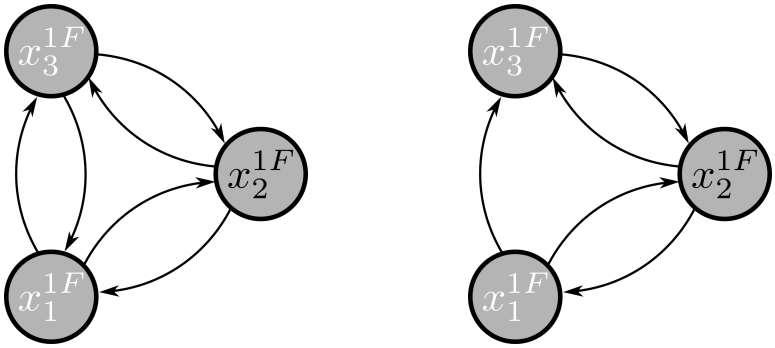
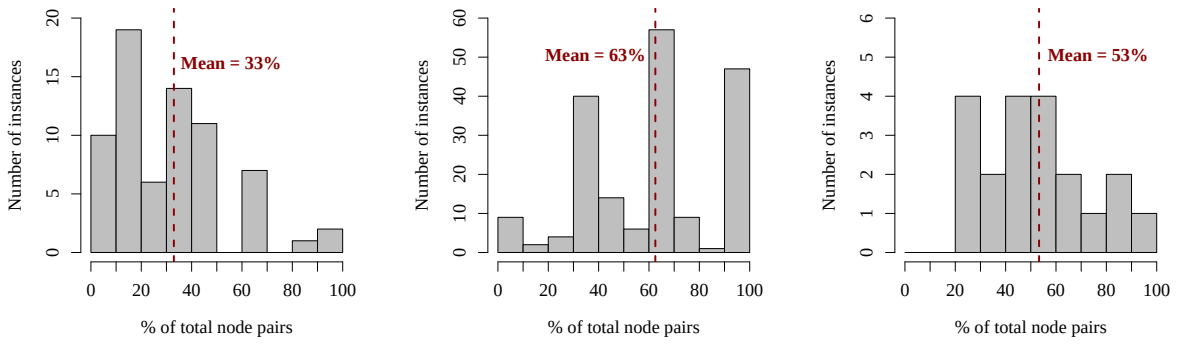


Figure 7.7: Example of the occurrence of indirect links in an FC instance. On the left x_3^{1F} and x_1^{1F} are directly connected by e_{31}^{1F}. On the right e_{31}^{1F} is not present thus inducing an indirect connection between x_3^{1F} and x_1^{1F}: $x_3^{1F} \rightarrow x_2^{1F} \rightarrow x_1^{1F}$.

This deviation was pervasive. On average 33% of the node-pairs in an FC instance were only connected in one direction. Even more node-pairs were affected in the hub archetypes. Figure 7.8 shows that the SH and DH instances had an average of 63% and 53% of the node-pairs connected uni-directionally. In all three archetypes there were multiple instances with 90%–100% of the node-pairs connected uni-directionally. In fact, in the SH archetype this was the second most prevalent bandwidth.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 7.8: Uni-directional node-pairs as a % of the total number of node-pairs per instance.

Having node-pairs connected only in one direction did not change the levels of damage that could be measured during a disruption simulation. Efficiency loss was still measured

as the increase in the average shortest path over all logical links; an instance was still considered disconnected when any single logical link was completely broken; and a network was still considered destroyed when no more logical links remained.

The net result was less densely connected logical layers, indicated by a lower link to node ratio for all three archetypes. For the FC archetype the occurrence of indirect connections affected the homogeneity of all of its network metrics, making their distributions broader and more similar to those of the hub archetypes than those of the theoretical instances. Section 7.5.1 and Chapter 8 illustrate these observations.

Prohibited links in the DH archetype

The third deviation from the theoretical assumptions related to the structure of the DH archetype. Strictly speaking, the DH archetype had two hub nodes, each surrounded by a number of spoke nodes. Each spoke node was connected only to its respective hub and the two hubs were connected to each other as illustrated in the network on the left in Figure 7.9. In the case study, DH instances were created by joining two adjacent SH instances. Sometimes a spoke node in one SH instance was connected to a spoke node in the other. Thus, when the two SH instances were combined, it resulted in a DH instance that had direct connections between spoke nodes as shown in the righthand network in Figure 7.9.

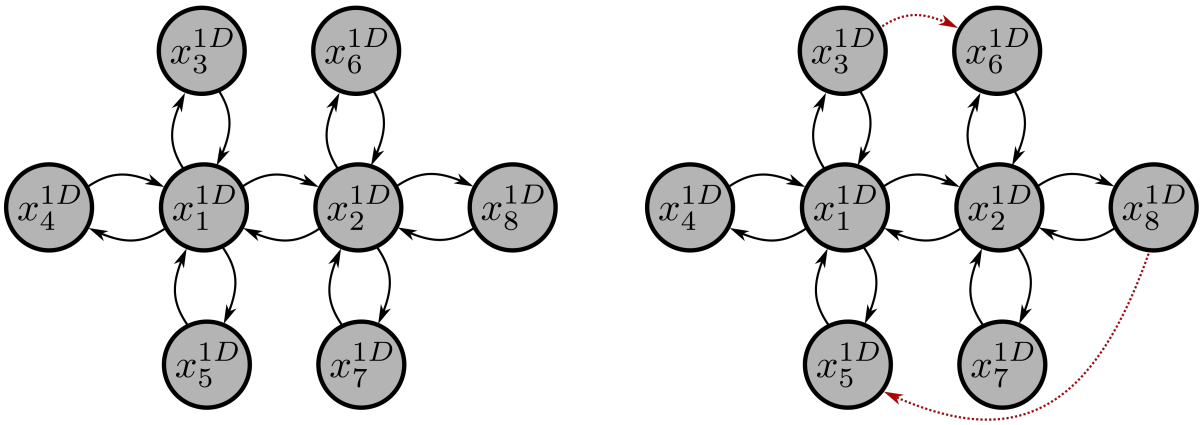


Figure 7.9: Example of direct links between the spoke nodes of a DH instance. In the instance on the left there would be no direct links between spoke nodes. In the case study, combining two adjacent SH instances sometimes resulted in the direct connection of spoke nodes for example $x_3^{1D} \rightarrow x_6^{1D}$ and $x_8^{1D} \rightarrow x_5^{1D}$ (right).

Compared to the previous two deviations, this deviation was less prevalent as it only affected 20% of the instances in one of the three archetypes. In these instances the number of logical links increased by 17%, on average.

One experimental implication of both the second and third deviations was that the number of links in the logical layer could not be calculated by virtue of the network archetype and number of nodes. Instead it had to be empirically determined for each instance.

Net effect on the number of logical links

If all node-pairs had bi-directional links and there were no prohibited links occurring in the DH instances, the number of logical links could be calculated based on the archetype

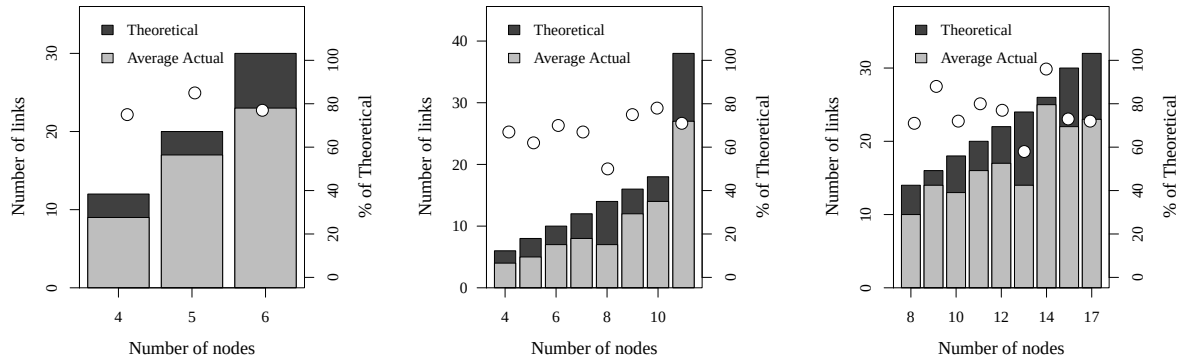
and number of nodes as shown in the equations below.

$$\text{FC: } N^{1F}(N^{1F}-1) \tag{7.1}$$

$$\text{SH: } 2(N^{1S}-1) \tag{7.2}$$

$$\text{DH: } 4(\frac{N^{1D}}{2}-1)+2 \tag{7.3}$$

The darker bars in Figure 7.10 indicate the theoretical number of logical links that should have been present for instances of a specific archetype and size if there were no deviations. The lighter bars indicate the average of the actual number of logical links present in instances of that size. For all three archetypes the actual number of logical links was less than the theoretical number of links. This confirms that, across the board, the logical layers of the case study instances were less densely connected than the theoretical instances.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 7.10: Theoretical number of links based on the number of nodes compared to the average of the actual links in each case study instance containing that number of nodes.

7.2.5 Validity of the formulation of the logical layer (G^{1K})

In this thesis three archetypes were thoughtfully formulated based on prevalent supply chain design theory and intuitive knowledge of supply chain network interactions. A number of theoretical assumptions were made when generating samples of each of these archetypes. This section investigated the validity of this formulation based on the real-life data available in the three case study areas.

The first observation was that these archetypes, or at least close approximations of these archetypes, were indeed present in the practice and captured between a fifth and a quarter of the total population of supply chain neighbourhoods. The remainder of the supply chain neighbourhoods were mixtures of theoretical archetypes.

The prevalence of the *mixed type* in practice can be discussed from two perspectives. The first perspective acknowledges that reality simply does not pan out as we plan. The partners in a supply chain may agree to design their network according to certain philosophies. These may be centralised distribution (approximating the SH archetype), decentralised distribution (multi-hub structures similar to the DH archetype) or fully collaborative distribution (similar to the FC archetype). Unfortunately, when the rubber

hits the road it is unlikely that the interactions between the facilities, at least in terms of freight movement, will obey these academic designs. The second perspective posits that supply chain network designs are really emergent and innovative — impervious to the enforcement of theoretical archetypes.

This is one of the key opportunities for future research identified by this thesis. The methodologies now exist to extract the building blocks of supply chain networks from big data. It is worthwhile to investigate the prevalent supply chain designs that emerge from practice and how these interconnect. A related question would be whether these designs are correlated to certain geographies, countries or economic structures are indifferent to these factors and therefore generalisable.

Despite the fact that the theoretical archetypes addressed only a portion of the population, that portion remained significant. Any findings based on these archetypes would be relevant to at least a fifth of the population.

The second observation was that the case study instances did not adhere to the theoretical assumptions in three specific ways. These deviations affected the size and densities of the instances. It was speculated that such differences could have influenced the performance of the suite of vulnerability metrics but it was impossible to predict what this influence would have been based on the characteristics of the logical layers alone. Next we present the methodology used to construct the physical layers (G^2) corresponding to each logical layer and the comparison of the real-life characteristics of these road networks to the theoretical representation of the bi-directional grid.

7.3 Constructing the physical layer

The ubiquity of GPS technology has led to the success of Volunteer Geographic Information (VGI) in the past decade. Non-experts and experts alike can now contribute to open source geographic information platforms. One such viable platform is OpenStreetMap (OSM). It is dedicated to creating and maintaining high-quality free geographic data. The quality of VGI can vary based on the number of contributors and the rigour enforced by the community. However, the quality and completeness of VGI in urban areas around the world is typically comparable to commercial geographic data (Graser et al., 2013). The OSM data for road networks in the three areas have been used in a number of studies by the Centre for Transport Development at the University of Pretoria. Therefore, it was assumed that the quality and completeness of this data were adequate for the purpose of this thesis.

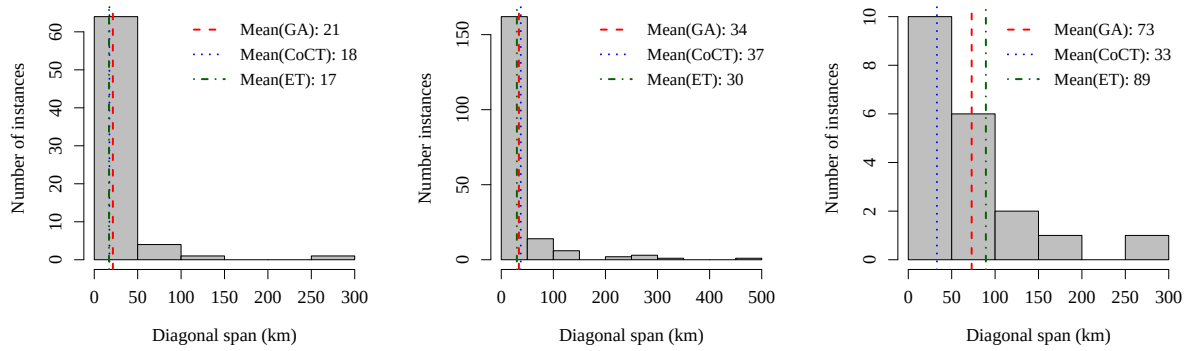
7.3.1 Extracting road networks from OpenStreetMap

The road networks within the areas were extracted from OSM using *Osmosis*, a Java application developed specifically for processing OSM data. During data extraction one had the option to exclude certain road types. In the commercial vehicle activity dataset there was no discrimination of vehicle type, thus we assumed that the vehicles varied from light delivery vehicles to the largest semi-trailer configurations. We further assumed that most of these vehicles would have used highways, primary and secondary roads and other major paved roads in an urban area. Suburban, private and dirt roads would have been avoided as much as possible. Therefore, when extracting the road networks for the three areas, filters were applied so that only major roadways were extracted.

7.3.2 Clipping network sections for case study instances

For each case study instance, a section of the road network had to be clipped from the larger area network. This was done to reduce the computational burden of conducting simulations on the case study instances. From Figure 7.11 it is evident that the diagonal span (in km) of the logical layers varied greatly across the different archetypes. FC instances covered the smallest area, SH instances covered, on average, approximately double the area of the FC instances. The DH instances in turn covered nearly double the area of the SH instances in GT and ET, but were smaller on average in CoCT.

In practice the scope for finding detours between two facilities is far greater than on the theoretical 10×10 grid. Nonetheless, there is a natural limit to the length of a detour that would make sense when serving a specific supply chain. A detour of 15km may be acceptable to a driver serving a DH instance with facilities strewn across the entire GT, but that same 15km detour would be a deal-breaker for a driver whose original route was less than 20km serving facilities in an FC instance.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 7.11: Diagonal span of the logical layers of the case study instances.

Clipping smaller sections from the area networks held two advantages. Firstly, as simulations destroyed the road network new shortest paths were limited to a smaller portion of the road network. This discouraged the routing of unrealistic shortest paths to keep facilities connected. Secondly, the size of the physical network layer G^2 built from the road network was greatly reduced — easing computation. Algorithm 4 explains how sections were clipped from the metropolitan network. In short, the distance between any node in X^{1K} and the edge of G^2 had to be 25km or 20% of the latitude/longitude range of X^{1K} — whichever was greater.

7.3.3 Deviation from theoretical assumptions

Although using a bi-directional grid to represent urban road networks is acceptable practice in similar studies (Ortigosa and Menendez, 2014), it was obvious that the physical layers of the case study instances would deviate from this simplistic model.

The theoretical 10×10 grid had 100 nodes representing road intersections. Each node was connected to its adjacent node(s) by a pair of directed road segments (links) as shown in Figure 7.12. The lengths of the links were uniform.

There were three features to inspect in the case study instances. The first was how close these physical layers were to a grid-like structure. The second was how uniform

Algorithm 4: Clipping a section from a area road network to create G^2 for a case study instance

Input : Area road network, X^{1K}
Output: G^2

$minLat \leftarrow$ minimum latitude coordinate in X^{1K};
$maxLat \leftarrow$ maximum latitude coordinate in X^{1K};
$minLon \leftarrow$ minimum longitude coordinate in X^{1K};
$maxLon \leftarrow$ maximum longitude coordinate in X^{1K};
$bufPercent \leftarrow$ buffer percentage
//Determine minimum and maximum latitude of bounding box;
$tempMinLat \leftarrow minLat + (maxLat - minLat) * bufPercent$;
// Latitude signs specific to road networks in the Southern Hemisphere;
$d_{minLat} \leftarrow$ Haversine distance between $tempMinLat$ and $minLat$;
if $d_{minLat} \geq 25km$ **then**
    $bbMinLat \leftarrow tempMinLat$;
    $bbMaxLat \leftarrow maxLat - (maxLat - minLat) * bufPercent$;
else
    $bbMinLat \leftarrow$ latitude that is 25km South of $minLat$;
    $bbMaxLat \leftarrow$ latitude that is 25km North of $maxLat$;
// Determine minimum and maximum longitude of bounding box;
$tempMinLon \leftarrow minLon - (maxLon - minLon) * bufPercent$;
// Longitude signs specific to road networks in the Eastern Hemisphere;
$d_{minLon} \leftarrow$ Haversine distance between $tempMinLon$ and $minLon$;
if $d_{minLon} \geq 25km$ **then**
    $bbMinLon \leftarrow tempMinLon$;
    $bbMaxLon \leftarrow maxLon + (maxLon - minLon) * bufPercent$;
else
    $bbMinLon \leftarrow$ longitude that is 25km West of $minLon$;
    $bbMaxLon \leftarrow$ longitude that is 25km East of $maxLon$;
$X^2 \leftarrow$ all nodes with coordinates within $(bbMinLat, bbMaxLat; bbMinLon, bbMaxLon)$;
$E^2 \leftarrow$ all links with both incident nodes in X^2;
$G^2 = (X^2, E^2)$
return G^2

the link lengths were. The third feature was the density of these layers compared to the theoretical grid.

Regular grid structure

The degree distribution of a bi-directional grid is relatively homogenous. The 4 corner nodes will always have the lowest degree of 4 (2 incoming and 2 outgoing links). The nodes on the top, bottom, left and right edges will have a degree of 6 and all other nodes will have a degree of 8. Changing the size of the grid will only increase/decrease the number of nodes that have a degree of 6 and 8. As the degree distribution of the regular grid is relatively stable for any number of nodes, it was used as a point of comparison to assess the grid-likeness of the case study instances.

Figure 7.13a shows the box-and-whisker plot for the degree distribution of the 10 × 10 grid. The box-and-whisker plots for the degree distributions of the physical layers clipped from the GT network are shown on the same scale in Figure 7.13b. The degree distributions of the physical layers were remarkably similar to each other with a minimum

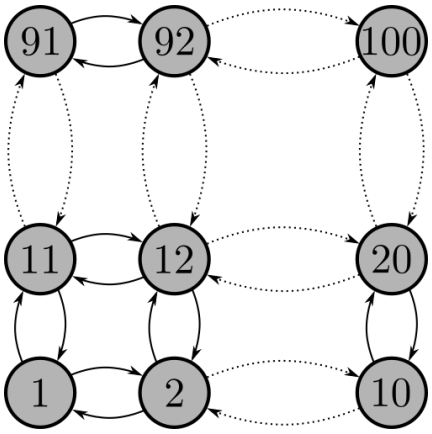
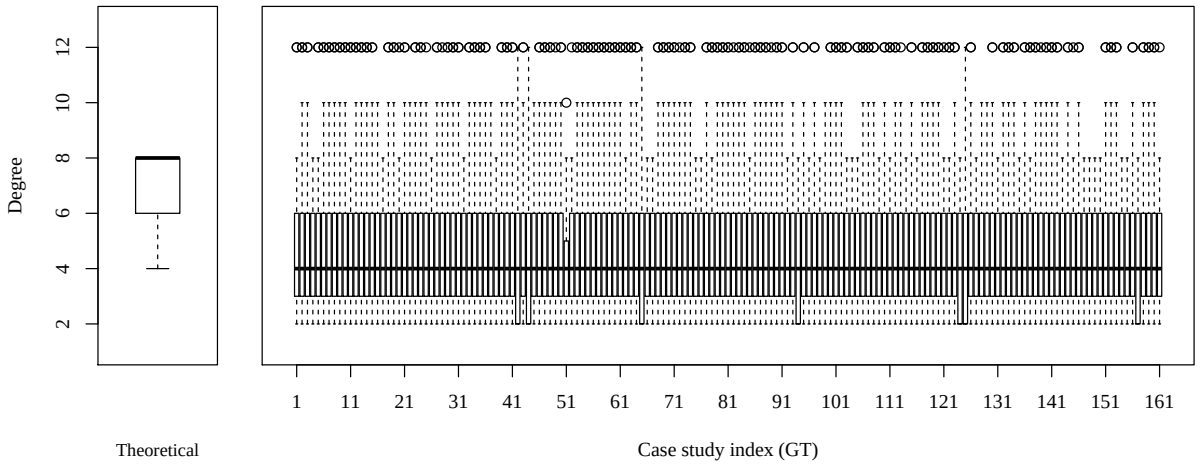


Figure 7.12: G^2 — the 10×10 directed, unweighted representation of the road network.

value of 2, a mean of 4 and a maximum of 12. The bulk of the nodes had degrees between 3 and 6. Although the case study instances did not mimic the degree distribution of the theoretical grid per se, they certainly exhibited a measure of regularity.

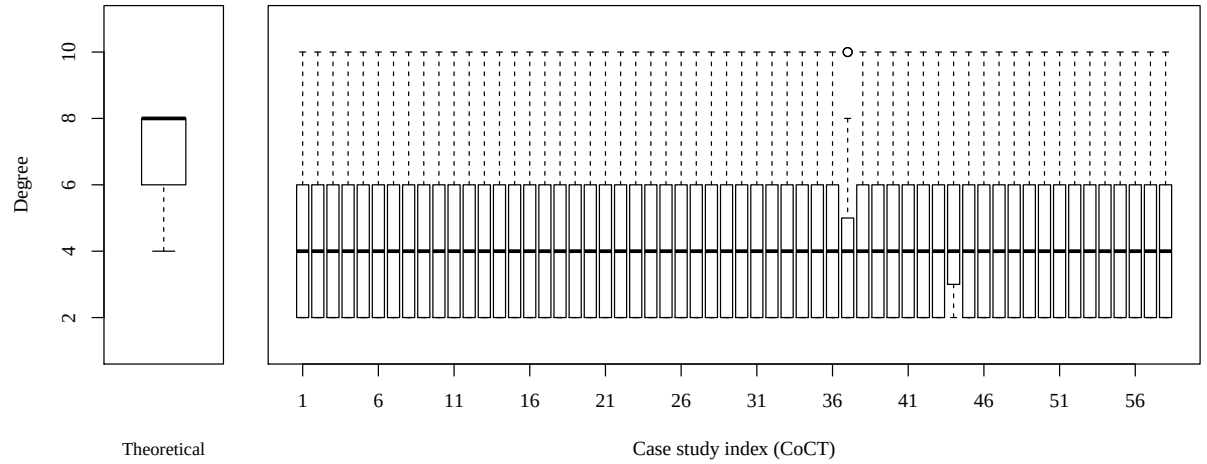


(a) Degree distribution of the theoretical grid.

(b) Degree distributions of the physical layer of the case study instances in GT.

Figure 7.13: Degree distributions of the physical layer of the case study instances in GT compared to that of the bi-directional 10 × 10 grid.

From Figure 7.14 it is apparent that the physical layers clipped from CoCT were also greatly similar in terms of their degree distributions. Akin to the layers in GT, the mean degree was 4. Most nodes had a degree between 2 and 6 with a maximum degree of 10. Again these distributions did not match that of the bi-directional grid exactly, but exhibited impressive regularity.

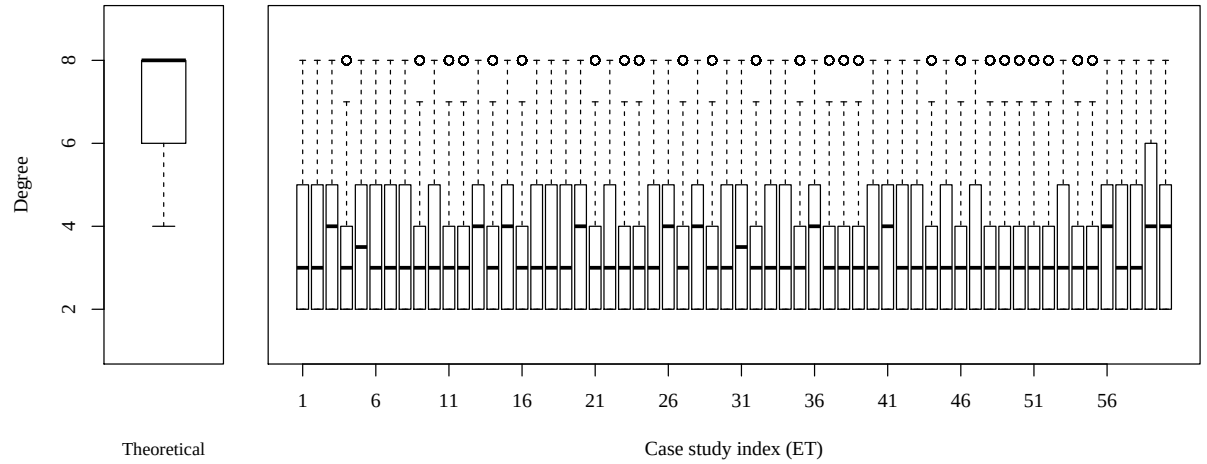


(a) Degree distributions of the theoretical grid.

(b) Degree distributions of the physical layer of the case study instances in CoCT.

Figure 7.14: Degree distributions of the physical layer of the case study instances in CoCT compared to that of the bi-directional 10×10 grid.

Finally the degree distributions of the physical layers clipped from ET are compared to the grid's distribution in Figure 7.15. Here the distributions showed slightly more variation, yet were narrower and had a lower maximum degree compared to the other areas. These distributions also exhibited regularity with a minimum degree of 2, a maximum degree of 8 and the bulk of the distributions between 2 and 5. Again the physical layers from ET were also not an exact match to the structure of the bi-directional grid.



(a) Degree distributions of the theoretical grid.

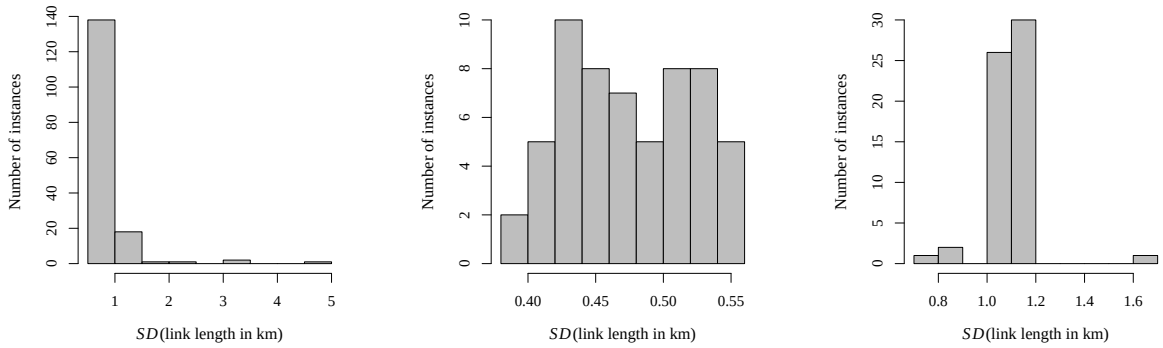
(b) Degree distributions of the physical layer of the case study instances in ET.

Figure 7.15: Degree distributions of the physical layer of the case study instances in ET compared to that of the bi-directional 10×10 grid.

In all three areas the physical layers of the case study instances were not bi-directional grids, but they did all display the regularity that was expected of a road network.

Uniform link length

In the theoretical grid, each link had a length of 1. Only in very exceptional downtown road networks would it be reasonable to expect that road segments are of uniform length. To assess the level of (non)uniformity of the case study instances, the standard deviation of the link length (in km) was recorded for each. A histogram of these standard deviations is plotted for each area in Figure 7.16. By comparison, the standard deviation of the theoretical grid was zero as all links had the same length. From the figure it is apparent that the standard deviations of the case study instances fell mostly within 2km. In light of the diagonal spans of the instances which ranged from 17km to 89km, the variation in link length was not considered exceptional.



(a) GT instances (b) CoCT instances (c) ET instances

Figure 7.16: Standard deviation of the link length in G^2 in each of the case study areas.

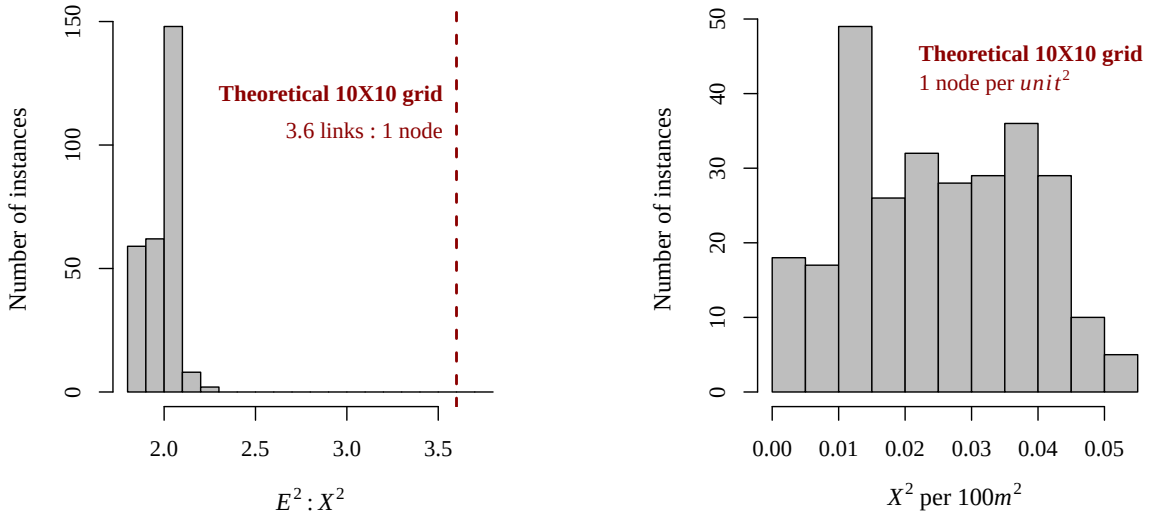
Density of G^2

The final feature to compare was the density of the physical layers measured against the density of the theoretical grid. Density was measured in two ways:

Links to node ratio: This ratio was an aggregate metric that indicated how densely connected the network was. The theoretical grid had 360 links and 100 nodes and thus a ratio of 3.6:1. In Figure 7.17a the highest ratios from the case study instances didn't even exceed 2.5:1.

The theoretical grid was unitless. To find a comparative unit of length in the case study, the question was asked "what length of a road segment could a logistics facility reasonably occupy?" Imagining a warehouse or retail store, $1m$ or $10m$ was obviously too short while a facility spanning $1km$ in an urban area seemed far-fetched. Therefore $100m$ was defined as the unit of measure for the following metric.

Nodes per area: The number of nodes divided by the area covered by the physical layer in $100m^2$ units. This is a more conventional way of measuring "density". The theoretical grid spanned $100units^2$ and contained 100 nodes, resulting in a value of 1. Measured in this way, the case study instances were more than 20 times less dense than the theoretical grid with a maximum value below 0.06 (Figure 7.17b).



(a) Links to node ratio

(b) Nodes per area

Figure 7.17: Comparing the density of the case study physical layers to that of the theoretical grid.

There are different approaches to comparing network densities. In this section we assessed two metrics and both showed that the physical layers of the case study instances were materially less dense than the theoretical grid. This meant that from the outset the underlying road network offered fewer alternatives in getting from any *Point A* to *Point B*.

In summary, although the case study instances did not mimic a bi-directional grid there was notable regularity in the structure of the physical layers and the link length did not vary greatly. The greatest divergence from the theoretical grid was in terms of density and it was expected that this would have dire consequences for network vulnerability.

After extracting the logical layers (G^{1K}) from freight movement data and clipping the physical layers (G^2) from the larger area road networks, these layers had to be associated to create a multilayered network ($\mathcal{M}$) for each of the 279 case study instances.

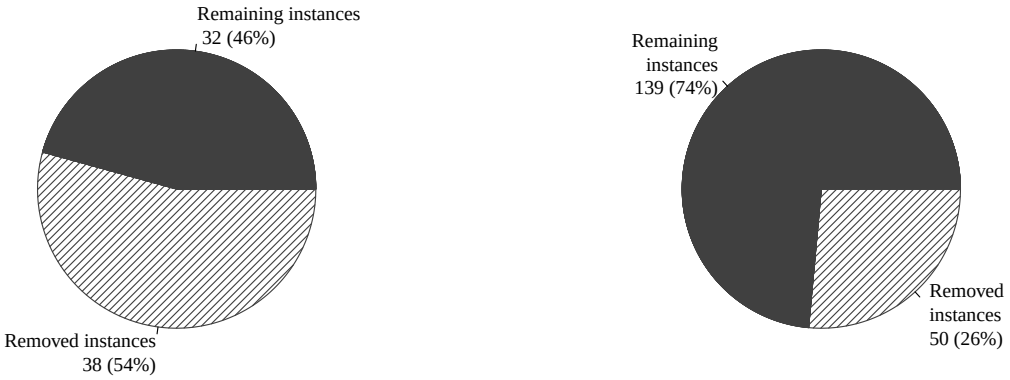
7.4 Creating multilayered case study instances

Layering G^{1K} on G^2 required that each node in X^{1K} be associated with a node in X^2 that was geographically closest. When creating the theoretical instances it was enforced that each x_i^{1K} be assigned to a unique x_s^2. To preserve the realism, we did not enforce this constraint in the case study instances. Therefore it did occur in a number of instances that multiple nodes in X^{1K} were associated with the same node in X^2, especially for networks that covered a small geographic area.

7.4.1 Duplicate associations

When two logistics facilities were associated with the same road intersection (i.e. x_i^{1K} and x_i^{1K} were both assigned to x_s^2 where $i \neq j$), then the length of the shortest path between these two nodes was considered to be N/A. Instances that had three or fewer unique associations effectively reduced that case study instance to a supply chain that was too

simple for the purposes of this thesis. Such instances were removed from the sample. None of the instances of the DH archetype needed to be removed due to this criteria but more than half of the SH instances and a quarter of the FC instances had to be removed (Figure 7.18).



(a) SH archetype

(b) FC archetype

Figure 7.18: Percentage of case study instances removed due to excessive duplicate associations.

7.4.2 Final sample of case study instances

After removing instances with too many duplicate associations, a sample of 191 case study instances remained. The SH archetype represented 73% of the sample followed by the FC archetype at 17% and the DH archetype at 10%. Intuitively, the majority of instances of each of the archetypes came from the largest area, GT (Table 7.5).

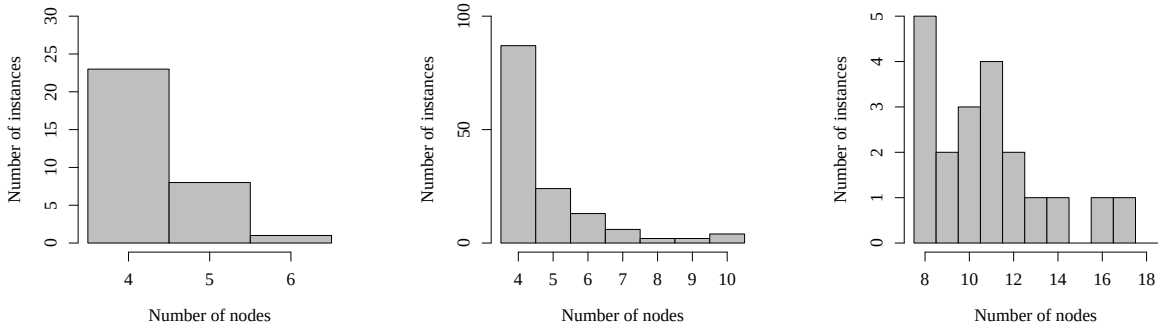
Table 7.5: Final sample of FC, SH and DH instances per area.

	GT	**CoCT**	**ET**	**Total**
FC	15 47%	12 37%	2 16%	32 100%
SH	81 58%	30 22%	28 20%	139 100%
DH	15 75%	2 10%	3 15%	20 100%

From the real-life data it was apparent that for supply chain neighbourhoods that did exhibit a definitive archetype[3] a simple hub design was preferred to a collaborative or

[3]Remember that the sample was biased to those supply chain neighbourhoods that *did* exhibit a definitive archetype and were not of *mixed type*.

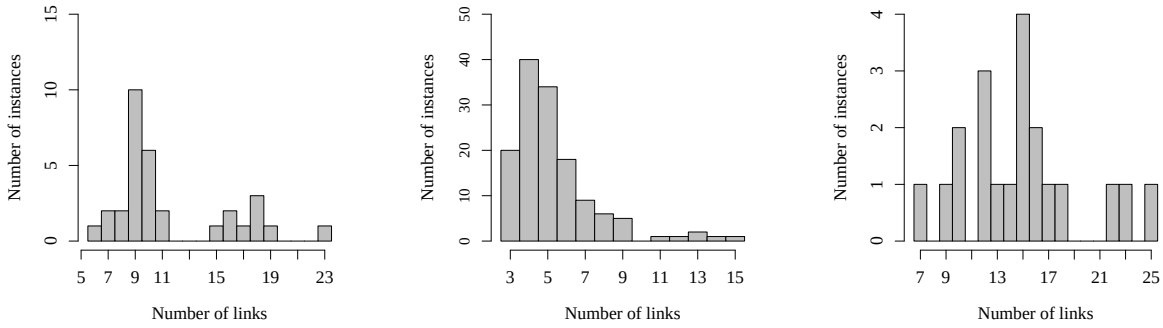
a multi-hub design. Urban supply chains are essentially distribution channels dispersing goods from a central point (production facility, warehouse or distribution centre) to many points of consumption (retail stores). Therefore, a collaborative design may not make much sense functionally on a neighbourhood level. A multi-hub design on a neighbourhood level may also be overcomplicating the relationships, especially considering that the number of nodes per neighbourhood ranged between 4 and 20 (Figure 7.19).



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 7.19: Size distribution of the final sample of case study instances in terms of logical nodes. (One outlier of SH not visible on graph: GT_{26} has 20 logical nodes.)

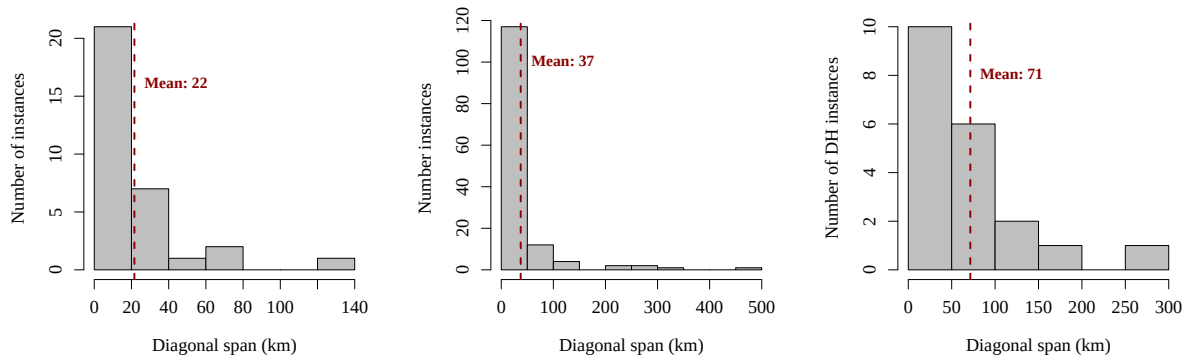
The distribution of the number of logical links illustrated the impact of the deviations on the number of links as discussed in Section 7.2.4. The distribution of links was only slightly higher than the distribution of nodes instead of being a multiple of the number of nodes as suggested by the theoretical formulas in (7.1)–(7.3). The supply chain neighbourhoods represented in the sample were far less densely connected than the theoretical archetypes.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 7.20: Size distribution of the final sample of case study instances in terms of logical links. (One outlier of SH not visible on graph: GT_{26} has 27 logical links.)

Finally, the diagonal span of the final sample of case study instances (Figure 7.21) showed that the FC instances were still most tightly clustered geographically, followed closely by the SH archetype. The DH archetype still covered the largest area by a great margin. On average its diagonal span was nearly three times that of the FC instances and twice that of the SH instances.



(a) FC archetype (b) SH archetype (c) DH archetype

Figure 7.21: Diagonal span of the final sample of logical layers for the case study instances.

Once the sample of multilayered case study instances had been created, the collection of shortest path sets, $\mathcal{C}(\mathcal{S}_{ij})$, had to be defined for each.

7.5 Shortest path sets for case study instances

In the theoretical instances G^2 was a bi-directional 10×10 grid where each link had a length of one. This symmetry resulted in multiple shortest paths of equal length between any two non-adjacent nodes in X^2. Contrarily, in the case study instances G^2 was asymmetrical and link lengths varied to reflect true road segment length (refer again to the standard deviations of link length in Figure 7.16). Thus, the probability of finding multiple shortest paths of exactly the same length between two non-adjacent nodes in X^2 was negligible.

In realistic scenarios, an alternative shortest path need not have exactly the same length as the original shortest path to be considered a viable alternative. Therefore, in the case study instances the requirement that all shortest path alternatives in $\mathcal{S}_{ij}$ have exactly the same length was relaxed. Instead, if the difference between the shortest path length and that of an alternative path was within some acceptable margin, the alternative path was also included in $\mathcal{S}_{ij}$.

Identifying alternative paths with lengths similar to that of the original shortest path presented a computational challenge. R's `igraph` package (Csardi and Nepusz, 2006) was used to construct the case study instances and determine the shortest paths. Existing shortest path algorithms could not accommodate the concept of "length tolerance" that we wished to incorporate. These algorithms could only identify multiple shortest paths if the paths had exactly the same length. One option was to use the algorithm that calculates all unique simple paths between two nodes and then extract only those paths with lengths that fell within some tolerable variance. This approach would have yielded the comprehensive set of all alternative paths that fell within the stated tolerance. Unfortunately, when tested on just one node-pair from the smallest of the case study instances, this algorithm took longer than two days to produce the set of all simple paths. The exponential nature of the algorithm's computation time negated its usefulness given the size of G^2 and the computational resources available. A custom algorithm was required to compose $\mathcal{S}_{ij}$ for each node-pair.

The algorithm is explained using the simplified network as shown in Figure 7.22. Note that this network is not an example of the FC, SH or DH archetypes, but rather a fictitious

network specifically chosen for its ability to explain the working of the algorithm. The network has nine nodes denoted by $x_q, q \in \{1, 2, \ldots, 9\}$ that are connected by directed links as indicated in the figure. Each link is defined by:

$$e_{qr} = \begin{cases} 1 & \text{if } x_q \text{ is directly connected to } x_r \\ & q, r \in \{1, 2, \ldots, 9\} \\ 0 & \text{otherwise} \end{cases} \tag{7.4}$$

The length of each link is indicated by a blue link label in the figure.

For this illustration we wish to determine a set of shortest paths between source node x_1 and target node x_4 using the steps outlined below:

Step 1: Determine the shortest path between x_1 and x_4.

Table 7.6: $\mathcal{S}_{14}$ after Step 1

Description	**Path**	L_{ij}	**Figure**
Original	$e_{17} \to e_{78} \to e_{84}$	3	7.22a

Step 2: Sequentially remove one link from the original shortest path at a time and determine the new shortest path(s).

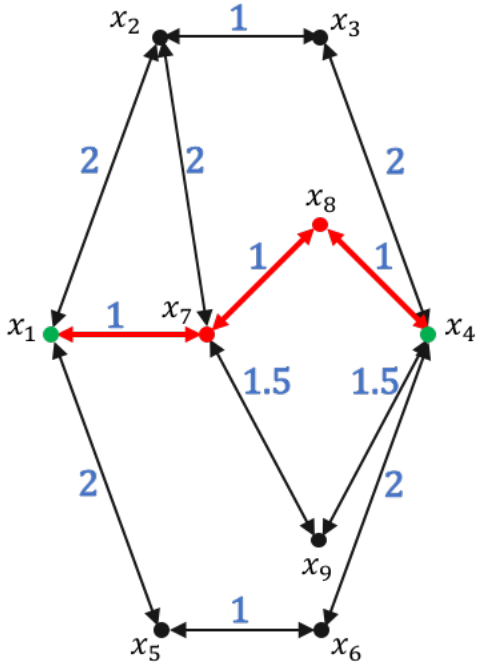
Table 7.7: $\mathcal{S}_{14}$ after Step 2

Description	**Link removed**	**Path**	L_{ij}	**Figure**
Original	–	$e_{17} \to e_{78} \to e_{84}$	3	7.22a
Alternative 1	e_{17}	$e_{12} \to e_{23} \to e_{34}$	5	7.22b
Alternative 2		$e_{15} \to e_{56} \to e_{64}$	5	
Alternative 3	e_{78}	$e_{17} \to e_{79} \to e_{94}$	4	7.22c
Alternative 4	e_{84}	$e_{17} \to e_{79} \to e_{94}$	4	7.22d

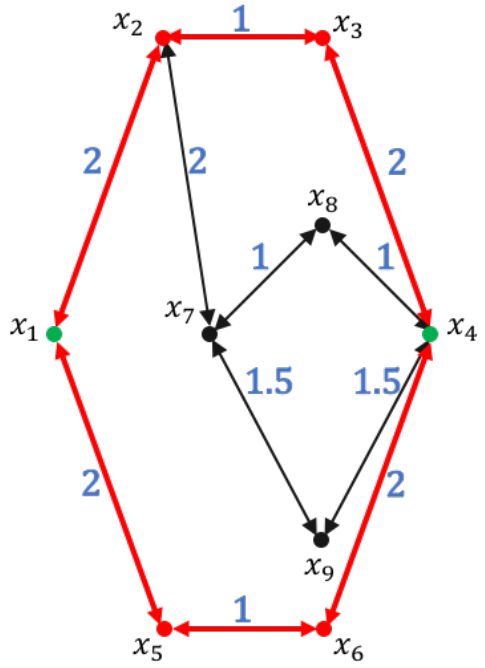
Step 3: Remove Alternative 4 as it is a duplicate of Alternative 3.

Table 7.8: $\mathcal{S}_{14}$ after Step 3

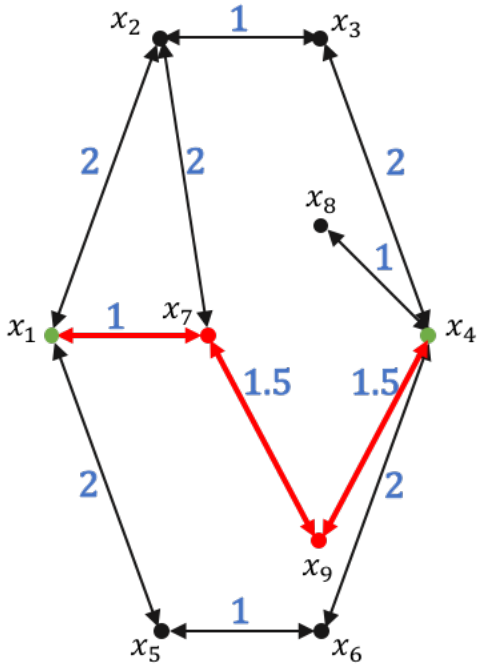
Description	**Path**	L_{ij}
Original	$e_{17} \to e_{78} \to e_{84}$	3
Alternative 1	$e_{12} \to e_{23} \to e_{34}$	5
Alternative 2	$e_{15} \to e_{56} \to e_{64}$	5
Alternative 3	$e_{17} \to e_{79} \to e_{94}$	4



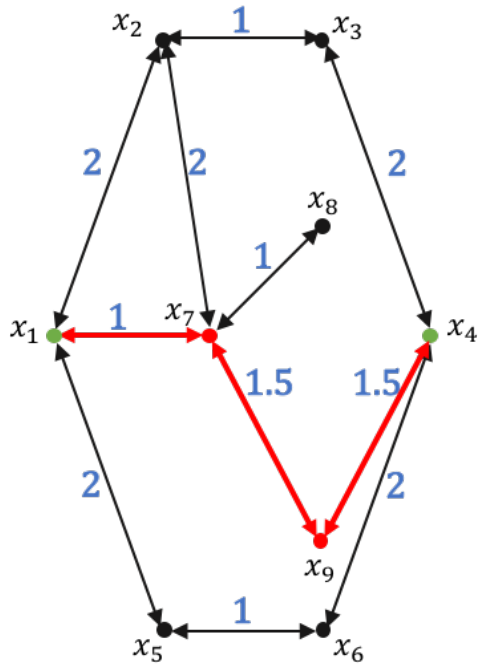
(a) Step 1: Determine original shortest path between x_1 and x_4.



(b) Step 2: Remove e_{17} and find alternative shortest paths.



(c) Step 2: Remove e_{78} and find alternative shortest paths.



(d) Step 2: Remove e_{84} and find alternative shortest paths.

Figure 7.22: An illustration of the first two steps of the custom shortest path collection algorithm used for the case study instances.

Step 4: Remove shortest paths that fall outside the length tolerance. Assume that the length tolerance was 50%. Then all alternative paths that are shorter than 150% of the original shortest path length ($3 \times 150\% = 4.5$) are retained.

Table 7.9: $\mathcal{S}_{14}$ after Step 4

Description	**Path**	L_{ij}
Original	$e_{17} \rightarrow e_{78} \rightarrow e_{84}$	3
Alternative 3	$e_{17} \rightarrow e_{79} \rightarrow e_{94}$	4

This example illustrates the algorithm in the case where there was only one original shortest path. In the exceptional cases where there were more than one original shortest path, the algorithm had to break all original paths simultaneously to force the search for new alternatives. Thus, one link from each original path had to be removed simultaneously. In such cases all the unique combinations of links were first enumerated as part of Step 1. Then in Step 2, each of the combinations were removed in turn to determine alternative paths.

This algorithm was executed to determine $\mathcal{S}_{ij}$ in each of the case study instances.

7.5.1 Shortest path set statistics of the initial networks

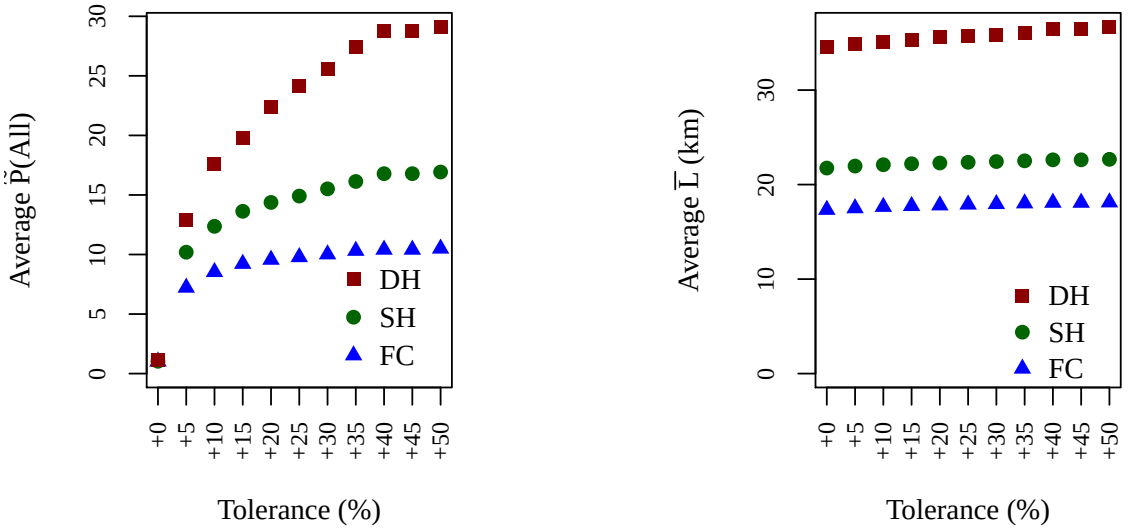
The larger the length tolerance chosen for the shortest path algorithm, the more alternative paths would have been included in the shortest path sets. Although more would've been better from a redundancy point of view, there was a limit to how large the tolerance could be to still be realistic. As a starting point it was assessed how the median of the shortest path set size and average shortest path length varied with different tolerances.

The algorithm was executed with a tolerance varying between +0% (no tolerance) and +50%. After each execution $\tilde{P}(\text{All})$ and $\bar{L}$ was determined for each case study instance. Figure 7.23a plots the average of $\tilde{P}(\text{All})$ across all instances that correspond to a specific tolerance. Figure 7.23b plots the average of $\bar{L}$ across all instances.

At a +0% tolerance instances had, on average, 1 path in their shortest path sets for each of the three archetypes. Increasing the tolerance had the greatest effect on the DH archetype. The average of $\tilde{P}(\text{All})$ rose sharply to more than 25 paths before reaching a slight plateau from +40% onwards. The average of $\tilde{P}(\text{All})$ was less sensitive in the hub archetypes. There was an initial jump when the tolerance was increased to +5% but after that the increase was gradual.

This difference could be explained by the geographic span of the instances. The bigger the area covered by the instance, the larger the road network reflected in G^2 and the more links were included in initial shortest paths as facilities were further apart. Therefore, the algorithm went through more iterations to remove each shortest path link *and* there were more alternatives available due to the larger G^2 network. The DH instances covered the largest area, followed by the SH instances and then the FC instances as shown in Figure 7.21.

The conclusion was that a small length tolerance had a decisive impact on increasing the number of paths in the shortest path sets, but larger tolerances offered diminishing returns in terms of increasing $\tilde{P}(\text{All})$. The diminishing returns were a result of the bounding box implemented when clipping the road networks using Algorithm 4. As the length



(a) Average $\tilde{P}$(All)

(b) Average $\bar{L}$

Figure 7.23: Impact of length tolerance on the average shortest path set size and average shortest path length.

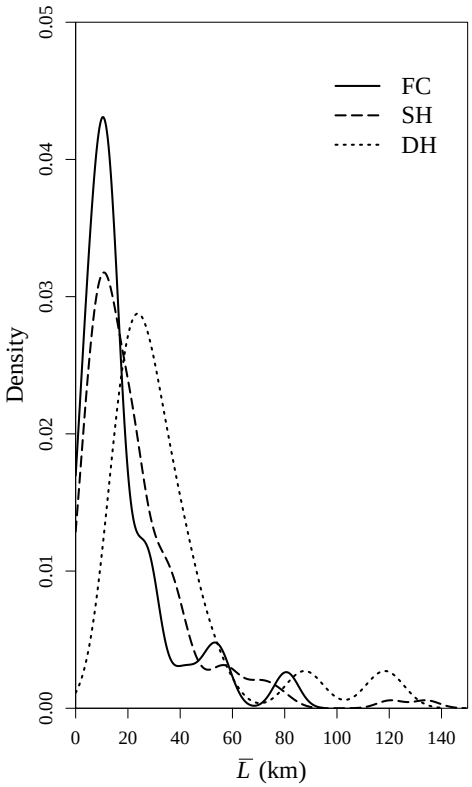
tolerance increased, fewer and fewer shortest path sets could benefit as there were no additional paths available on G^2 within the bounding box.

The (un)availability of alternative paths also explained the minimal increase in the average of $\bar{L}$ (Figure 7.23b). In the case of a densely connected road network one would have expected the average of $\bar{L}$ to increase by the same % as the length tolerance, at least until the bounding box constrained additional alternatives. However, at a length tolerance of 50% the increase was only 4.5%, 4.3% and 6.2% for the FC, SH and DH archetypes, respectively.

These results showed that an increase in length tolerance had a negligible impact on the average shortest path length and only a small impact on the sizes of the shortest path sets. In light of these results and based on intuition regarding freight transport in these areas, the length tolerance was set at 25%.

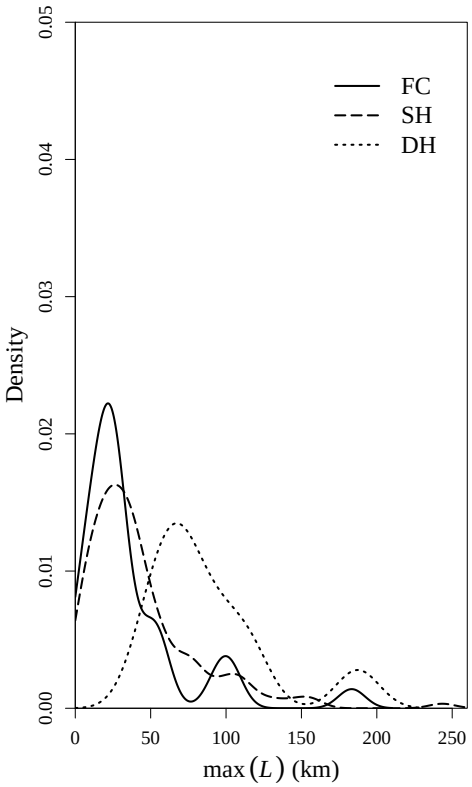
Initial distributions of $\bar{L}$

The initial distributions of $\bar{L}$ for the case study instances are shown in Figure 7.24. In comparing these distributions to those of the theoretical distributions repeated in Figure 7.25 we noticed two differences.



Archetype	Min	Mean	Max	Kurtosis
FC	0.9	17.9	80.6	6.7
SH	1.3	22.4	134.2	11.2
DH	14.6	35.7	118.3	7.0

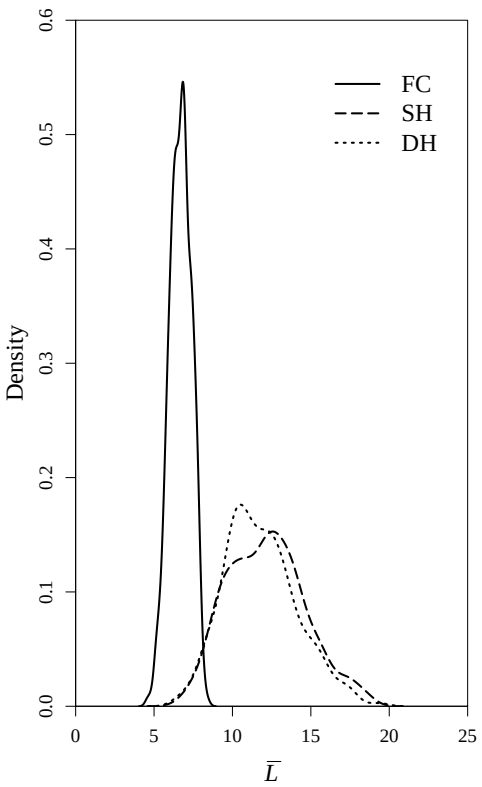
(a) Average shortest path length, $\bar{L}$ (km)



Archetype	Min	Mean	Max	Kurtosis
FC	1.6	36.5	183.3	8.7
SH	2.4	43.1	243.5	8.9
DH	44.3	88.3	188.4	4.3

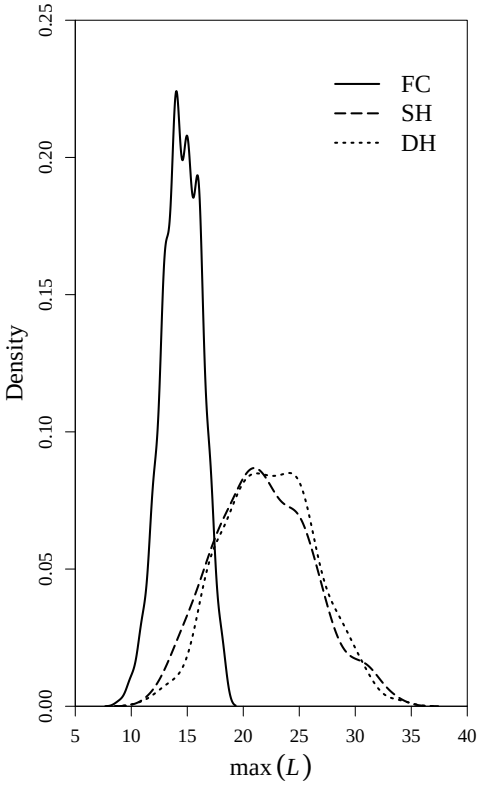
(b) Diameter, $\max(L)$ (km)

Figure 7.24: Case Study: Initial distributions of the diameter and average shortest path lengths.



Archetype	Min	Mean	Max	Kurtosis
FC	4.6	6.8	8.4	2.6
SH	6.5	12.2	18.8	2.7
DH	6.4	11.8	19.5	3.0

(a) Average shortest path length, $\bar{L}$



Archetype	Min	Mean	Max	Kurtosis
FC	9.0	14.5	18.0	2.7
SH	12.0	21.8	34.0	2.7
DH	11.0	22.3	34.0	2.7

(b) Diameter, max (L)

Figure 7.25: Theoretical: Initial distributions of the diameter and average shortest path lengths.

Grouping of the distributions: In the theoretical instances the distributions for the hub archetypes were very similar while the FC archetype was narrower and shifted to the left. This implied that FC instances were distinctly more efficient than the hub archetypes and that the hub archetypes were very similar.

In the case study instances we saw that the FC archetype was no longer that different to the hub archetypes. At the same time the difference between the hub archetypes had now increased slightly with the DH archetype shifted further to the right.

The second deviation discussed in Section 7.2.4 regarding the bi-directionality of links offered an explanation for the change in the FC archetype. The FC instances now included many indirect connections to compensate for the absence of bi-directional links. Indirect connections resulted in longer shortest paths between two nodes. This increased $\bar{L}$ and therefore the distribution of the FC archetype was closer to that of the hubs although it was still the narrowest of the three.

The DH archetype's distribution was furthest to the right. One contributing factor was that many indirect connections between spoke nodes had a logical path length of three to start with. This made the shortest paths longer as they had to route via more nodes. The greater contributing factor, however, was that the diagonal

span of the DH archetype was so much broader than that of the other archetypes. Facilities were spread further apart and therefore shortest paths were longer.

In the theoretical instances we could conclude that one archetype was more efficient than another by virtue of the distribution of $\bar{L}$. This same conclusion could not be made from the case study instances due to the variance in the diagonal span.

Distribution shape: The theoretical distributions showed a clear central tendency with kurtosis values close to 3. A kurtosis value of 3 indicated that a distribution was as likely as the normal distribution to produce outliers. The case study distributions did not show a central tendency at all. They were left-skewed with long right tails. Kurtosis values exceeding 3 also indicated that the occurrence of outliers was far more prevalent. Another observation was that the FC and DH distributions had distinct modes in their right-tails.

The left tendency indicated that the majority of instances had relatively short average paths. The mean of $\bar{L}$ was smaller than the mean of the diagonal span (Figure 7.21) for each archetype. The means of the diameter of each archetype were 68%, 62% and 24% larger than the mean of the diagonal span for the FC, SH and DH archetypes, respectively. Therefore we concluded that initially there wasn't a great degree of zigzagging or roundabout travel on the road networks between facilities, but that shortest paths were straight-forward.

The right tail and occurrence of outliers was again explained by the distribution of the diagonal span of the instances. Each of the distributions in Figure 7.21 show a prominent right tail. For each archetype there were a number of instances that covered a much larger area than the rest. Intuitively these instances would also have had much longer shortest paths between node-pairs.

In summary, two differences between the case study and theoretical instances resulted in vastly different distributions for $\bar{L}$. Firstly, the fact that the FC instances included a number of indirect connections reduced the distinction between this archetype's distribution and those of the hub archetypes. Secondly, case study instances covered a broad range of geographic areas whereas the bounds of the area covered in the theoretical instances were tighter with a minimum of $12 units^2$ and a maximum of $100 units^2$.

Initial distributions of the sum of set sizes

The sum of set sizes per instance is defined as:

$$\text{Total set size (All)} = \sum_{i,j \in \mathcal{S}_{ij} i \neq j} P_{ij} \tag{7.5}$$

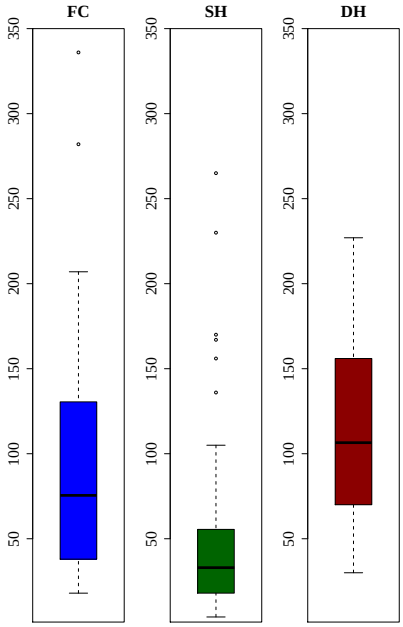
when considering *all* node-pairs, and

$$\text{Total set size (Dir)} = \sum_{i,j \in SD_{ij}; i \neq j} P_{ij} \tag{7.6}$$

when considering only directly connected node-pairs.

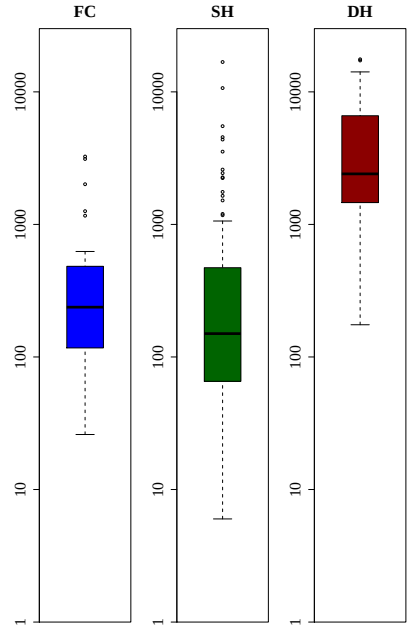
This was an aggregate indication of the inherent redundancy present in an instance before any disruptions had occurred. The initial distributions of the sum of set sizes for

the case study instances are shown in Figure 7.26 and were compared to those of the theoretical instances shown in Figure 7.27.



Archetype	Min	Median	Max	Kurtosis
FC	18	76	336	4.8
SH	4	33	265	12.1
DH	30	107	227	2.2

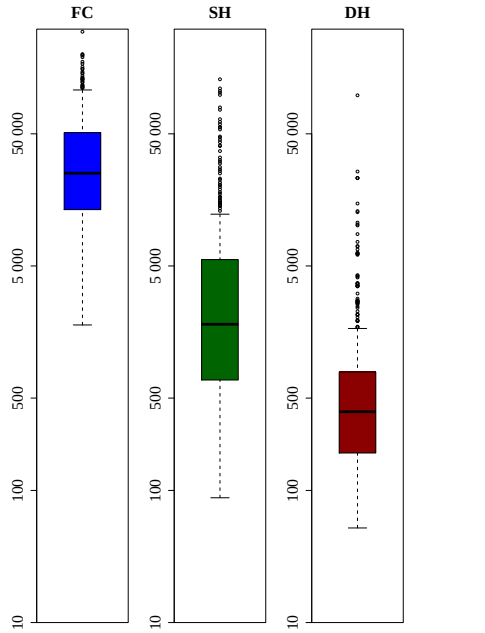
(a) Distributions of $\sum_{i,j \in SD_{ij}; i \neq j} P_{ij}$



Archetype	Min	Median	Max	Kurtosis
FC	26	238	3 252	8.1
SH	6	150	16 798	50.4
DH	175	2 406	17 608	3.7

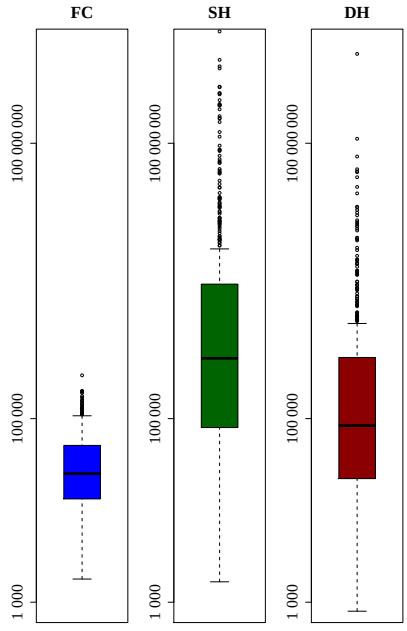
(b) Distributions of $\sum_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$

Figure 7.26: Case Study: Distributions of the sum of set sizes.



Archetype	Min	Median	Max	Kurtosis
FC	1 788	25 171	3.0×10^5	10
SH	88	1 810	1.3×10^5	30
DH	52	394	9.8×10^4	265

(a) Distributions of $\sum_{i,j \in SD_{ij}; i \neq j} P_{ij}$



Archetype	Min	Median	Max	Kurtosis
FC	1 788	25 171	3.0×10^5	10
SH	1 668	453 765	1.7×10^9	134
DH	798	84 021	9.4×10^8	465

(b) Distributions of $\sum_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$

Figure 7.27: Theoretical: Distributions of the sum of set sizes.

Again two notable differences were identified: the distributions' shape and range, and the ordering of the three archetypes.

Distribution shape and range: The theoretical distributions were orders of magnitude larger than the case study distributions. This meant that there were many more alternative shortest paths present in the theoretical instances than in the case study instances.

The density of G^2 and its grid-like structure played a prominent role in providing alternative shortest paths. The analysis in Section 7.3.3 showed that G^2 was considerably less dense in the case study instances than in the theoretical instances (Figure 7.17). Furthermore, although its degree distribution showed a regular structure, G^2 in the case study instances was not a bi-directional grid and intersections had fewer road segments connecting them to the rest of the network (Figures 7.13–7.14). In the theoretical instances it was the combinatorial effect of the grid structure that resulted in the very long right tails and high kurtosis values for the distributions. By comparison, the case study instances did not have significant right tails and their kurtosis values showed that the occurrence of outliers was less likely than that of the normal distribution.

Another factor influencing these distributions was the number of nodes and links in G^{1K}. These distributions represented a summative metric and therefore the absolute size of the instances made a difference. Figure 7.19 shows that all FC

and SH instances had fewer nodes than their theoretical counterparts which had 12 nodes. The DH instances were larger and some instances even exceeded the theoretical size of 12 nodes, but the mean was still lower at 11 nodes. In addition to having fewer nodes, the case study instances also had far fewer logical links than expected (refer to Figure 7.10) because most of the supply chain relationships were not bi-directional. Smaller instances (in terms of both nodes and links) resulted in fewer shortest path sets to add together.

One last consideration related to the algorithm used to construct the shortest path set. It could not be proven, *theoretically*, that the algorithm was guaranteed to construct a comprehensive set of all shortest paths that fell within the stated tolerance. Therefore, the shortest path sets were possibly incomplete. Exploring alternative algorithms and comparing their performance in terms of completeness and computational complexity was marked for future research.

In brief, the far lower density of G^2 and its deviation from the bi-directional grid structure combined with the smaller G^{1K} layers resulted in strikingly fewer shortest path alternatives per instance. Admittedly, it was also possible that the shortest path algorithm did not identify all possible paths.

Ordering of archetypes: In the theoretical instances the contrast between the archetypes was prominent. In the case study instances the distributions were closer together and the ordering of the archetypes was slightly different. The fact that the diagonal spans and instance sizes varied so greatly between the case study instances made comparisons dubious. For instance it was no longer valid to conclude that the DH archetype could be less vulnerable as it had more alternative paths because this distribution was greatly influenced by the fact that the DH instances spanned a broader area and had more logical nodes and links than the other two archetypes.

Another observation was that the FC and SH distributions were more similar in the case study. This was firstly due to their kindred diagonal spans and sizes and secondly due to the fact that the FC instances now included indirect connections instead of only direct connections.

In summary, making comparisons between and generalisations regarding the archetypes was far less justifiable in the case study as the qualities that influenced shortest path length and the sum of set sizes varied greatly. Having said that, it could still be concluded that across the board there were far fewer alternative shortest paths available to the case study instances. This was primarily due to the structure of the underlying road network. From these initial results it was expected that the case study instances would be more vulnerable than the theoretical instances had been. The following chapter presents the results from the random error simulation used to assess this vulnerability. It also presents and discusses the behaviour of the vulnerability metrics throughout the simulation.

Chapter 8

Case study: Link-based random error simulation

The previous chapter explained how 191 case study instances were constructed from real-life data in three urban areas namely the City of Cape Town (CoCT), eThekwini Metropolitan Municipality (ET) and Gauteng Province (GT). Comparing these case study instances to the theoretical instances it was apparent that many of the assumptions made in constructing the theoretical instances no longer held true for the case study instances — both in terms of the logical layer (G^{1K}) and the physical layer (G^2).

These deviations from the theoretical assumptions were reflected in the initial distributions of the shortest path set size and average shortest path length of the case study instances. From their characteristics it was postulated that the case study instances would be far more vulnerable to random link disruption than the theoretical instances.

To test the vulnerability a link-based random error simulation was executed on the case study instances. This simulation was similar to the one executed on the theoretical instances in Chapter 5. The simulation started with an undisturbed case study instance. One percent of the initial number of links in G^2 were randomly selected and removed. If all the direct and indirect connections between nodes in G^{1K} were still intact, the instance was still considered connected and a further 1% of the links were randomly removed. So the instance was progressively disrupted until it became disconnected. Shortest path statistics and the vulnerability metrics were measured after each progressive disruption.

The next section discusses the efficiency loss and rate of disconnection of the case study instances. Thereafter the behaviour of the vulnerability metrics is presented and their correlation to efficiency loss and robustness is discussed. The chapter concludes with a discussion on the validity of the vulnerability metrics when dealing with real-life data.

8.1 Results of the link-based random error simulation

The levels of damage endured by the case study instances were the same as those described for the theoretical instances in Section 4.1.1. An instance experienced efficiency loss when the average shortest path length of the overall network was increased. This implied that, on average, it would take longer for freight to be transported from origin facilities to destination facilities. An instance was considered disconnected when two facilities that were directly connected could no longer ship freight from one to the other. In network

terms this was when one of the links in G^{1K} were broken. However, a supply chain would presumably continue shipping freight between its other facilities after it becomes disconnected. An instance was therefore only considered destroyed when there were no more links in G^{1K} intact.

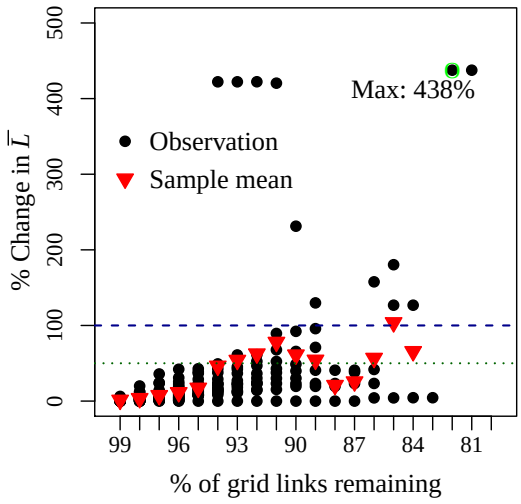
In this thesis we only continued simulations until the point of disconnection. In practice, the disconnection of two facilities would spark mitigating action by a supply chain. One example could be that other facilities are tasked to support the disconnected facilities. This would change the structure of G^{1K}. Because it was difficult to anticipate the mitigating action that could have been taken and its effect on the instance's structure, continuing simulations beyond disconnection was questionable in terms of validity.

8.1.1 Efficiency loss before disconnection

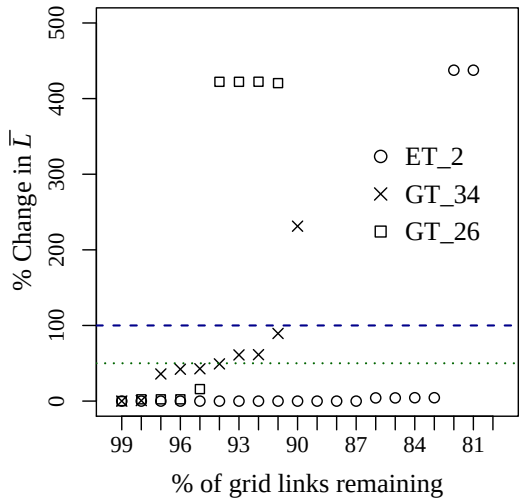
The shortest path statistics were measured after each progressive disruption. Efficiency loss was determined by measuring the percentage change in $\bar{L}$ from the initial undisturbed network to $\bar{L}$ right before the final disruption that disconnected it.

Fully Connected (FC) archetype

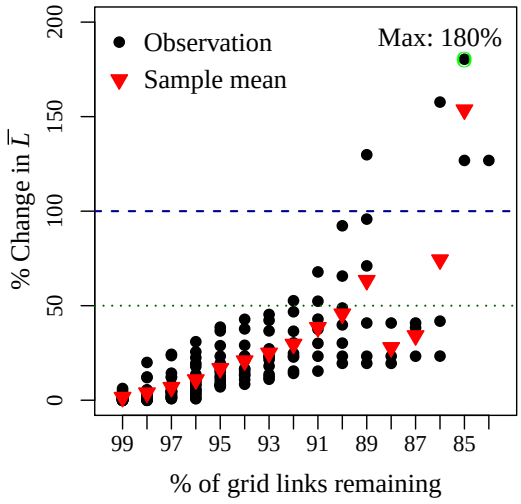
At first the trend in efficiency loss for the FC archetype was problematic. Figure 8.1a shows the efficiency loss for all 32 FC instances. After each disruption the mean of the efficiency loss for all instances that were not yet disconnected was measured. This is indicated by the red triangles in the graph. The trend in the sample mean would suggest that overall efficiency was lost during the first few disruptions but from the 9^{th} disruption onwards $\bar{L}$ actually became shorter and thus the instances more efficient. What was also observed was that there were some exceptional efficiency loss percentages in excess of 400%.



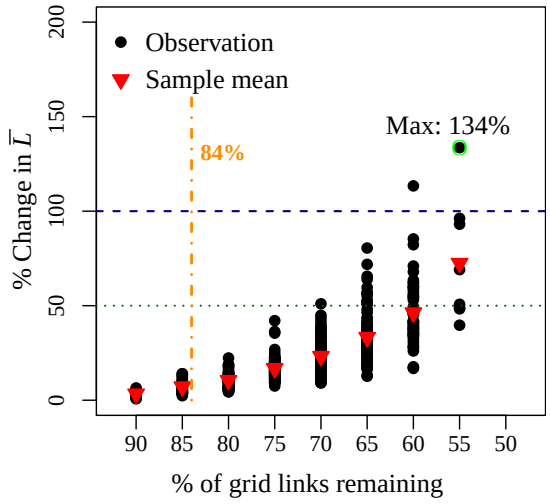
(a) FC efficiency loss for all 32 case study instances.



(b) Efficiency loss of three exceptional FC instances.



(c) FC efficiency loss for 29 case study instances (exceptional instances removed).



(d) FC efficiency loss for 500 theoretical instances. The vertical line at 84% indicates the limit of the x-axis of 8.1c.

Figure 8.1: Case Study vs Theoretical: Efficiency Loss (FC)

Upon further investigation three outlier instances[1] were identified (Figure 8.1b). ET_2 showed virtually zero efficiency loss until its final two disruptions where suddenly efficiency loss shot up to over 400%. GT_26 behaved similarly, shooting up suddenly to over 400% efficiency loss after the 5$^{\text{th}}$ disruption. GT_34 was less extreme but the sudden jump to over 200% efficiency loss still had an undue effect on the sample mean trend.

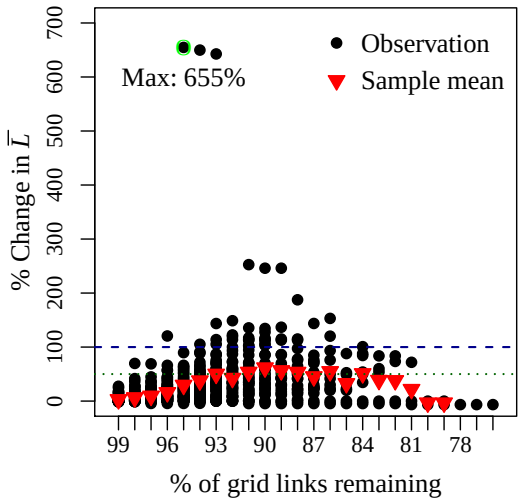
After removing the exceptional instances, the trend in efficiency loss was more intuitive (Figure 8.1c). Efficiency losses increased monotonically until the 11$^{\text{th}}$ disruption. Thereafter the sample mean became more erratic as the number of instances that were not yet disconnected dwindled.

[1]These instances are referred to by the an acronym for the urban area (CT, ET or GT) followed by the index number of that instance, e.g. CT_1 which refers to instance 1 in the set of CoCT instances.

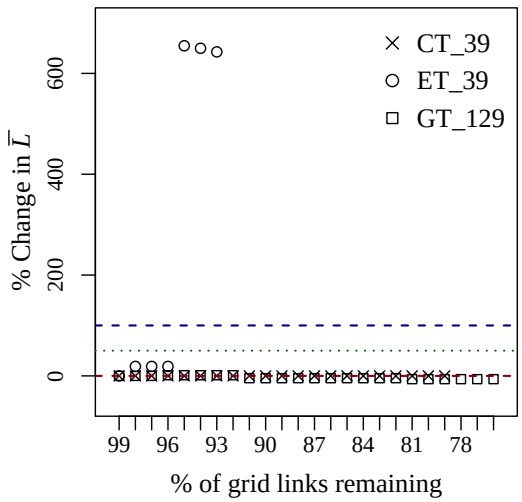
By comparison the efficiency loss for the theoretical FC instances is graphed in Figure 8.1d. The extent of efficiency loss was comparable, but the rate at which the case study instance lost efficiency was far greater. The orange line in Figure 8.1d indicates the point at which all the non-exceptional case study instances were disconnected. By the time the theoretical instances had lost only about 10% of their efficiency, the case study instances had lost more than 50% of their efficiency. This implied that random disturbance of the road network is very rapidly visible in these real-life instances through the increase of travel distance between facilities.

Single Hub (SH) archetype

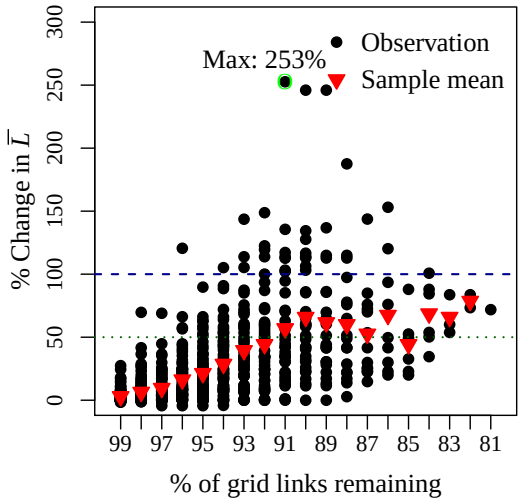
Similar to the FC archetype, the SH archetype exhibited a troublesome distribution of efficiency loss at first. In Figure 8.2a the trend showed a decrease in efficiency up until the 10^{th} disruption. After this it seemed that the efficiency actually increased again.



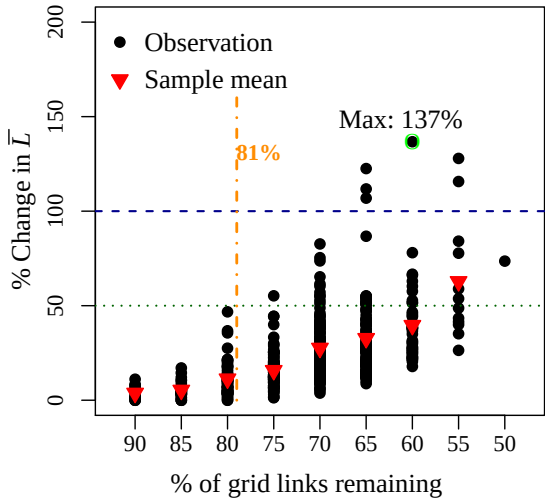
(a) SH efficiency loss for all 139 instances.



(b) Efficiency loss of three exceptional SH instances.



(c) SH efficiency loss for 136 instances (exceptional instances removed).



(d) SH efficiency loss for 500 theoretical instances. The vertical line at 81% indicates the limit of the x-axis of 8.2c.

Figure 8.2: Case Study vs Theoretical: Efficiency Loss (SH)

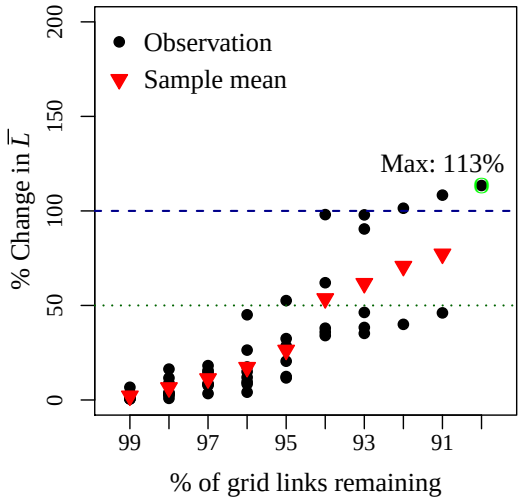
Further investigation identified three exceptional instances (Figure 8.2b). CT_39 literally had no efficiency loss up until the 21st disturbance and then suddenly became disconnected. This implied that the first 21 random disturbances did not affect the shortest paths at all, but in the 22nd disturbance one or more very critical links were removed, causing disconnection. Instance ET_39 showed reasonable efficiency losses at first which then suddenly jumped beyond 600% before becoming disconnected. This jump had an undue effect on the sample mean. Lastly, GT_129 showed minimal efficiency losses (below 2%) until the 8th disruption. After that it actually became more efficient than the initial network. This was possible because of the tolerance of 125% that allowed paths to be included in the shortest path sets that weren't the absolute shortest. The first 8 disruptions removed enough of these "longer" shortest paths to decrease the overall $\bar{L}$.

Figure 8.2c shows the efficiency loss of the remaining instances once the exceptions were removed. The sample means show a monotonically increasing trend up until the 12th disruption. Thereafter the trend became erratic with pertinent decreases. Although the sample size had decreased, a pertinent number of instances were still reflected in the sample mean. What this showed was that the efficiency loss patterns between instances were more heterogenous than the theoretical results shown in Figure 8.2d.

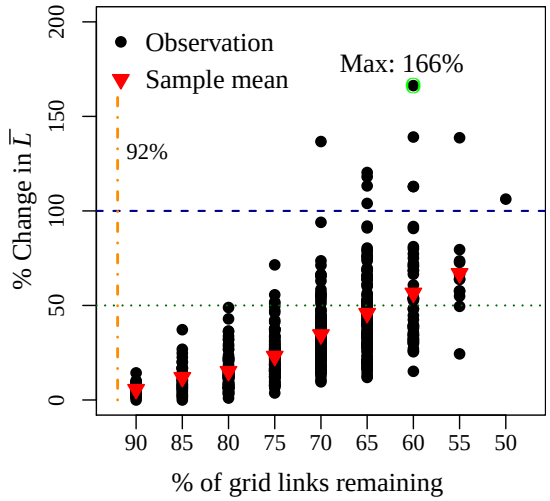
Similar to the FC archetype, the comparison between the case study and theoretical trend showed that in general the efficiency losses were in the same range, but that the case study instances lost efficiency far more rapidly.

Double Hub (DH) archetype

The sample of DH instances was the smallest of the three archetypes with only 20 instances. Out of these 20 instances there were no notable exceptions. The trend of the sample mean was monotonically increasing, indicating that in general efficiency was decreased as more road segments were removed from the underlying road network. Again the extent of efficiency loss was relatively similar to that of the theoretical instances shown in Figure 8.3b. Remarkably, all the case instances had incurred their total ambit of efficiency loss and had become disconnected before the first theoretical instances even lost any efficiency. This is illustrated by the orange vertical line in Figure 8.3b that precedes the first observations of efficiency loss.



(a) DH efficiency loss for all 20 instances.



(b) DH efficiency loss for 500 theoretical instances. The vertical line at 92% indicates the limit of the x-axis of 8.3a.

Figure 8.3: Case Study vs Theoretical: Efficiency Loss (DH)

Discussion

The first observation was that for two of the three archetypes there were exceptional instances whereas for the theoretical instances there were no notable exceptions. We knew from our analysis in Chapter 7 that unlike the theoretical instances there were pronounced differences in the structure of the case study instances. Particularly, the

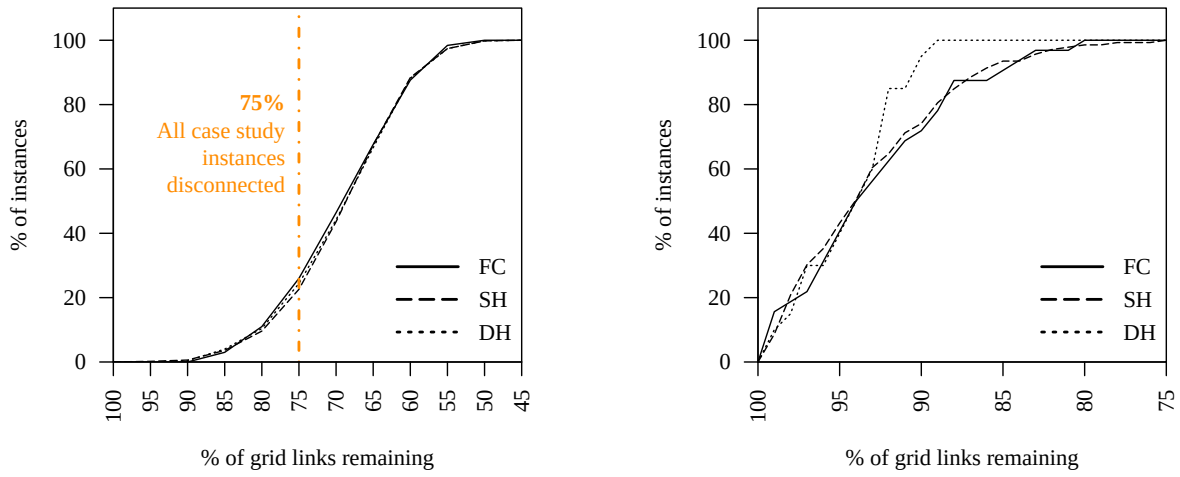
differences in the diagonal span and sizes of both the logical and physical layers could have been the primary contributors to the variance in efficiency loss patterns.

A second observation was that while efficiency gain (the increase in $\bar{L}$) was impossible in the theoretical instances, it was possible in the case study instances due to the tolerance used when establishing the shortest path sets. The result was that in a number of case study instances efficiency actually increased after an initial few disruptions. Therefore, the blanket statements that "the average shortest path continues increasing until an instance is disconnected", or that "the longer an instance survives the greater its overall efficiency losses" could no longer be stated with such confidence. However, it made no sense, realistically, that the more sparse the road network becomes, the shorter the paths between facilities would become. This observation was the result of the way in which efficiency loss is defined and measured. Finding more valid ways of measuring efficiency losses for the case study instances was listed as future work emanating from this thesis.

The final observation was that for all archetypes, efficiency loss occurred more rapidly than for the theoretical instances. This was explained by the fact that the physical layers of the case study instances were so much less dense than the bi-directional grid of the theoretical instances. The likelihood that a random selection of one percent of the road segments would include road segments necessary for the shortest paths was just that much greater in the case study instances. For the same reason, the case study instances were also disconnected far quicker, as the next section shows.

8.1.2 Disconnection of networks

The cumulative percentage of theoretical instances that were disconnected after each progressive disruption were remarkably similar for the three archetypes (Figure 8.4a). All the theoretical instances had 12 nodes in the logical layer, strictly bi-directional links between these nodes and were all limited in terms of span by the 10×10 grid. What differentiated the instances were their archetypes and the placement of the nodes on the grid. In Figure 8.4a we note that the distributions for the archetypes were nearly indistinguishable. This implied that the archetype really played very little role in whether an instance would become disconnected sooner rather than later.



(a) Theoretical

(b) Case Study

Figure 8.4: Case Study vs Theoretical: Cumulative distribution of instance disconnection.

On the other hand, in the case study the instances representing one archetype varied from the instances representing another. The greatest difference was in terms of the geographic span, but there were also differences in terms of the number of nodes and number of links in the logical layers and the density of the physical layers. Despite these important variations, the cumulative distributions of the case study instances were relatively similar as seen in Figure 8.4b. Again it was apparent that the archetype was not a major determinant in when an archetype would become disconnected.

Another observed difference between the theoretical and case study distributions was that the theoretical distribution displayed a distinct S-curve shape while the case study distributions did not. Initially the random error simulation had very little impact on the theoretical instances. After 15% removal of the grid links, the rate at which instances were disconnected hiked up until the last few diehard instances caused the distribution to plateau. In the case study, the rate of disconnection was steep from the start with a number of instances becoming disconnected immediately. The rate of disconnection gradually tapered before plateauing. To explain this shape we also considered that all the case study instances had been disconnected by the time that only 75% of the grid links remained. In contrast, only 24% of the theoretical instances had become disconnected when 75% of the grid remained. In a nutshell, the case study instances were far more vulnerable with very little resistance to the removal of road segments. This concurred with the observation in Section 7.5.1 that $\sum_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$ was orders of magnitudes smaller than that of the theoretical instances. There simply weren't as many alternative shortest paths available to the case study instances.

It was clear both from the efficiency loss and disconnection results that the impact of the random removal of road segments was severe and immediate. Next, we investigated whether the vulnerability metrics developed and tested using the theoretical instances behaved similarly in the case study instances.

8.1.3 Redundancy

Redundancy was defined as a measure of the number of alternative shortest paths that were available to an instance in Section 5.3. The size of the shortest path sets P_{ij} was the underlying characteristic used to quantify redundancy. There were two aspects of the distribution of P_{ij} that were of interest. The first was the centrality of the distribution. This gave an aggregate indication of the level of redundancy across all shortest path sets. The second aspect was the centrality of the left-tail of the distribution that contained those shortest path sets with the fewest alternatives. It was also deemed necessary to differentiate between the redundancy of the direct shortest path sets alone and the collection of all shortest paths sets. The summary of the aspects and their metrics from Section 5.3 is repeated below in Table 8.1 for convenience.

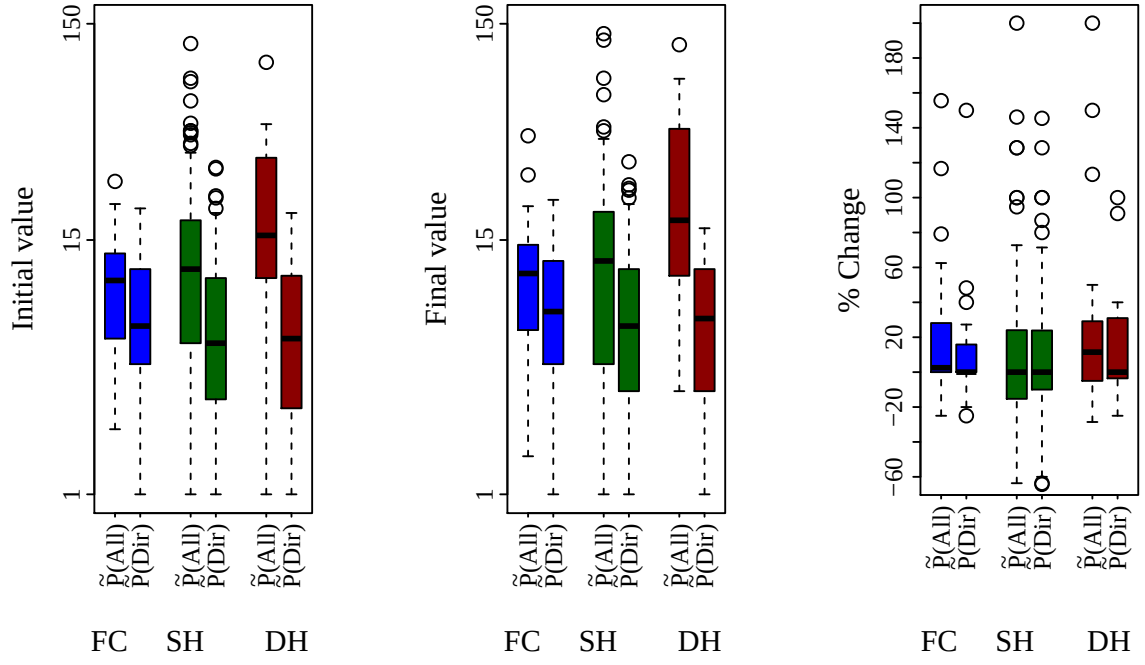
Table 8.1: Summary of redundancy metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\tilde{P}$(All)	(5.1)(5.3) $(P_{ij} \in \mathcal{C}(\mathcal{S}_{ij}))$
	SD_{ij}	$\tilde{P}$(Dir)	(5.1)(5.3) $(P_{ij} \in SD_{ij})$
Left-tail centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\tilde{P}^{25\%}$(All)	(5.6) $(P_{ij} \in \mathcal{C}(\mathcal{S}_{ij}))$
	SD_{ij}	$\tilde{P}^{25\%}$(Dir)	(5.6) $(P_{ij} \in SD_{ij})$

The case study distributions of $\tilde{P}$(All) and $\tilde{P}$(Dir) as well as $\tilde{P}^{25\%}$(All) and $\tilde{P}^{25\%}$(Dir) are compared to their theoretical counterparts (repeated from Section 5.3) in Figures 8.5 and 8.6, respectively. For this and the following section on overlap metrics, it is important to note that $SD_{ij} \not\equiv \mathcal{C}(\mathcal{S}_{ij})$ for the FC archetype. Therefore there were two distributions reflected in the graphs. This was contrary to the theoretical instances where $SD_{ij} \equiv \mathcal{C}(\mathcal{S}_{ij})$ and resulted from the bi-directionality deviation discussed in Section 7.2.4.

Distributions of $\tilde{P}$(All) and $\tilde{P}$(Dir)

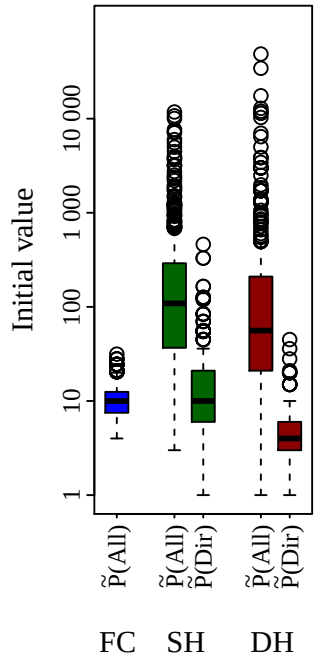
In Figure 8.5 the top three graphs display three measures of $\tilde{P}$(All) and $\tilde{P}$(Dir) for the case study instances. The initial values (Figure 8.5a) were measured in the undisturbed case study instances. The final values (Figure 8.5b) were measured right before instances became disconnected. The % change measured the difference between the initial and final values. The three bottom graphs reflect the same measures observed for the theoretical instances. Three observations were made when comparing the case study and theoretical distributions.



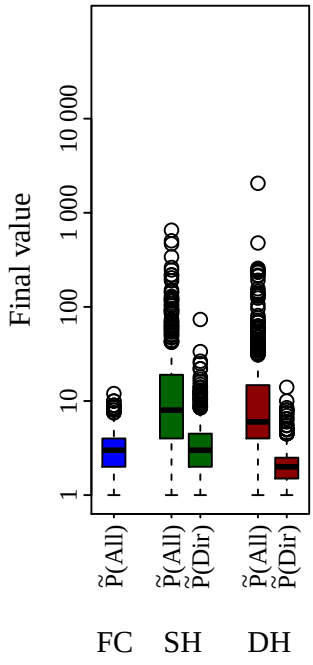
(a) Case Study: Initial value

(b) Case Study: Final value

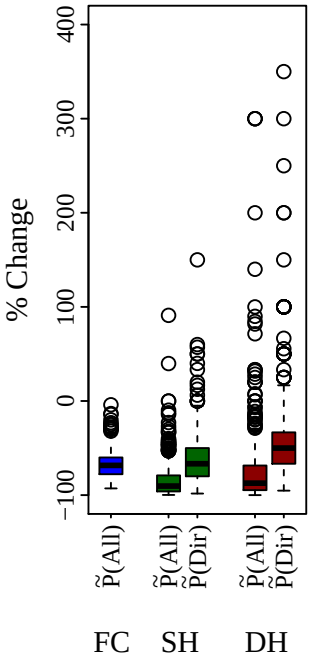
(c) Case Study: % Change



(d) Theoretical: Initial value



(e) Theoretical: Final value



(f) Theoretical: % Change

Figure 8.5: Case Study vs Theoretical: Distributions of the three measures of $\tilde{P}$(All) and $\tilde{P}$(Dir).

Distribution range: The theoretical range of the initial and final distributions was two orders of magnitude larger than that of the case study range. In addition the case study distributions were tighter with fewer outliers. These results concurred with

what was observed regarding $\sum_{i,j \in SD_{ij}; i \neq j} P_{ij}$ and $\sum_{i,j \in \mathcal{S}_{ij}; i \neq j} P_{ij}$ that the case study instances started out with remarkably fewer shortest path alternatives. These however, were summative metrics. The vulnerability metrics regarded aggregate measures of redundancy. This controlled for the impact that the smaller G^{1K} layers had and highlighted the effect of the lower density of G^2 and its deviation from the bi-directional grid structure. Furthermore, the final values of the theoretical instances indicated that even right before disconnection, the shortest path collections of these instances were still orders of magnitude larger than the case study instances.

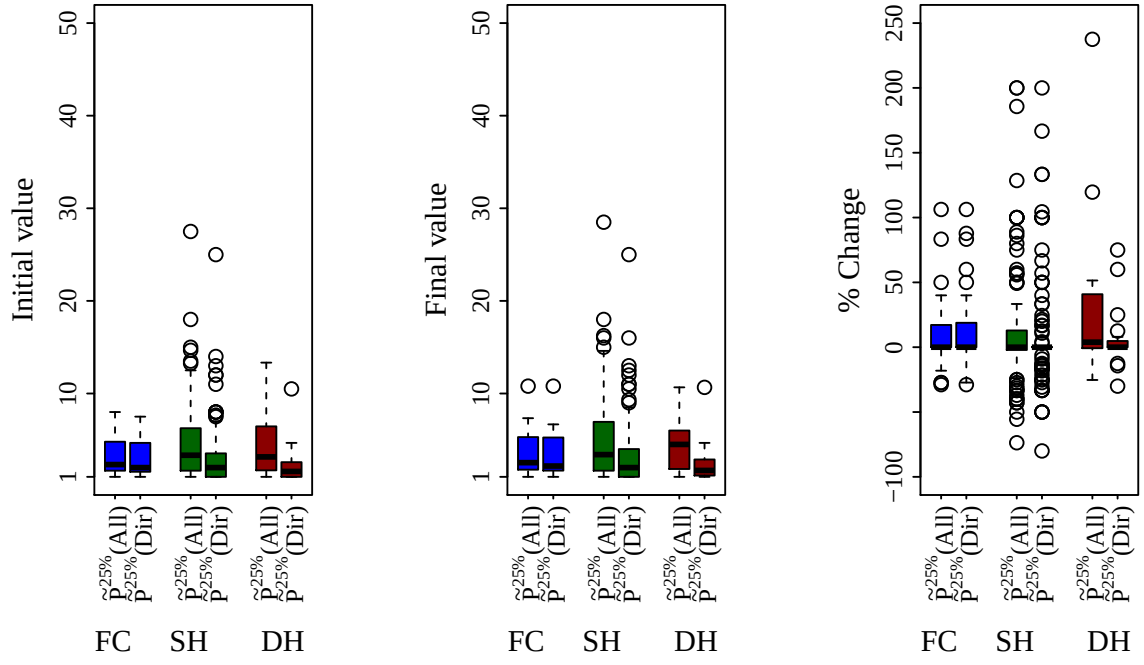
Ordering of archetypes: The ordering of the case study distributions was not greatly different to that of the theoretical distributions. Generally the hub archetypes had greater redundancy than the FC archetype. In the theoretical instances the SH archetype had more alternatives than the DH archetype. This changed ever so slightly in the case study instances were the DH archetype took the lead.

The geographic span of the case study instances influenced their redundancy scores. The larger area an instance spanned, the more of the road network it covered and thus the more detours were potentially available for inclusion in the shortest path sets. A Spearman's correlation test confirmed that $\tilde{P}$(All) had a strongly positive monotonic correlation to the geographic span ($\rho = 0.69$; p-value$< 2.2 \times 10^{-16}$). This correlation explained why the DH archetype had higher redundancy scores as its geographic spans were far greater than those of the SH archetype. In the theoretical instances, archetypes also had varying spans but all three were bound by the 10×10 grid and therefore the scope for variation was limited.

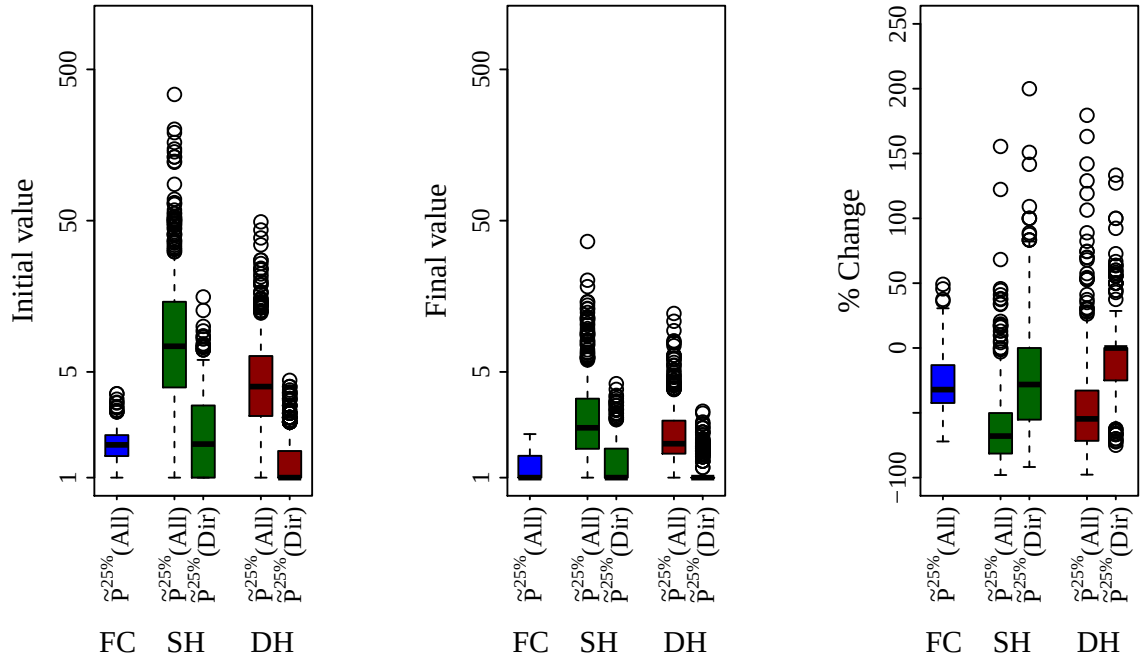
Minimal change from initial to final values: The change between the initial and final case study instances was almost imperceptible. The means of the % change in Figure 8.5c hovered around 0% with the exception of $\tilde{P}$(All) for the DH instances. This was compared to the % change in the theoretical instances which showed an average reduction between 50% and 100%. The case study instances simply became disconnected so quickly that there was hardly opportunity to whittle away at the shortest path sets.

Distributions of $\tilde{P}^{25\%}$(All) and $\tilde{P}^{25\%}$(Dir)

When zooming in to the least redundant instances, there was less difference between the theoretical and case study distributions.



(a) Case Study: Initial value (b) Case Study: Final value (c) Case Study: % Change



(d) Theoretical: Initial value (e) Theoretical: Final value (f) Theoretical: % Change

Figure 8.6: Case Study vs Theoretical: Distributions of the three measurements of $\tilde{P}^{25\%}$(All) and $\tilde{P}^{25\%}$(Dir).

Distribution range: Although the initial redundancy of the theoretical instances (Figure 8.6d) was still greater than that of the case study instances (Figure 8.6a), the bulk of the instances now fell within a more comparable range. There was even

greater similarity between the final distributions (Figures 8.6b and 8.6e) as the redundancy of the theoretical instances was greatly reduced while the case study instances remained almost static.

One notable difference was that the theoretical distributions had many more outliers. This was owed to the combinatorial nature of the shortest paths on the theoretical grid. Because the road networks in the case study instances were not bi-directional grids, there was less opportunity for such a combinatorial explosion in shortest path alternatives.

Minimal change from initial to final values: The change between the initial and final case study instances was almost imperceptible, which was contrary to the marked reduction in the theoretical instances. Again this was due to the fact that the case study instances became disconnected so rapidly.

Correlation to efficiency loss and robustness

In Chapter 6 statistical tests were executed to test the correlation of the vulnerability metrics to efficiency loss and robustness. Those tests were repeated for the case study instances. However, because there was so little change in the redundancy metrics from their initial values to their final values, the tests were only conducted for the initial values.

In the theoretical instances most measures of the redundancy metrics showed a significant negative correlation to both efficiency loss and robustness (see Tables 6.1 to 6.3). When repeating the correlation tests for the case study instances, only two relations were significant. In the FC archetype, $\tilde{P}^{25\%}$(All) was negatively correlated to efficiency loss whereas in the SH archetype it was negatively correlated to robustness. Although significant, these correlations were weak.

In summary the case study instances were remarkably less redundant than the theoretical instances. There was also less significant correlations of the redundancy metrics to efficiency loss and robustness. This combined with the fact that redundancy remained practically unchanged up until disconnection aligned with the findings of the theoretical instances. Redundancy was not a stand-alone indicator of vulnerability in the case study instances under random disruptions.

8.1.4 Overlap

Overlap was defined in Section 5.4 as the degree to which the shortest path sets of an instance had road segments in common. We used the concept of relative link betweenness to quantify the fraction of shortest paths that a specific road segment featured on. A distinction was made between the link betweenness when taking into account only the directly connected node-pairs and when taking into account all node-pairs. The former was called the Elemental link betweenness (Elemental-B) (4.2) score and $\boldsymbol{B}_{elemental}$ was defined as the set of Elemental-B scores for links in G^2 in descending order. The Overall link betweenness (Overall-B) (4.1) score was the latter and defined the link betweenness based on the complete collection of shortest paths. Thus, $\boldsymbol{B}_{overall}$ was defined as the set of Overall-B scores for links in G^2 in descending order.

There were two aspects of the the distributions of $\boldsymbol{B}_{elemental}$ and $\boldsymbol{B}_{overall}$ that were of interest. Firstly, the centrality of the distribution was an aggregate indication of the degree of shortest path overlap in an instance. The second aspect sought to quantify how dependent the shortest paths of a single instance were on a critical few road segments.

The road segments that were shared the most by the shortest paths had the highest betweenness scores and were found in the right tail of the distribution. A longer right tail indicated that a few links had scores *remarkably* higher than the majority of the other road links. Therefore, a larger right-tail range indicated the presence of a few critical links that, when removed, would have a dire impact on the instance's connectivity. The range of the right-tail, measured from the 75^{th} percentile to the maximum value, was the metric defined to quantify this aspect. The summary of the aspects and their metrics from Section 5.4 is repeated below in Table 8.2 for convenience.

Table 8.2: Summary of overlap metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\overline{B}_{\text{overall}}$	(5.7)
	SD_{ij}	$\overline{B}_{\text{elemental}}$	(5.8)
Right-tail range	$\mathcal{C}(\mathcal{S}_{ij})$	$R(B_{\text{overall}})^{75\%}$	(5.11)
	SD_{ij}	$R(B_{\text{elemental}})^{75\%}$	(5.12)

The case study distributions of $\overline{B}_{\text{overall}}$ and $\overline{B}_{\text{elemental}}$ as well as $R(B_{\text{overall}})^{75\%}$ and $R(B_{\text{elemental}})^{75\%}$ are compared to their theoretical counterparts (repeated from Section 5.4) in Figures 8.7 and 8.8, respectively.

Distributions of $\overline{B}_{\text{overall}}$ and $\overline{B}_{\text{elemental}}$

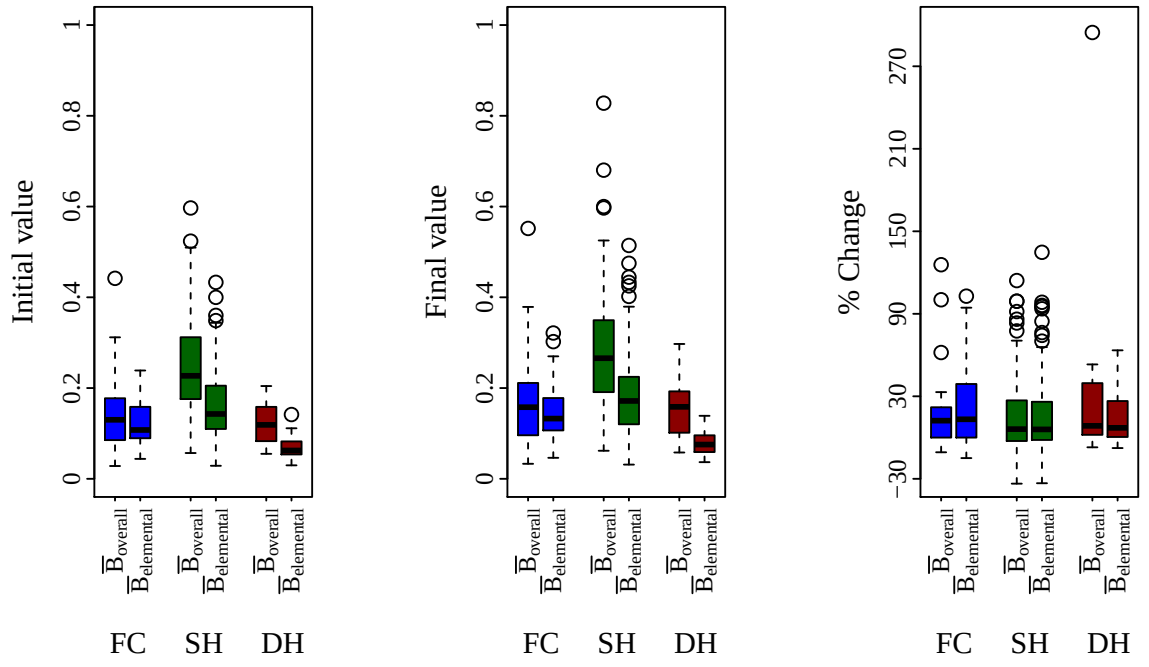
There existed a significant relationship between redundancy and overlap in the case study instances. If an instance had many shortest path alternatives, it would also have had higher values for $\tilde{P}$(All) and $\tilde{P}$(Dir). This would have increased the denominator when calculating Overall-B or Elemental-B, thus reducing the betweenness scores. The Spearman correlation test confirmed that there was a strong and significant negative relationship between the centrality metrics of redundancy and overlap both initially and just before instances became disconnected (see Table 8.3).

Table 8.3: Correlation between redundancy and overlap in the case study instances.

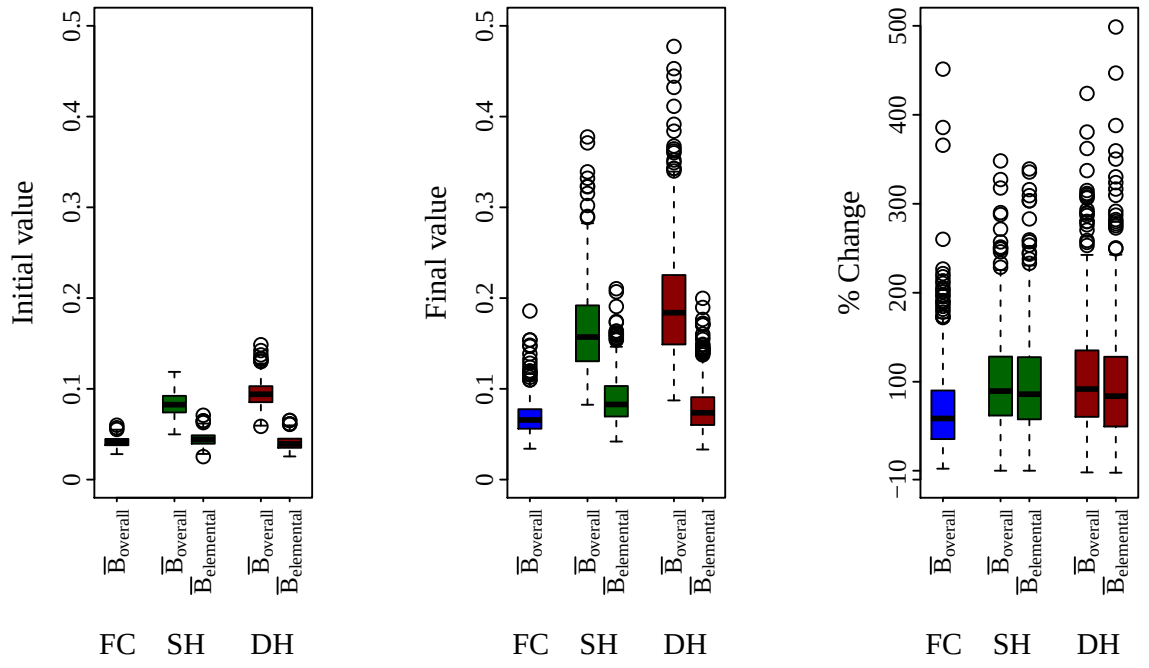
x	**y**	$\boldsymbol{\rho}$	**p-value**
Initial values			
$\tilde{P}$(All)	$\overline{B}_{\text{overall}}$	-0.52	9.97×10^{-15}
$\tilde{P}$(Dir)	$\overline{B}_{\text{elemental}}$	-0.64	$< 2.2 \times 10^{-16}$
Final values			
$\tilde{P}$(All)	$\overline{B}_{\text{overall}}$	-0.47	4.3×10^{12}
$\tilde{P}$(Dir)	$\overline{B}_{\text{elemental}}$	-0.59	$< 2.2 \times 10^{-16}$

The top three graphs of Figure 8.7 plot the distributions for the initial values, final values and % change of the case study instances, respectively. The theoretical counterparts

are repeated from Section 5.4 in the bottom three graphs.



(a) Case Study: Initial value (b) Case Study: Final value (c) Case Study: % Change



(d) Theoretical: Initial value (e) Theoretical: Final value (f) Theoretical: % Change

Figure 8.7: Case Study vs Theoretical: Distributions of the three measurements of $\overline{B}_{\text{overall}}$ and $\overline{B}_{\text{elemental}}$.

Distribution range: The case study distributions showed much greater overlap than

the theoretical instances initially (Figures 8.7a and 8.7d). The negative correlation between redundancy and overlap already offered one explanation for this observation. Because the case studies had much smaller shortest path collections, the link betweenness of any road segment featuring in one or more paths was greater than it would have been if the collections were as big as those of the theoretical instances.

Another pertinent explanation for the higher overlap observed in the case study instances was that overlap was induced by the shortest path algorithm. The algorithm started with an absolute shortest path that served as a "core path". Each iteration made only a small perturbation to this core path to search for alternatives. Therefore, the alternatives were in a way "anchored" to the initial core path. As a result most of the shortest paths in the final shortest path set included segments of the core path. The road segments that constituted the core path would thus have featured in many paths, inflating their link betweenness. On the flip-side, road segments that were not in the core path would have featured on far fewer shortest paths and thus have had a very low link betweenness.

In addition to the general elevation in overlap, the case study instances also had broader initial distributions, implying that overlap was more instance specific. This made sense given the significant variation between the case study instances.

Overlap increased markedly in the theoretical instances (Figures 8.7e) due to the pertinent reduction in the number of shortest path alternatives before disconnection. There was thus more similarity between the final values of the case study instances (Figures 8.7b) and the theoretical instances, although the case study instances still had higher overlap overall.

Ordering of archetypes: The correlation between redundancy and overlap then also explained why in the case study distributions the DH archetype had lower distributions than the SH archetype when the opposite was true in the theoretical instances. The DH instances had higher redundancy because of their greater diagonal span (refer to the discussion in Section 8.1.3) and therefore less overlap.

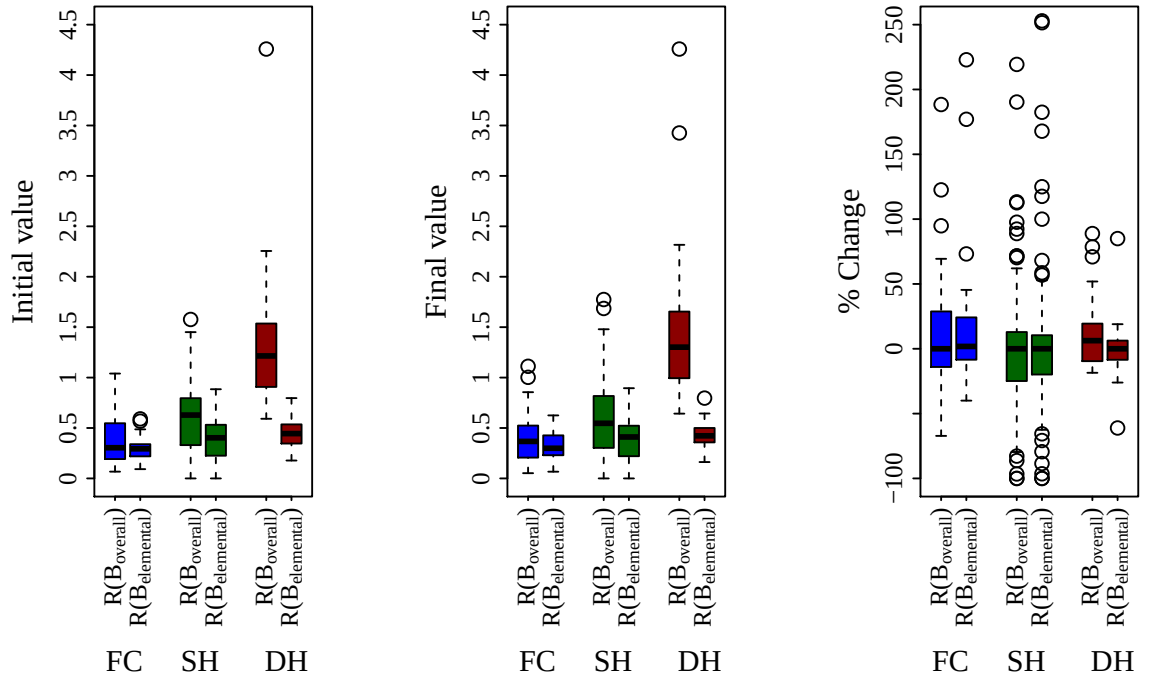
Minimal change from initial to final values: As was the case with the redundancy metrics, the change between the initial and final values of the case study instances were negligible as the instances became disconnected so rapidly.

Distributions of $R(B_{\mathbf{overall}})^{75\%}$ and $R(B_{\mathbf{elemental}})^{75\%}$

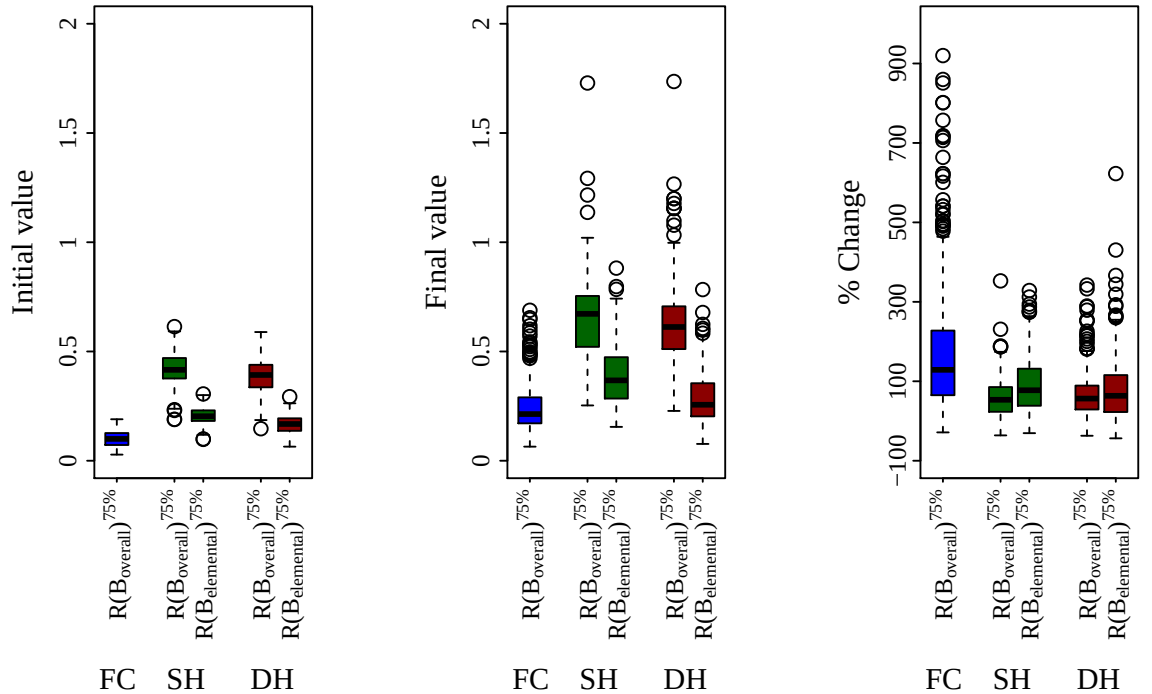
The range of the right tail of the distribution of overlap scores gave an indication of how dependent the shortest path sets were on a few critical road segments. The range was measured as the difference between the road segment with the highest betweenness score and the 75$^{\text{th}}$ percentile value. Keep this in mind when considering the y-axes in Figure 8.8. Those values do not indicate the absolute betweenness score of instances, but the distance between their most between road segments and the majority of the road segments.

Section 5.4.3 explained that if a road segment had a betweenness value greater than 1, it implied that it featured more than once in one or many of the shortest paths. This was called "doubling back". Depending on the degree of overlap in an instance, it was therefore possible to have instances where the most between road segments had a score much higher than one. It was also possible that the range of the right tails in those instances would then be greater than one. This was evident already in the final values of the theoretical

instances (Figure 8.8e) where the $R(B_{\text{overall}})^{75\%}$ scores of some instances exceeded one. The two notable differences between the theoretical and case study instances were the magnitude of the ranges and the ordering of the archetypes.



(a) Case Study: Initial value (b) Case Study: Final value (c) Case Study: % Change



(d) Theoretical: Initial value (e) Theoretical: Final value (f) Theoretical: % Change

Figure 8.8: Case Study vs Theoretical: Distributions of the three measurements of $R(B_{\text{overall}})^{75\%}$ and $R(B_{\text{elemental}})^{75\%}$.

Distribution range: The ranges of the right tails were much greater in the case study instances than the theoretical instances. Thus, the case study instances were more dependent on their most critical links.

The doubling back phenomenon was only possible in instances that had indirectly connected node-pairs. For that specific reason it was not observed in the theoretical FC instances. Conversely, the magnitude of the right tail ranges in the case study's FC archetype suggests that there were some instances where doubling back occurred.

Ordering of archetypes: When evaluating the centrality of the overlap distributions for the case study instances, the DH archetype had lower overlap than the SH archetype overall. Here, however, when considering the right tails, the DH archetype had notably larger ranges. This owed to the structure of the logical layer of the DH archetype. A number of the indirectly connected node-pairs had three logical links from one node to another, making it possible for road segments to feature up to three times on one shortest path. This made it possible for the most between road segments to have even higher scores, resulting in a larger right tail range.

Minimal change from initial to final values: Again there was very little change between the initial and final values. This indicated that even for the instances that survived a few disturbances, the impact on the importance of their most between road segments was inconsequential.

Correlation to efficiency loss and robustness

The correlation between the initial values of the overlap metrics to efficiency loss and robustness was determined. Again it didn't make sense to measure correlation for the % change and final values as the change from the initial to final values was negligible. In the theoretical instances the initial values of the overlap metrics had insignificant or very weak correlations to efficiency loss and robustness. Meanwhile, the % change and final values showed strongly positive correlations.

In the case study instances the DH archetype showed a strongly positive correlation between $\overline{B}_{\text{overall}}$ and efficiency loss. This suggested that the same phenomena was at play in the case studies. Instances with higher overlap were more concentrated, covering less of the road network. Therefore random disruptions were less likely to affect their shortest path sets. Unfortunately, the correlations for the FC and SH archetypes were insignificant or too weak to be compared to the theoretical instances.

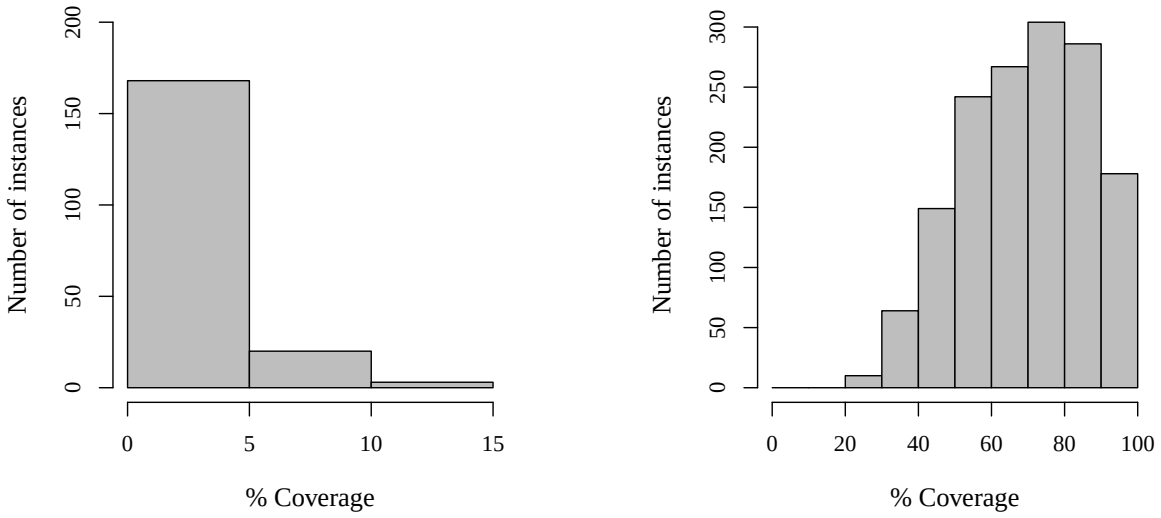
In summary the case study instances displayed a greater degree of overlap in their shortest path sets. The significant correlations of the overlap metrics to efficiency loss and robustness combined with the fact that overlap remained practically unchanged up until disconnection echoed the theoretical findings. Overlap was not a stand-alone indicator of vulnerability in the case study instances under random disruptions.

8.1.5 Correlation between G^2 coverage, redundancy and overlap

The realisation that redundancy and overlap did not capture vulnerability as expected was first noted in Chapter 6 when conducting the statistical tests for the theoretical instances. There it was found that the degree to which the shortest paths covered the 10×10 grid was correlated to redundancy, overlap and the damage of the instances. We concluded

that the likelihood of the random simulation selecting grid links in the shortest paths had an influence on vulnerability.

The degree to which the shortest path sets covered the road network in the case study instances was determined. Figure 8.9 shows that the coverage in the case study instances was not even comparable to the theoretical instances. The minimum coverage of a theoretical instance was greater than 20%, meanwhile the maximum coverage of a case study instance was less than 15%. Notwithstanding, the correlation of coverage to specific response variables was remarkably alike.



(a) Case study instances

(b) Theoretical instances

Figure 8.9: Comparison of G^2 coverage in the case study and theoretical instances.

Similar to the theoretical instances, the coverage held a significant, strongly positive correlation to redundancy and a significant, strongly negative correlation to overlap in all three archetypes. The more shortest paths an instance had, the greater its coverage of the underlying road network. Whereas the greater the overlap of these alternative paths, the more concentrated they were and the smaller the coverage of the road network.

The coverage also held significant negative correlation to robustness in the FC and SH archetypes. The more of the road network was covered, the sooner it became disconnected. However this relation was weak. In the case of the DH archetype, the correlation was insignificant.

Unlike the theoretical instances, significant correlation of coverage to efficiency loss could not be established.

Table 8.4 lists the correlation values of the case study instances next to those of the theoretical instances to enable comparison.

Table 8.4: Comparison of Spearman's correlation (ρ) of the coverage of G^2 to redundancy, overlap, efficiency loss and robustness in the theoretical and case study instances. Insignificant correlations (p-value < 0.05) are indicated by *insig.*

Response variable	Archetype	Theoretical	Case study
$\tilde{P}$(All)	FC	0.54	0.82
	SH	0.68	0.71
	DH	0.64	*insig.*
$\overline{B}_{\text{overall}}$	FC	-0.34	-0.53
	SH	-0.21	-0.61
	DH	-0.42	-0.55
Efficiency Loss	FC	-0.21	*insig.*
	SH	-0.30	*insig.*
	DH	-0.30	*insig.*
Robustness	FC	-0.18	-0.35
	SH	-0.24	-0.35
	DH	-0.15	*insig.*

8.1.6 Efficiency Step-Change

In Section 5.5 the concept of efficiency step-change was defined. The change in the average shortest path length (L_{ij}) of a specific node-pair between two successive disruptions was considered the step-change of that node-pair (5.13). We then converted the step-change of that node-pair to a relative figure by dividing it by the shortest path length before the disruption (5.14). To aggregate all these relative step-changes for a specific instance, we divided their sum by the number of node-pairs (5.15). There was thus only one metric defined to measure the efficiency step-change, as indicated in Table 8.5 repeated from Section 5.5.

Table 8.5: Summary of efficiency step-change metrics.

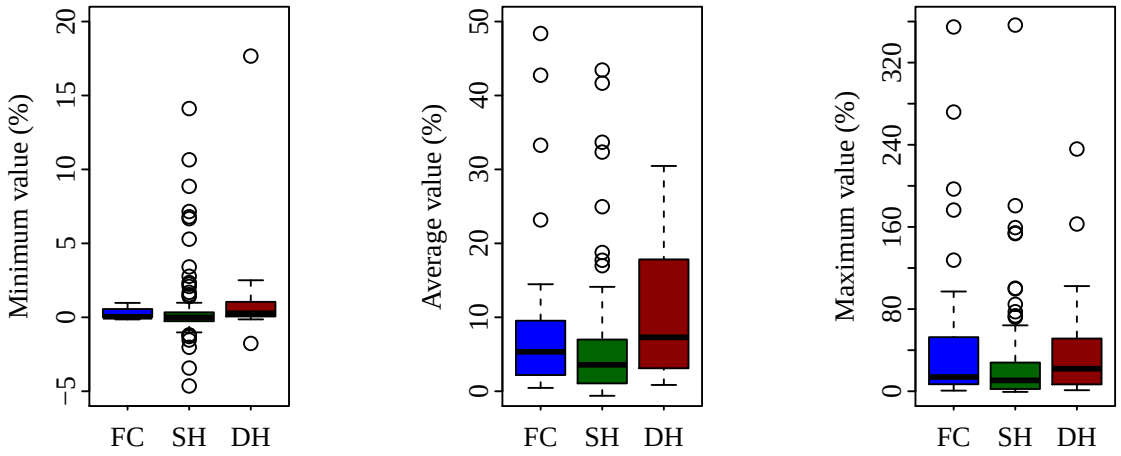
Aspect	Scope	Metric	Equation
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$Rel\overline{\Delta L}(t; t+z)$	(5.15)

Distributions of $Rel\overline{\Delta L}(t; t+1)$

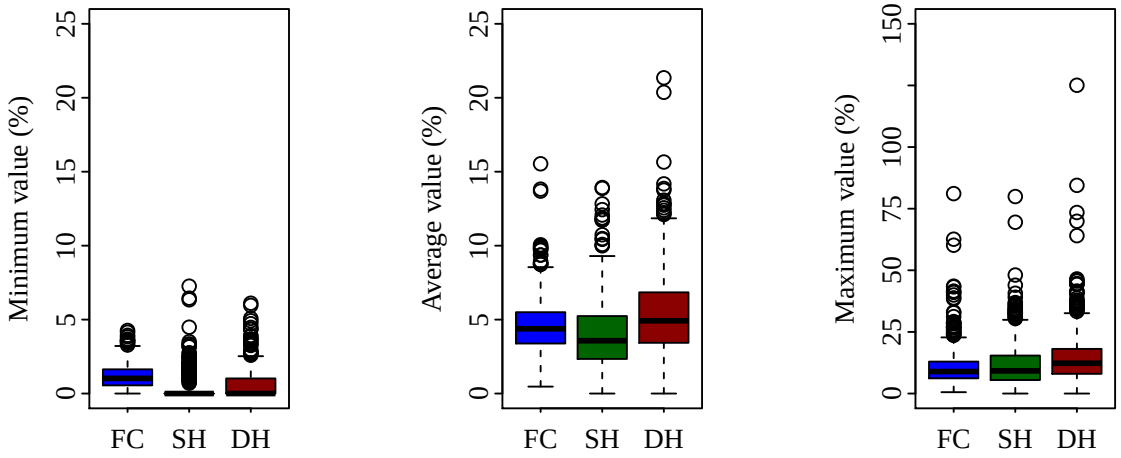
Measurements of $Rel\overline{\Delta L}(t; t+1)$ were different to the measurements of the redundancy and overlap metrics. Instead of considering the initial value, final value and the % change between the two, the minimum (Figures 8.10a and 8.10d), average (Figures 8.10b and 8.10e) and maximum values (Figures 8.10c and 8.10f) of $Rel\overline{\Delta L}(t; t+1)$ were determined over the time series from the initial to the final network.

In the theoretical instances it was impossible for the shortest paths to become shorter after a disruption. Thus, the minimum values of $Rel\overline{\Delta L}(t; t+1)$ were always greater than

or equal to zero (Figure 8.10d). In the case study instances, L_{ij} was really the average of all the shortest paths identified that fell within the stated tolerance of 125%. It was thus possible that a disruption could break one or more of the longer paths included in the shortest path set resulting in a lower value for L_{ij}. If this negative step-change was relatively large or occurred for a number of node-pairs, it resulted in a negative value for $Rel\overline{\Delta L}(t;t+1)$. In Figure 8.10a negative values were observed for a few instances of both the SH and DH archetypes.



(a) Case Study: Minimum value (b) Case Study: Average value (c) Case Study: Maximum value



(d) Theoretical: Minimum value (e) Theoretical: Average value (f) Theoretical: Maximum value

Figure 8.10: Case Study vs Theoretical: Distributions of the three measurements of $Rel\overline{\Delta L}(t;t+1)$.

The fact that the case study instances had higher average values (Figures 8.10b) could also be ascribed to the shortest path sets containing alternatives of differing lengths. Unlike with the theoretical instances, it was possible for L_{ij} to change before the entire shortest path set was empty. In fact, it was highly probable that every time an alternative path was removed from the shortest path set, L_{ij} would change. On the contrary, in the

theoretical instances, L_{ij} only increased when an entire shortest path set had been emptied and replaced by a new shortest path set. Therefore, step-changes were more frequent in the case study instances, leading to slightly higher average values.

The case study instances also exhibited higher maximum values (Figure 8.10c) than the theoretical instances (Figure 8.10f). This time the observation was attributed to the bi-directional grid of the theoretical instances. The uniform link length, density of the grid and geographic bound made it less likely for step-change to be as large as in the case study instances.

Correlation to efficiency loss and robustness

In the case of the efficiency step-change, the correlation of the minimum, average and maximum values to efficiency loss and robustness was determined. Here the case study results echoed that of the theoretical instances. The average and maximum values showed strongly positive correlations to robustness for the FC archetype and to efficiency loss and robustness for the hub archetypes.

In summary the efficiency step-changes were more frequent and larger than those experienced by the theoretical instances. This was primarily a by-product of the shortest path algorithm and the characteristics of G^2. Again the strong correlations were intuitive, but it was not clear that efficiency step-change could predict disconnection or efficiency loss.

Overall the theoretical instances showed much stronger correlation of the vulnerability metrics to efficiency loss and robustness. Despite these correlations, the vulnerability metrics held no discriminatory power in the theoretical instances. In light of the correlation results from the case study instances, repeating the tests to determine discriminatory power seemed futile.

8.2 Validity of the vulnerability metrics determined from real-life data

The most prominent insight from repeating the random error simulation for the case study instances was that in practice these instances were much more vulnerable. They became disconnected quickly because they had relatively few alternative paths and a less dense road network. Notwithstanding, the instances experienced the same magnitude of efficiency loss as the theoretical instances before disconnection.

A question was raised regarding the validity of how efficiency loss was defined and measured. The heterogenous nature of the shortest path sets in the case study instances occasionally resulted in efficiency *gains*. This occurred when the alternative paths that were longer than the core path, but fell within the tolerance, were removed. Considering the real-life implications we asserted that even though the metric showed an efficiency *gain*, it was not realistic as the core path remained the same length. It was suggested that the formulation of efficiency loss is perhaps not the most valid for practice where acceptable detours are not exactly the same length. An alternative proposition to measuring efficiency loss was to only reflect the length of the core path in L_{ij}, even if the shortest path set contains other longer detours.

The patterns observed and behaviour of redundancy, overlap and efficiency step-change were validly explained phenomena in the case study instances. The differences between

the distributions of the theoretical instances and the case study instances were discussed. These differences sprung primarily from the fact that the case study instances themselves were not exact replicas of the theoretical instances — as described in Chapter 7. Regardless of these differences, we confirmed that the metrics as formulated still captured the concept of redundancy, overlap and efficiency step-change as defined in Chapter 5 when it came to real-life data.

Finally, we considered the correlation tests. The case study instances failed to yield as many significant results as the theoretical instances did. The first reason was that only the initial values could be used for redundancy and overlap as these distributions did not change much before disconnection. The second and third reasons were presumably the smaller sample sizes and the heterogeneity among the case study instances. However, the few significant results that were observed were congruent with Chapter 6. Not one of the vulnerability metrics was a stand-alone indicator of vulnerability. In fact, it would seem that vulnerability itself was multi-faceted and that key influential factors not captured by the vulnerability metrics (such as road network coverage) are yet to be explored.

Chapter 9

Conclusion and future work

Today's supply chains face increasing volatility on many fronts. From the shop-floor where machines break and suppliers fail to the boardrooms where unanticipated price inflation erodes profitability. Turbulence is the new normal.

To remain competitive and weather these (daily) storms, supply chains need to move away from an efficiency mindset towards a resilience mindset. For over a little more than a decade industry and academia have awakened to this reality. Academic literature and case studies show that there is no longer a shortage of resilience strategies and designs. Unfortunately, industry still lacks the tools with which to assess and evaluate the effectiveness of such strategies and designs. Without the ability to quantify the benefit it is impossible to motivate the cost.

This thesis aimed to add one piece to the puzzle of quantifying supply chain vulnerability. Specifically, it focussed on supply chains within urban areas. It sought to quantify to what degree an urban supply chain's network design (internal configuration) and its dependence on the underlying road network (external circumstances) made it more or less vulnerable to disruptions.

Multilayered Complex Network Theory (CNT) held promise as a modelling approach that could capture the complexity of the dependence between a *logical* supply chain network and the *physical* road network that underpins it. Such an approach addressed two research gaps in complex network theory applications. In the supply chain arena CNT applications have reaped many benefits but the majority of studies regarded single-layer networks that model only supply chain relations. There were no studies found where the dependence of supply chain layers on underlying physical infrastructure was modelled in a multilayered manner. Road network applications offered many more multilayered applications but these primarily focussed on passenger transport, not freight transport.

The thesis statement was formulated as follows:

> *Metrics related to the shortest path sets of the multilayered supply chain/road network formulation can quantify the inherent vulnerability of a specific supply chain to its choice of internal configuration and the underlying urban road network's integrity.*

In order to evaluate this statement the following objectives were achieved:

1. Development of a multilayered complex network model that captured the dependence of a supply chain network on an urban road network.

2. Identification of the characteristics of this model that described the nature of the supply chain's vulnerability to the integrity of the urban road network.

3. Development of metrics that could quantify a supply chain's inherent vulnerability based on its internal configuration and the underlying road network.

4. Evaluation of the validity of the suite of vulnerability metrics through statistical analysis and a real-life case study.

These objectives were achieved using a *design research* methodology as described by Manson (2006). Design research is

> "a process of using knowledge to design and create useful artefacts, and then using various rigorous methods to analyse why, or why not, a particular artefact is effective" (Manson, 2006).

Two artefacts were developed namely a multilayered complex network formulation of a supply chain's dependence on its underlying road network, and a suite of vulnerability metrics that were proposed to quantify the inherent vulnerability of the supply chain. The artefacts were evaluated in two ways. Using the multilayer formulation we generated large samples for three different supply chain network archetypes. The distributions of the topological characteristics were investigated and compared to verify the formulation. We then used random error simulations and statistical correlation tests to assess the performance of the suite of vulnerability metrics. Further feedback of the artefacts' utility was obtained when these were applied to a case study of three South African urban areas. Feedback from the case study added to the operation and goal knowledge obtained through the study.

The design of the artefacts itself, although novel, was not considered research but it was through the insights derived during analysis of the artefacts' performance that we contributed to the body of knowledge.

9.1 Key findings from the thesis

The multilayered network formulation $\mathcal{M}$ is a useful and intuitive artefact to capture the dependence between a supply chain's internal configuration and the road network that underpins it. It is capable of combining both the *logical* and *physical* connectedness of the multilayered system in a manner that enables quantitative experimentation and analyses. Strict assumptions were applied in the generation of the theoretical instances. Later, in the case study instances the evidence was overwhelming that these strict assumptions do not hold in practice. Nonetheless, the formulation of $\mathcal{M}$ was still a useful and intuitive artefact in studying the case study instances.

To develop the suite of vulnerability metrics, three targeted attack simulations were executed on the theoretical instances. These simulations isolated the characteristics of the the collection of shortest path sets ($\mathcal{C}(\mathcal{S}_{ij})$) believed to be most critical to network vulnerability. From these simulations the concept of network skeletons was ruled out as an indicator of vulnerability. The multilayered nature of $\mathcal{M}$ dilutes skeleton structures that are present in the individual network layers. Instead, the role played by grid links in $\mathcal{C}(\mathcal{S}_{ij})$ in terms of betweenness is a far more convincing indication of vulnerability.

Disturbances in practice are neither fully targeted nor completely random. Rather, these are somewhere on the continuum between targeted and random. There existed no data that could characterise exactly *where* on this continuum disturbances should be plotted. Thus, using completely random disturbances to assess the performance of the suite of vulnerability metrics was regarded as sufficiently conservative.

The damage caused by the random disturbances in terms of efficiency loss and disconnection were indifferent to the network archetype. Under random disturbances the internal configuration of the supply chain network has very little bearing on when efficiency will be lost or facilities will become disconnected. This finding is contrary to the accepted knowledge in single-layer CNT application which states that hub archetypes are more robust to random errors. However, under targeted attack the hub archetypes were far more vulnerable than the Fully Connected (FC) archetype. Here the observation corroborated what is known from single-layer studies.

A definite plot twist occurred when testing the correlation of the vulnerability metrics to efficiency loss and robustness. Although more than one metric was strongly correlated to efficiency loss and/or robustness for redundancy and overlap, the direction of the correlations were unexpected. Vulnerability under random link disturbances is not a straight-forward product of redundancy and overlap. Instead it is multi-faceted and the probability of removing a link that features in shortest path holds pertinent influence. The initial values of the vulnerability metrics were also surprisingly uncorrelated. Therefore, it is impossible to gauge the inherent vulnerability of a network looking only at the initial, undisturbed network.

The thesis continued with a case study of three urban areas in South Africa. The intention was to test the validity of the developed artefacts when applied to real-life data. Many noteworthy findings emerged regarding the structure of supply chain neighbourhoods in urban areas.

Firstly, it was found that the three theoretical archetypes used in the thesis are prevalent in practice, accounting for between a fifth and a quarter of the population of supply chain neighbourhoods in the case study. However, the remainder of the supply chain neighbourhoods were of *mixed type*. The prevalence of the *mixed type* in practice was viewed from two perspectives. The first perspective acknowledges that reality simply does not pan out as we plan. The partners in a supply chain may agree to design their network according to certain philosophies. Unfortunately, when the rubber hits the road it is unlikely that the interactions between the facilities, at least in terms of freight movement, will obey these academic designs. The second perspective posits that supply chain network designs are really emergent and innovative — impervious to the enforcement of theoretical archetypes.

Secondly, it was found that there is little difference between the size and structure of the supply chain neighbourhoods across the three urban areas. Supply chain neighbourhoods in Gauteng Province (GT), City of Cape Town (CoCT) and eThekwini Metropolitan Municipality (ET) seem indifferent to the varying geographies, demographics and economic activities present in these areas.

Thirdly, supply chain neighbourhoods in practice very seldomly if ever exhibit the simplifying theoretical assumptions used to generate the theoretical instances. Nonetheless, the formulation of $\mathcal{M}$ is still valid and useful in modelling the real-life instances as it is capable of describing any conceivable *logical* network layer when layered on a *physical* network layer.

Moving on to the road network layer, the fourth finding was that although the road network shows impressive regularity, it is not a bi-directional grid as it is often modelled in literature. The degree distributions are homogenous but do not mimic that of a bi-directional grid. More importantly though is that the road network has a much lower density in reality than the bi-directional grid suggests.

A link-based random error simulation performed on the case study instances revealed a

number of worthwhile findings. Firstly, the rate at which instances become disconnected in practice is indifferent to the archetypes. This corroborated the earlier finding regarding the theoretical instances. Therefore, under random link disruption the internal configuration of the supply chain has little bearing on how quickly it will become disconnected.

Secondly, it was noted that the way in which efficiency loss was measured was invalid for real-life instances. As soon as shortest paths were not exactly the same length (as is the case in practice), this measure misrepresents the damage to the network in terms of the average shortest path length.

Thirdly, the case study instances were far more vulnerable than the theoretical instances. The fact that these instances were also far less redundant and had much greater degrees of overlap anticipated this result. However, as was the case in the theoretical instances, statistical tests again proved that neither redundancy nor overlap were stand-alone indicators of vulnerability. Again it was found that the coverage of the road network by the shortest paths played a pertinent role in quantifying vulnerability under random disruptions. Therefore, in practice vulnerability itself is multi-faceted and key influential factors not captured by the vulnerability metrics (such as road network coverage) are yet to be explored.

In light of this summary of key findings, a reflection on the performance of the artefacts is presented to close the *design research* loop.

9.2 Reflection on the performance of the artefacts

The multilayered network formulation $\mathcal{M}$ is a valid and useful representation of a supply chain's dependence on its underlying road network. In both the theoretical and case study instances it enabled the quantification of the relationship to allow experimentation and analysis. In this first version of the artefact a simplifying assumption was made that the logical layer is unweighted. Certainly this could be refined in future iterations. The capacity of the underlying road network could also somehow be reflected in the physical layer. A road segment with multiple lanes would be less likely to fail completely than a single-lane road segment.

The suite of vulnerability metrics were derived from the results of targeted attack simulations on theoretical instances. This approach thoughtfully identified three aspects of vulnerability namely redundancy, overlap and efficiency step-change. Both the theoretical and case study instances show that these concepts were adequately modelled by the metrics.

Unfortunately, under a random disruption strategy this suite of vulnerability metrics fails to be a robust predictor or even quantifier of vulnerability. This is primarily owed to the fact that random disturbances are the most unpredictable. Much of an instance's vulnerability comes down to "luck of the draw". This "luck of the draw" is linked to the degree to which the shortest paths of an instance covers the road network. Therefore, under random disruption vulnerability is multi-faceted and not adequately covered by the suite of metrics.

Despite the poor performance of the vulnerability metrics under random disturbances, the rationale behind these metrics is sound. There remains a strong possibility that these metrics will be good quantifiers and predictors of vulnerability when the disruption simulation is less random — as is the case in practice. Three suggestions are made to refine this artefact: The first is to increase the ambit of the suite and include the coverage of G^2 as a vulnerability aspect. The second is to test for multivariate correlations and

covariances to better understand the interplay between metrics. The final suggestion is to test these metrics using a more realistic disturbance simulation.

9.3 Research contribution and limitations

9.3.1 Research contribution

According to Hevner et al. (2004) effective design research must contribute in three ways: the development of a design artefact that addresses an unsolved problem; the expansion of the knowledge foundation in a domain through creative development of (evaluated) constructs; and the creative development of evaluation methodologies.

Therefore, this thesis contributes to research in the following ways:

Artefact development: The formulation of the multilayered complex network model captures the dependence of a supply chain on underlying infrastructure. The vulnerability metrics that emanate from this formulation capture the redundancy, overlap and efficiency step-change of the multilayered network.

The vulnerability metrics were developed based on the performance of theoretical instances under targeted disturbances. However, in the case study it was shown that real-life supply chains are very different and respond very differently compared to their theoretical counterparts. An unexpected contribution by this thesis is thus the suggestion that future research in this field start with empirical networks sourced from practice and develop artefacts based on reality as opposed to first studying the performance of theoretical artefacts.

Foundational knowledge: Novel insights and perspectives were gained regarding the influence of a supply chain's internal configuration and the integrity of the road network on its inherent vulnerability. Furthermore, the case study analysis offers a rich description of supply chain neighbourhoods present in three urban areas in South Africa. These descriptions are also novel and are unanticipated research contributions made by the thesis.

Evaluation methodologies: The thesis used purely targeted disturbances to derive the vulnerability metrics and then switched to purely random disturbances to test the performance of the vulnerability metrics. The confounding performance of the vulnerability metrics under random disruptions highlights that it is crucial to develop a simulation strategy that more closely reflects reality to use when testing the vulnerability metrics. Road network disruptions are neither completely random nor specifically targeted. Important segments with greater traffic loads are more *likely* to be disrupted, but the reality is that disruptions such as accidents, equipment failure or road maintenance could really occur anywhere on the network. Therefore, a blended disruption strategy on the continuum between targeted and random disruptions would be more valid in evaluating the vulnerability metrics. The thesis thus contributes to research by showing that the typical approaches used in CNT vulnerability studies are not sufficient to solve this problem. One cannot model road disruptions as exclusively targeted or purely random. A more tailored approach is required.

This study adds its voice to ongoing research that seeks to develop methods and models that quantify supply chain vulnerability. By improving the quality of quantitative

information on the topic, better trade-off decisions can be made between supply chain resilience and efficiency.

9.3.2 Limitations of the thesis

Supply chain vulnerability drivers can emanate from a number of sources. As described in Section 1.1.1, Peck (2005) categorised these drivers into four levels. This thesis focussed on only one driver from the second level of vulnerability drivers namely asset and infrastructure dependencies. Within this level it focussed only on the vulnerability of a supply chain as induced by its dependence on underlying transport infrastructure. The study was further confined to supply chain networks within urban areas that used only road transport. This exact scope provides but one piece to a much larger supply chain vulnerability puzzle and the findings should be appreciated within that context.

The developed artefacts are subject to four significant simplifying assumptions:

- Links are unweighted in both layers of the multilayered networks. Real-life road networks could be weighted in terms of capacity or traffic density while supply chains could be weighted in terms of freight volumes or even number of shipments. Shudong et al. (2012) and Zadeh and Rajabi (2013) illustrated the importance of using weighted links for more accurate prioritisation of critical links.

- The 10×10 grid chosen to represent the *physical* layer is small. Even so the computational burden of calculating the collections of shortest paths was prohibitive. Concerns are raised that such a small *physical* layer induces edge effects that could either mask or distort results.

- The samples of 500 randomly generated network instances per archetype are not negligible. However, the population of possible network instances could be orders of magnitude bigger, depending on the similarity constraints imposed. Calculating theoretical upper bounds would strengthen claims of representation.

- Shortest paths are not always the routes of choice. This model assumes that commercial vehicles would choose the shortest path in terms of distance. In practice routes that are longer in terms of distance could be shorter in terms of travel time or longer routes could be chosen for other reasons. In addition elements such as elevation and gradient could add to the perceived distance of a route.

- Cascading phenomena of road network failures are disregarded. Failure on one road segment can easily have spill-over effects to other road segments (Feng et al., 2017; Shudong et al., 2012).

- The road network does not recover. In the simulations we assume that there is no post-disruption rewiring but that disrupted road segments remain disrupted. This is certainly not the case in practice.

- The overall connectedness of the networks was defined purely in terms of the shortest path sets of node-pairs. It did not consider truck routes and how these routes would combine sets of facilities. Incorporating this into the formulation could have implications for metrics and measurements based on efficiency.

These simplifications were necessary to put forward a first iteration of the thesis artefacts. In their comparison between richer systems-based road vulnerability metrics and the simplified topology-based metrics, Dehghani et al. (2014) illustrated how a simpler approach can still yield very useful results. That was the case in this thesis as well.

Limitations regarding the representativity of the case study data must also be noted:

- The freight movement data is both commodity and vehicle "blind". Supply chains could have been better characterised and even categorised according to economic sector if this were not the case.
- The data covers only one month namely February 2014. It is debatable whether freight movement activities in this one month are representative of the whole year.
- The data only covers three urban areas in one country. This is certainly not representative of urban areas worldwide. Furthermore, the study does not control for different geographies, demographics and economy structures and the influences these might have on supply chain structure.
- The road network was filtered according to assumptions of the road type that would be used. Map matching data for the freight movements were not available to corroborate these assumptions.
- The shortest path sets were determined by a custom algorithm. This algorithm does not guarantee the identification of all allowbale alternative paths.

The key findings, reflection on the artefacts and the limitations mentioned above give rise to a pipeline of future work that could further expand this research.

9.4 Future work

This thesis has lain sturdy groundwork for the quantification of supply chain vulnerability as it relates to the urban road network. Three parallel streams of future work are envisioned:

Refinement of the artefact formulations: The first suggested refinement is the incorporation of link weights into both the formulation of $\mathcal{M}$ and the vulnerability metrics. In the logical layer link weight would denote the strength of the relationship. In the physical layers it would denote capacity of the road segment. Incorporating the capacity does not imply that traffic flow should be included. That would shift this research from a topology-based study to a systems-based study, which is not the intention. Instead, knowing the capacity of a road segment may make it more or less prone to disruptions. Link weights, especially the logical link weights, should also be incorporated into the vulnerability metrics.

The second suggested refinement is the expansion of the suite of vulnerability metrics to account for the "luck of the draw" as described earlier in this chapter. The probability of a disruption affecting the shortest paths of an instance at all seems to have notable bearing on its vulnerability. The coverage of G^2 is one possible metric that could account for this, but this topic requires further exploration.

The third suggestion is to expand the library of supply chain network archetypes beyond the FC, Single Hub (SH) and Double Hub (DH) archetypes. From the case

study it was evident that at least three quarters of the supply chain neighbourhoods are of *mixed type*. These types need to be characterised and represented in the formulation of $\mathcal{M}$ in order to make it more representative.

Expansion of the theoretical model sizes and samples: With more efficient computational techniques experiments should be repeated on theoretical models with much larger *physical* layers. This would control for edge effects that could have influenced the results in this thesis. The size of the samples of network instances used should also be determined taking cognisance of theoretical upper bounds to ensure that they are large enough to be representative. With a greater variety of theoretical models (in terms of the *physical* layer size) and larger samples, asymptotic behaviour of the collection of shortest paths could be studied.

Expansion and refinement of the case study: The insights drawn from the case study show that real-life networks behave very differently to their theoretical counterparts, primarily because they do not obey the simplifying constraints imposed on theoretical models. We believe future research could be much more expedient if analyses focus on building theory from real networks instead. The data exists within the Centre for Transport Development at the University of Pretoria to expand the case study over years and across many more urban areas. This would yield larger samples for experimentation and analysis.

If the focus of this research now shifts to developing theory from real-life networks, then the refinement of the shortest path algorithm is a priority. The current algorithm was custom designed to build a shortest path set given a core path and an allowable tolerance. This approach does not guarantee that a comprehensive set of paths is determined. Furthermore, the algorithm produces paths that are strongly anchored to the core path. It could be possible to find more diverse paths using other approaches.

Another consideration in refining the shortest path algorithm is to assess whether "least kilometres travelled" really translates to the preferred path in practice. It may make more sense to define the shortest paths in terms of travel time or number of turns.

Refinement of the disruption simulation and statistical analysis: In this first iteration, a link-based random disturbance simulation was used as it was regarded the most conservative method to test performance. Admittedly, this approach does not simulate how road disruptions occur in practice. Road network disruptions are neither completely random nor specifically targeted. Important segments with greater traffic loads are more *likely* to be disrupted, but the reality is that disruptions such as accidents, equipment failure or road maintenance could really occur anywhere on the network. Further research is required to develop a topology-based disruption strategy that lies between targeted and random disruptions and more closely approximates everyday road disruptions.

Again, if the focus shifts to theory-building from real-life networks, then the manner in which the first level of damage, *efficiency loss*, is measured should be revised. The current metric produces a slight misrepresentation when given heterogenous shortest path sets.

Finally, the statistical analysis could be refined and expanded. The current single variate correlation analysis can be upgraded to multivariate analysis. The Kolmogorov-Smirnov test (KS-test) could also be replaced by a more refined approach to test the similarity of distributions.

Bibliography

Agigi, A., Niemann, W., and Kotzé, T. (2016). Supply chain design approaches for supply chain resilience: A qualitative study of South African fast-moving consumer goods grocery manufacturers. *Journal of Transport and Supply Chain Management*, 10(1):1–15.

Albert, R. and Barabási, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics*, 74(1):47–97.

Albert, R., Jeong, H., and Barabási, A.-L. (2000). Error and attack tolerance of complex networks. *Nature*, 406(6794):378–381.

Anderson, P. (1999). Complexity theory and organization science. *Organization Science*, 10(3):216–232.

Barabási, A. and Albert, R. (1999). Emergence of scaling in random networks. *Science*, 286(5439):509–512.

Barthélemy, M. (2011). Spatial networks. *Physics Reports*, 499(1–3):1–101.

Barthélemy, M. and Flammini, A. (2008). Modeling urban street patterns. *Physical Review Letters*, 100(13):138702.

Basole, R. C. and Bellamy, M. A. (2014). Supply network structure, visibility, and risk diffusion: A computational approach. *Decision Sciences*, 45(4):753–789.

Basole, R. C., Bellamy, M. A., Park, H., and Putrevu, J. (2016). Computational analysis and visualization of global supply network risks. *IEEE Transactions on Industrial Informatics*, 12(3):1206–1213.

Bellamy, M. and Basole, C. (2013). Network analysis of supply chain systems: A systematic review and future research. *Systems Engineering*, 16(2):235–249.

Berche, B., Von Ferber, C., Holovatch, T., and Holovatch, Y. (2009). Resilience of public transport networks against attacks. *The European Physical Journal B*, 71(1):125–137.

Berche, B., Von Ferber, C., Holovatch, T., and Holovatch, Y. (2012). Transportation network stability: A case study of city transit. *Advances in Complex Systems*, 15(1):1250063.

Bhatia, G., Lane, C., and Wain, A. (2013). *Building Resilience in Supply Chains*. Risk Response Network. World Economic Forum.

Boccaletti, S., Bianconi, G., Criado, R., Del Genio, C. I., Gómez-Gardeñes, J., Romance, M., Señdina Nadal, I., Wang, Z., and Zanin, M. (2014). The structure and dynamics of multilayer networks. *Physics Reports*, 544(1):1–122.

Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.-U. (2006). Complex networks: Structure and dynamics. *Physics Reports*, 424(4–5):175–308.

Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., García del Amo, A., and Boccaletti, S. (2013). Modeling the multi-layer nature of the European air transport network: Resilience and passengers re-scheduling under random failures. *The European Physical Journal Special Topics*, 215(1):23–33.

Chacon, N., Doherty, S., Hayashi, C., Green, R., and Lever, I. (2012). *New Models for Addressing Supply Chain and Transport Risk*. Risk Response Network. World Economic Forum.

Choi, T., Dooley, K., and Rungtusanatham, M. (2001). Supply networks and complex adaptive systems: Control versus emergence. *Journal of Operations Management*, 19(3):351–366.

Choi, T. Y. and Hong, Y. (2002). Unveiling the structure of supply networks: Case studies in Honda, Acura, and DaimlerChrysler. *Journal of Operations Management*, 20(5):469–493.

Christopher, M. and Holweg, M. (2011). "Supply Chain 2.0": Managing supply chains in the era of turbulence. *International Journal of Physical Distribution & Logistics Management*, 41(1):63–82.

Christopher, M. and Holweg, M. (2017). Supply Chain 2.0 revisited: A framework for managing volatility — induced risk in the supply chain. *International Journal of Physical Distribution & Logistics Management*, 47(1):2–17.

Christopher, M. and Peck, H. (2004). Building the resilient supply chain. *The International Journal of Logistics Management*, 15(2):1–14.

Cohen, R., Erez, K., Avraham, D., and Havlin, S. (2000). Resilience of the internet to random breakdowns. *Physical Review Letters*, 85(21):4626–4628.

Cohen, R., Erez, K., Avraham, D., and Havlin, S. (2001). Breakdown of the internet under intentional attack. *Physical Review Letters*, 86(16):3682.

Costa, L. d. F., Rodrigues, F., Travieso, G., and Villas Boas, P. (2007). Characterization of complex networks: A survey of measurements. *Advances in Physics*, 56(1):167–242.

Crucitti, P., Latora, V., and Porta, S. (2006). Centrality measures in spatial networks of urban streets. *Physical Review E*, 73(3):036125.

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

Danziger, M. M., Bashan, A., Berezin, Y., Shekhtman, L. M., and Havlin, S. (2014). An introduction to interdependent networks. *International Conference on Nonlinear Dynamics of Electronic Systems*, CCIS 438:189–202.

Dehghani, M., Flintsch, G., and McNeil, S. (2014). Impact of road conditions and disruption uncertainties on network vulnerability. *Journal of Infrastructure Systems*, 20(3):04014015.

Demšar, U., Špatenková, O., and Virrantaus, K. (2008). Identifying critical locations in a spatial network with graph theory. *Transactions in GIS*, 12(1):61–82.

Dorogovtsev, S. N. and Mendes, J. F. F. (2002). Evolution of networks. *Advances in Physics*, 51(4):1079–1187.

Duan, Y. and Lu, F. (2014). Robustness of city road networks at different granularities. *Physica A*, 411:21–34.

Ducruet, C. (2013). Network diversity and maritime flows. *Journal of Transport Geography*, 30:77–88.

Ducruet, C., Rozenblat, C., and Zaidi, F. (2010). Ports in multi-level maritime networks: evidence from the Atlantic (1996–2006). *Journal of Transport Geography*, 18:508–518.

Erdös, P. and Rényi, A. (1959). On random graphs. *Publicationes Mathematicae Debrecen*, 6:290–297.

Feng, Y., Sun, B., and Zeng, A. (2017). Cascade of links in complex networks. *Physical Letters A*, 381(4):263–269.

Fortunato, S. (2010). Community detection in graphs. *Physics Reports*, 486(3–5):75–174.

Gallotti, R. and Barthelemy, M. (2014). Anatomy and efficiency of urban multimodal mobility. *Scientific Reports*, 4:6911.

Gallotti, R. and Barthelemy, M. (2015). The multilayer temporal network of public transport in Great Britain. *Scientific Data*, 2:140056.

Gauteng Online (2017). The Economy of Gauteng. Available online at http://www.gautengonline.gov.za/Business/Pages/TheEconomyofGauteng.aspx (accessed 4 August).

Girvan, M. and Newman, M. (2002). Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12):7821–7826.

Grady, D., Thiemann, C., and Brockman, D. (2012). Robust classification of salient links in complex networks. *Nature Communications*, 3:864.

Graser, A., Straub, M., and Dragaschnig, M. (2013). Towards an open source analysis toolbox for street network comparison: Indicators, tools and results of a comparison of OSM and the official Austrian Reference Graph. *Transactions in GIS*, 18(4):510–526.

He, S., Li, S., and Ma, H. (2009). Effect of edge removal on topological and functional robustness of complex networks. *Physica A*, 388:2243–2253.

Hearnshaw, E. and Wilson, M. (2013). A complex network approach to supply chain network theory. *International Journal of Operations & Production Management*, 33(4):442–469.

Heckmann, I., Comes, T., and Nickel, S. (2015). A critical review on supply chain risk — definition, measure and modeling. *Omega*, 52:119–132.

Hevner, A. R., March, S. T., Park, J., and Ram, S. (2004). Design science in information systems research. *Management Information Systems Quarterly*, 28(1):75–105.

Hohenstein, N.-O., Feisel, E., Hartmann, E., and Giunipero, L. (2015). Research on the phenomenon of supply chain resilience: A systematic review and paths for further investigation. *International Journal of Physical Distribution & Logistics Management*, 45:90–117.

Holme, P. (2002). Edge overload breakdown in evolving networks. *Physical Review E*, 66(3):036119.

Hu, Y. and Zhu, D. (2009). Empirical analysis of the worldwide maritime transportation network. *Physica A*, 388:2061–2071.

Jiang, B. (2007). A topological pattern of urban street networks: Universality and peculiarity. *Physica A*, 384:647–655.

Jiang, B. and Claramunt, C. (2004). Topological analysis of urban street networks. *Environment and Planning B*, 31(1):151–162.

Joubert, J. and Axhausen, K. W. (2013). A complex network approach to understand commercial vehicle movement. *Transportation*, 40(3):729–750.

Joubert, J. W. and Axhausen, K. W. (2011). Inferring commercial vehicle activities in Gauteng, South Africa. *Journal of Transport Geography*, 19:115–124.

Joubert, J. W. and Meintjes, S. (2015a). Computational considerations in building inter-firm networks. *Transportation*, 42(5):857–878.

Joubert, J. W. and Meintjes, S. (2015b). Repeatability & reproducibility: Implications of using GPS data for freight activity chains. *Transportation Research Part B*, 76:81–92.

Joubert, J. W. and Viljoen, N. M. (2017). Multilayer complex networks, v3. *Mendeley Data*. Available online from `http://dx.doi.org/10.17632/268byhmvv5.3`.

Kaluza, P., Kölzsch, A., Gastner, M. T., and Blasius, B. (2010). The complex network of global cargo ship movements. *Journal of the Royal Society Interface*, 7:1093–1103.

Kamalahmadi, M. and Parast, M. (2016). A review of the literature on the principles of enterprise and supply chain resilience: Major findings and directions for future research. *International Journal of Production Economics*, 171:116–133.

Kilubi, I. (2016a). Investigating current paradigms in supply chain risk management — a bibliometric study. *Business Process Management Journal*, 22(4):662–692.

Kilubi, I. (2016b). The strategies of supply chain risk management — a synthesis and classification. *International Journal of Logistics Research and Applications: A Leading Journal of Supply Chain Management*, 19(6):604–629.

Kim, Y., Choi, T. Y., Yan, T., and Dooley, K. (2011). Structural investigation of supply networks: A social network analysis approach. *Journal of Operations Management*, 29(3):194–211.

Kivelä, M., Arenas, A., Barthelemy, M., Gleeson, J. P., Moreno, Y., and Porter, M. A. (2014). Multilayer networks. *Journal of Complex Networks*, 2(3):203–271.

Klibi, W. and Martel, A. (2012). Modeling approaches for the design of resilient supply networks under disruptions. *International Journal of Production Economics*, 135:882–898.

Kurant, M. and Thiran, P. (2006a). Extraction and analysis of traffic and topologies of transportation networks. *Physical Review E*, 74(3):036114.

Kurant, M. and Thiran, P. (2006b). Layered complex networks. *Physical review letters*, 96(13):138701.

Lee, K.-M., Min, B., and Goh, K.-I. (2015). Towards real-world complexity: and introduction to multiplex networks. *European Physical Journal B*, 88(2):48.

Lordan, O., Sallan, J., Simo, P., and Gonzalez-Prieto, D. (2015). Robustness of airline alliance route networks. *Communications in Nonlinear Science and Numerical Simulation*, 22(1–3):587–595.

Lotero, L., Cadillo, A., Hurtado, R., and Gómez-Gardeñes, J. (2016). Several multiplexes in the same city: The role of socioeconomic differences in urban mobility. In Garas, A., editor, *Interconnected Networks*, chapter 9, pages 149–168. Springer International Publishing.

Manson, N. (2006). Is operations research really research? *ORiON*, 22(2):155–180.

March, J. G. and Shapira, Z. (1987). Managerial perspectives on risk and risk taking. *Management Science*, 33(11):1404–1418.

Masucci, A., Smith, D., Crooks, A., and Batty, M. (2009). Random planar graphs and the London street network. *European Physical Journal B*, 71(2):259.

Mattsson, L.-G. and Jenelius, E. (2015). Vulnerability and resilience of transport systems — a discussion of recent research. *Transportation Research Part A*, 81:16–34.

Motter, A. E., Nishikawa, T., and Lai, Y.-C. (2002). Range-based attack on links in scale-free networks: Are long-range links responsible for the small-world phenomenon? *Physical Review E*, 66(6):065103.

Nair, A. and Vidal, J. M. (2011). Supply network topology and robustness against disruptions — an investigation using multi-agent model. *International Journal of Production Research*, 49(5):1391–1404.

Newman, M. (2003). The structure and function of complex networks. *SIAM Review*, 45(2):167–256.

Nie, T., Guo, Z., Zhao, K., and Lu, Z. (2015). New attack strategies for complex networks. *Physica A*, 424:248–253.

OpenStreetMap contributors (2017). Planet dump retrieved from https://planet.osm.org. `https://www.openstreetmap.org`.

Ortigosa, J. and Menendez, M. (2014). Traffic performance on quasi-grid urban structures. *Cities*, 36:18–27.

Parshani, R., Rozenblat, C., Ietri, D., Ducruet, C., and Havlin, S. (2011). Inter-similarity between coupled networks. *Europhysics Letters*, 92(6):68002.

Pastoras-Satorras, R. and Vespignani, A. (2001). Epidemic spreading in scale-free networks. *Physical Review Letters*, 86(14):3200.

Pathak, S., Day, J., Nair, A., Sawaya, W., and Kristal, M. (2007). Complexity and adaptivity in supply networks: Building supply network theory using a complex adaptive systems perspective. *Decision Sciences*, 38(4):547–580.

Peck, H. (2005). Drivers of supply chain vulnerability: An integrated framework. *International Journal of Physical Distribution & Logistics Management*, 35(4):210–232.

Peck, H. (2006). Reconciling supply chain vulnerability, risk and supply chain management. *International Journal of Logistics: Research and Applications*, 9(2):127–142.

Porta, S., Crucitti, P., and Latora, V. (2006). The network analysis of urban streets: A dual approach. *Physica A*, 369(2):853–866.

Porta, S., Latora, V., Wang, F., Rueda, S., Strano, E., Scellato, S., Cardillo, A., Belli, E., Cárdenas, F., Cormenzana, B., and Latora, L. (2012). Street centrality and the location of economic activities in Barcelona. *Urban Studies*, 49(7):1471–1488.

Purdy, G. (2010). ISO 31000:2009 — Setting a new standard for risk management. *Risk Analysis*, 30(6):881–886.

Rao, S. and Goldsby, T. (2009). Supply chain risk: A review and typology. *The Interantional Journal of Logistics Management*, 20(1):97–123.

Rice, J. B. and Caniato, F. (2003). Building a secure and resilient supply network. *Supply Chain Management Review*, 7(5):22–30.

Salehi, M., Sharma, R., Marzolla, M., Magnani, M., Siyari, P., and Montesi, D. (2015). Spreading processes in multilayer networks. *IEEE Transactions on Network Science and Engineering*, 2(2):65–83.

Sen, P., Dasgupta, S., Chatterjee, A., Sreeram, P., Mukherjee, G., and Manna, S. (2003). Small-world properties of the Indian railway network. *Physical Review E*, 67(3):036106.

Serrano, M. A., Boguná, M., and Vespignani, A. (2009). Extracting the multiscale backbone of complex weighted networks. *Proceedings of the National Academy of Sciences of the United States of America*, 106(16):6483–6488.

Sheffi, Y. and Rice, J. B. (2005). A supply chain view of the resilient enterprise. *MIT Sloan Management Review*, 47(1):41–48.

Shekhtman, L. M., Bagrow, J. P., and Brockmann, D. (2014). Robustness of skeletons and salient features in networks. *Journal of Complex Networks*, 2(2):110–120.

Shudong, L., Lixiang, L., Yixian, Y., and Qun, L. (2012). Revealing the process of edge-based-attack cascading failures. *Nonlinear Dynamics*, 69(3):837–845.

Solé-Ribalta, A., Gómez, S., and Arenas, A. (2016). Congestion induced by the structure of multiplex networks. *Physical Review Letters*, 116(10):108701.

Statistics South Africa (2016). Community Survey 2016 Provinces at a glance. Technical Report ISBN 9780621446661, Statistics South Africa, Pretoria, South Africa.

Strano, E., Cardillo, A., Iacoviello, V., Latora, V., Messora, R., Porta, S., and Scellato, S. (2009). Street centrality vs. commerce and service locations in cities: A kernel density correlation case study in Bologna, Italy. *Environment and Planning B*, 36(3):450–465.

Strano, E., Shai, S., Dobson, S., and Barthélemy, M. (2015). Multiplex networks in metropolitan areas: generic features and local effects. *Journal of the Royal Society Interface*, 12(111):20150651.

Tang, C. (2006). Robust strategies for mitigating supply chain disruptions. *International Journal of Logistics: Research and Applications*, 9(1):33–45.

Thadakamalla, H., Raghavan, U., Kumara, S., and Albert, R. (2004). Survivability of multiagent-based supply networks: A topological perspective. *IEEE Intelligent Systems*, 19(5):24–31.

Tomko, M., Winter, S., and Claramunt, C. (2008). Experiential hierarchies of streets. *Computers, Environment and Urban Systems*, 32(1):41–52.

Travieso, G. and da Fontoura Costa, L. (2012). Evaluating links through spectral decomposition. *Journal of Statistical Mechanics: Theory and Experiment*, 2012(P01015).

Tsiotas, D. and Polyzos, S. (2015). Decomposing multilayer transportation networks using complex network analysis: A case study for the Greek aviation network. *Journal of Complex Networks*, 3(4):642–670.

Tukamuhabwa, B. R., Stevenson, M., Busby, J., and Zorzini, M. (2015). Supply chain resilience: Definition, review and theoretical foundations for further study. *International Journal of Production Research*, 53(18):5592–5623.

United Nations, Department of Economic and Social Affairs, Population Division (2015). World Urbanization Prospects: The 2014 Revision. Technical Report ST/ESA/SER.A/366, United Nations, New York.

Van Heerden, Q. and Joubert, J. W. (2014). Generating intra and inter-provincial commercial vehicle activity chains. *Procedia - Social and Behavioral Sciences*, 125:136–146.

Viljoen, N. M. and Joubert, J. W. (2016). The vulnerability of the global container shipping network to targeted link disruption. *Physica A*, 462:396–409.

Viljoen, N. M. and Joubert, J. W. (2017). Multilayered complex network datasets for three supply chain archetypes on an urban road grid. Working paper 062, Centre for Transport Development, University of Pretoria.

Wagner, S. and Bode, C. (2006). An empirical investigation into supply chain vulnerability. *Journal of Purchasing and Supply Management*, 12(6):301–312.

Wagner, S. and Neshat, N. (2010). Assessing the vulnerability of supply chains using graph theory. *International Journal of Production Economics*, 126:121–129.

Watts, D. and Strogatz, D. (1998). Collective dynamics of small-world networks. *Nature*, 393(6684):440.

Wu, T., Huang, S., Blackhurst, J., Zhang, X., and Wang, S. (2013). Supply chain risk management: An agent-based simulation to study the impact of retail stockouts. *IEEE Transactions on Engineering Management*, 60(4):676–686.

Zadeh, A. S. M. and Rajabi, M. A. (2013). Analyzing the effect of the street network configuration on the efficiency of an urban transportation system. *Cities*, 31:285–297.

Zhang, G.-Q., Wang, D., and Li, G.-J. (2007). Enhancing the transmission efficiency by edge deletion in scale-free networks. *Physical Review E*, 76(1):017101.

Zhao, K., Kumar, A., Harrison, T. P., and Yen, J. (2011). Analyzing the resilience of complex supply network topologies against random and targeted disruptions. *IEEE Systems Journal*, 5(1):28–39.

Zhuo, Y., Peng, Y., Liu, C., Liu, Y., and Long, K. (2011). Traffic dynamics on layered complex networks. *Physica A*, 390:2401–2407.

Zsidisin, G., Ellram, L., Carter, J., and Cavinato, J. (2004). An analysis of supply risk assessment techniques. *International Journal of Physical Distribution & Logistics Management*, 34(5):397–413.

Zsidisin, G., Panelli, A., and Upton, R. (2000). Purchasing organization involvement in risk assessments, contingency plans, and risk management: An explanatory study. *Supply Chain Management: An International Journal*, 5(4):187–198.

Appendix A

Summary of mathematical formulations

A.1 Glossary

Table A.1: Glossary of indices

Indices	Description
α, β	Indices of network layers in generic formulation
b, c	Indices of network layers in generic formulation
i, j	Node indices in the logical network G^{1K}
n	Index of ordered sets $\boldsymbol{P}^{25\%}$, $\boldsymbol{B}_{overall}$ and $\boldsymbol{B}_{elemental}$
s, t	Node indices in the physical network G^2
u, v	Node indices of the betweenness networks G^{γ} and G^{ς}
K	Index of supply chain network archetypes

Table A.2: Glossary of mathematical symbols and elements with reference to defining equations.

Symbol	Description	Eq. ref
$A^{[1K,2]}$	Adjacency matrix of interlayer connections	(A.29)
$a_{is}^{1K,2}$	Elements of the adjacency matrix of interlayer connections	(A.30)
$\boldsymbol{B}_{elemental}$	Set of decreasing Elemental-B(e_{st}^2) scores	(A.60)
$\overline{B}_{\text{elemental}}$	Mean of $\boldsymbol{B}_{elemental}$	(A.63)
$\boldsymbol{B}_{overall}$	Set of decreasing Overall-B(e_{st}^2) scores	(A.61)
$\overline{B}_{\text{overall}}$	Mean of $\boldsymbol{B}_{overall}$	(A.62)
B_n	Elements of the ordered set $\boldsymbol{B}_{overall}$ or $\boldsymbol{B}_{elemental}$	(A.63) (A.62)
$B_{elemental}^{75\%}$	Value of the cut-off of the 75$^{\text{th}}$ percentile of $\boldsymbol{B}_{elemental}$	(A.65)
$B_{overall}^{75\%}$	Value of the cut-off of the 75$^{\text{th}}$ percentile of $\boldsymbol{B}_{elemental}$	(A.64)

Table A.3: Glossary of mathematical symbols and elements with reference to defining equations (continued).

Symbol	Description	Eq. ref
$\mathcal{C}$	Set of interconnections of $\mathcal{M}$	(A.5) (A.28)
$\mathcal{C}(\mathcal{S}_{ij})$	Collection of shortest path sets for all node-pairs	(A.34)
$c_{st}(i,j)$	Binary consensus variable to determine if e^2_{st} features in the $\mathcal{S}_{ij}$	(A.51) (A.52)
E_α	Intralayer link set of G_α	(A.9)
E_β	Intralayer link set of G_β	(A.13)
$E_{\alpha,\beta}$	Interlayer link set of G_α and G_β	(A.14)
E_b	Intralayer link set of G_b	(A.4)
E^{1K}	Intralayer link set of logical layer G^{1K}	(A.20)
e^{1K}_{ij}	Intralayer logical link connecting x^{1K}_i and x^{1K}_j in G^{1K}	(A.21)
E^2	Intralayer link set of physical layer G^2	(A.26)
e^2_{st}	Intralayer physical link connecting x^2_s and x^2_t in G^2	(A.27)
$E^{1K,2}$	Interlayer link set of G^{1K} and G^2	(A.29)
E^γ	Intralayer link set of Elemental-B network G^γ	(A.47)
e^γ_{uv}	Intralayer link that features in SD_{ij}	(A.47)
E^ζ	Intralayer link set of Overall-B network G^ζ	(A.41)
e^ζ_{uv}	Intralayer link that features in $\mathcal{C}(\mathcal{S}_{ij})$	(A.41)
Elemental-B(e^2_{st})	Link betweenness score of e^2_{st} based only on SD_{ij}	(A.50)
$\mathcal{G}$	Family of individual network layers	(A.2)(A.16)
G_α	Individual network layers of the generic formulation	(A.7)
G_β		(A.11)
G_b		(A.4)
G^{1K}	Logical network layer	(A.17)
G^2	Physical network layer	(A.22)
G^γ	Elemental-B network layer	(A.44)
G^ζ	Overall-B network layer	(A.38)
L_{ij}	Shortest path length between node-pair (x^{1K}_i, x^{1K}_j)	(A.37)
$\bar{L}$	Average shortest path across all node-pairs in G^{1K}	(A.36)
$L_{ij}(t)$	Shortest path length between node-pair (x^{1K}_i, x^{1K}_j) at a specific time t during a simulation	(A.68)
$L_{ij}(t+z)$	Shortest path length between node-pair (x^{1K}_i, x^{1K}_j) at a specific time $t+z$ during a simulation	
$\Delta L_{ij}(t;t+z)$	Difference in the shortest path length of a node-pair (x^{1K}_i, x^{1K}_j) between time t and $t+z$	

Table A.4: Glossary of mathematical symbols and elements with reference to defining equations (continued).

Symbol	Description	Eq. ref
$\max(\boldsymbol{B}_{\text{overall}})$	Maximum Overall-B value	(A.66)
$\max(\boldsymbol{B}_{\text{elemental}})$	Maximum Elemental-B value	(A.67)
$\mathcal{M}$	Multilayered network	(A.1)
M	Number of individual network layers	(A.3)
mid	Midpoint of the ordered set of set sizes ($\boldsymbol{P}$)	(A.55)(A.57)
N_α	Number of nodes in G_α	(A.8)
N_β	Number of nodes in G_β	(A.12)
N^{1K}	Number of nodes in G^{1K}	(A.19)
N^2	Number of nodes in G^2	(A.24)
N^γ	Number of nodes in G^γ	(A.46)
N^ζ	Number of nodes in G^ζ	(A.40)
Overall-B(e^2_{st})	Link betweenness score of e^2_{st} based on $\mathcal{C}(\mathcal{S}_{ij})$	(A.43)
Overall-S(e^2_{st})	Link salience score of e^2_{st} based on $\mathcal{C}(\mathcal{S}_{ij})$	(A.53)
P_{ij}	Number of shortest paths in $\mathcal{S}_{ij}$	(A.35)
$\boldsymbol{P}$	Ordered set of shortest path set sizes P_{ij}	(A.58)
$\tilde{P}$	Median of $\boldsymbol{P}$ (generally)	
$\tilde{P}$(All)	Median of $\boldsymbol{P}$ (considering $\mathcal{C}(\mathcal{S}_{ij})$)	(A.54)(A.56)
$\tilde{P}$(Dir)	Median of $\boldsymbol{P}$ (considering SD_{ij})	
$\boldsymbol{P}^{25\%}$	Set of P_{ij} that fall within the 25$^{\text{th}}$ percentile of $\tilde{P}$	(A.58)
$\tilde{P}^{25\%}$	Mean of $\boldsymbol{P}^{25\%}$ (generally)	
$\tilde{P}^{25\%}$(All)	Mean of $\boldsymbol{P}^{25\%}$ (considering $\mathcal{C}(\mathcal{S}_{ij})$)	(A.59)
$\tilde{P}^{25\%}$(Dir)	Mean of $\boldsymbol{P}^{25\%}$ (considering SD_{ij})	
P_n	Elements of the ordered set of $\boldsymbol{P}^{25\%}$	(A.59)
P_{mid}	Left value of the median of even set or midpoint value of an uneven set $\tilde{P}$	(A.54)(A.56)
$P_{\text{mid+1}}$	Right value of the median of even set $\tilde{P}$	(A.54)
$Rel\Delta L_{ij}(t;t+z)$	Relative step-change in $L_{ij}(t;t+z)$	(A.69)
$Rel\Delta \bar{L}(t;t+z)$	Average of $Rel\Delta L_{ij}(t;t+z)$ over all node-pairs in G^{1K}	(A.70)
$R(B_{\text{elemental}})^{75\%}$	Range between $B^{75\%}_{elemental}$ and $\max(\boldsymbol{B}_{elemental})$	(A.67)
$R(B_{\text{overall}})^{75\%}$	Range between $B^{75\%}_{overall}$ and $\max(\boldsymbol{B}_{overall})$	(A.66)
SD_{ij}	Shortest path set of a directly connected node-pair in G^{1K}	(A.31)
SI_{ij}	Shortest path set of an indirectly connected node-pair in G^{1K}	(A.32)
$\mathcal{S}_{ij}$	Shortest path set of any node-pair in G^{1K}	(A.33)

Table A.5: Glossary of mathematical symbols and elements with reference to defining equations (continued).

Symbol	Description	Eq. ref
w_{uv}^{γ}	Weight of link e_{uv}^{γ} which is the number of times it features in SD_{ij}	(A.49)
w_{uv}^{ζ}	Weight of link e_{uv}^{γ} which is the number of times it features in $\mathcal{C}(\mathcal{S}_{ij})$	(A.42)
X_{α}	Node set of G_{α}	(A.8)
x_b^{α}	Node in node set X_{α}	(A.8)
X_{β}	Node set of G_{β}	(A.12)
x_b^{β}	Node in node set X_{β}	(A.12)
X_b	Node set of G_b	(A.4)
X^{1K}	Node set of logical network G^{1K}	(A.18)
x_i^{1K}	Node in node set X^{1K}	(A.18)
X^2	Node set of physical network G^2	(A.23)
x_i^2	Node in node set X^2	(A.23)
X^{γ}	Node set of G^{gamma}	(A.45)
x_i^{γ}	Node in node set X^{γ}	(A.45)
X^{ζ}	Node set of G^{zeta}	(A.39)
x_i^{ζ}	Node in node set X^{ζ}	(A.39)

A.2 Mathematical formulations

A.2.1 Generic multilayered network

(Refer to Section 3.2.1.)

$$\mathcal{M} = (\mathcal{G}, \mathcal{C}) \tag{A.1}$$

where

$$\mathcal{G} = \{G_b; b \in \{1, 2, \ldots, M\}\} \tag{A.2}$$
$$M \equiv \text{ number of individual graph layers} \tag{A.3}$$
$$G_b = (X_b, E_b) \tag{A.4}$$

and

$$\mathcal{C} = \{E_{b,c} \subseteq X_b \times X_c; \forall b, c \in \{1, 2, \ldots, M\}, b \neq c\} \tag{A.5}$$

Now let

$$\alpha, \beta \text{ refer to layers of } \mathcal{G} | \alpha, \beta \in \{1, 2, \ldots, M\} \text{ and } \alpha \neq \beta \tag{A.6}$$

then

$$G_{\alpha} = (X_{\alpha}, E_{\alpha}) \tag{A.7}$$

with nodes

$$X_\alpha = \left\{x_1^\alpha, x_2^\alpha, \ldots, x_{N_\alpha}^\alpha\right\} \text{ where } N_\alpha \text{ was the number of nodes} \tag{A.8}$$

and *intralayer* links

$$E_\alpha \subseteq X_\alpha \times X_\alpha \tag{A.9}$$
$$\tag{A.10}$$

Then similarly,

$$G_\beta = (X_\beta, E_\beta) \tag{A.11}$$
$$X_\beta = \left\{x_1^\beta, x_2^\beta, \ldots, x_{N_\beta}^\beta\right\} \text{ where } N_\beta \text{ was the number of nodes} \tag{A.12}$$
$$E_\beta \subseteq X_\beta \times X_\beta \tag{A.13}$$

The *interlayer* connections were

$$E_{\alpha,\beta} \subseteq X_\alpha \times X_\beta; \forall \alpha, \beta \in \{1, 2, \ldots, M\}, \alpha \neq \beta\} \tag{A.14}$$

A.2.2 Customised multilayered network formulation

(Refer to Section 3.2.2.)

For this thesis we adapted the generic formulation. One universal adaptation was that the indices that named the different network layers (α and β in the generic formulation) were superscripts in the customised formulation instead of subscripts. This was necessary to avoid confusion with node indices which were (as per convention) indicated as subscripts.

$$\mathcal{M} = (\mathcal{G}, \mathcal{C}) \tag{A.15}$$
$$\mathcal{G} = (G^{1K}, G^2) \qquad \text{where } K \in \{F, S, D\} \tag{A.16}$$

G^{1K} was the logical layer and G^2 was the physical layer of $\mathcal{M}$ and F: FC archetype; S: SH archetype and D: DH archetype.

Logical layer[1]:

$$G^{1K} = (X^{1K}, E^{1K}) \qquad \forall K \in \{F, S, D\} \tag{A.17}$$

with nodes

$$X^{1K} = \{x_1^{1K}, x_2^{1K}, \ldots, x_{N^{1K}}^{1K}\} \qquad \forall K \in \{F, S, D\} \tag{A.18}$$

where

$$N^{1K} = 12 \qquad \forall K \in \{F, S, D\} \tag{A.19}$$

and links

$$E^{1K} = \{e_{ij}^{1K}\}\ \forall i,j \in \{1,2,\ldots,N^{1K}\} | i \neq j, \qquad \forall K \in \{F,S,D\} \tag{A.20}$$

where

$$e_{ij}^{1K} = \begin{cases} 1 & \text{if } x_i^{1K} \text{ was connected to } x_j^{1K} \\ & \forall K \in \{F,S,D\} \\ 0 & \text{otherwise.} \end{cases} \tag{A.21}$$

Physical layer:

$$G^2 = (X^2, E^2) \tag{A.22}$$

with nodes

$$X^2 = \{x_1^2, x_2^2, \ldots, x_{N^2}^2\} \text{ where,} \tag{A.23}$$
$$N^2 = m \times n = 100 \tag{A.24}$$
$$\tag{A.25}$$

and links

$$E^2 = \{e_{st}^2\} \qquad \forall s,t \in \{1,2,\ldots,N^2\} | s \neq t \tag{A.26}$$
$$e_{st}^2 = \begin{cases} 1 & \text{if } x_s^2 \text{ was connected to } x_t^2 \\ 0 & \text{otherwise.} \end{cases} \tag{A.27}$$

The interconnections of $\mathcal{M}$ were defined by:

$$\mathcal{C} = E^{1K,2} \tag{A.28}$$

where $E^{1K,2}$ was defined by the adjacency matrix

$$A^{[1K,2]} = (a_{is}^{1K,2}) \tag{A.29}$$

and

$$a_{is}^{1K,2} = \begin{cases} 1, & \text{if } (x_i^{1K}, x_s^2) \in E^{1K,2} \\ & \forall i \in \{1,2,\ldots,N^{1K}\}, \text{ and} \\ & \forall s \in \{1,2,\ldots,N^2\} \\ 0, & \text{otherwise} \end{cases} \tag{A.30}$$

A.2.3 Collection of shortest path sets

(Refer to Section 3.4.)

All metrics pertaining to shortest paths referred to a specific realisation of $\mathcal{M}$, therefore in the definitions that follow the superscripts relating to the layers and network archetypes were dropped for simplicity's sake.

A shortest path set was a collection of shortest paths of equal length connecting two

nodes x_i^{1K} and x_j^{1K}.

Shortest path sets between **directly connected node-pairs:**

$$SD_{ij} = \{s^1, s^2, \ldots, s^{P_{ij}}\} \qquad \forall e_{ij}^{1K} \in E^{1K} \tag{A.31}$$

Shortest path sets between **indirectly connected node-pairs:**

$$SI_{ij} = \{s^1, s^2, \ldots, s^{P_{ij}}\} \qquad \forall e_{ij}^{1K} \notin E^{1K} \tag{A.32}$$

thus

$$\mathcal{S}_{ij} = \{SD_{ij}, SI_{ij}\} \qquad \forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j \tag{A.33}$$

Collection of shortest path sets:

$$\mathcal{C}(\mathcal{S}_{ij}) = \bigcup_{i,j} \mathcal{S}_{ij} \qquad \forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j \tag{A.34}$$

where

$$P_{ij} \equiv \text{number of alternative shortest paths between node } x_i^{1K} \text{ and } x_j^{1K} \qquad \forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j \tag{A.35}$$

The efficiency of $\mathcal{M}$ was the average shortest path and was defined by:

$$\bar{L} = \frac{\Sigma_{i,j,i\neq j} L_{ij}}{N^{1K}(N^{1K} - 1)} \text{ where } i, j \in \{1, 2, \ldots, N^{1K}\} \tag{A.36}$$

where

$$L_{ij} \equiv \text{length of a shortest path between node } x_i^{1K} \text{ and } x_j^{1K} \qquad \forall i, j \in \{1, 2, \ldots, N^{1K}\}, i \neq j \tag{A.37}$$

A.2.4 Targeted attack simulations

Overall link betweenness (Overall-B)

(Refer to Section 4.1.3.)
The partial grid network comprising all links that feature in $\mathcal{C}(\mathcal{S}_{ij})$ was defined by:

$$G^{\zeta} = (X^{\zeta}, E^{\zeta}) \tag{A.38}$$

with nodes

$$X^{\zeta} \subseteq X_2 \mid x_u^{\zeta} \in \mathcal{C}(\mathcal{S}_{ij}), \qquad u \in \{1, 2, \ldots, N^{\zeta}\} \tag{A.39}$$

where

$$N^{\zeta} \equiv \text{number of unique nodes in } X^{\zeta} \tag{A.40}$$

and links

$$E^{\zeta} \subseteq E_2 \mid e_{uv}^{\zeta} \in \mathcal{C}(\mathcal{S}_{ij}) \qquad u, v \in \{1, 2, \ldots, N^{\zeta}\} \tag{A.41}$$

and link weights

$$w_{uv}^{\zeta} \equiv \text{ number of occurrences of } e_{uv}^{\zeta} \text{ in } \mathcal{C}(\mathcal{S}_{ij}) \tag{A.42}$$

Overall-B was calculated from w_{uv}^{ζ} as follows:

$$\text{Overall-B}(e_{st}^2) = \begin{cases} \frac{w_{uv}^{\zeta}}{\sum\limits_{i,j;i\neq j} P_{ij}} & \text{if } e_{st}^2 \equiv e_{uv}^{\zeta} \text{ and thus } e_{st}^2 \in E^{\zeta}; \\ & \text{where } P_{ij} \text{ was the shortest path set size,} \\ & u, v \in \{1, 2, \ldots, N^{\zeta}\}, \\ & i, j \in \{1, 2, \ldots, N^{1K}\}, \text{ and} \\ & s, t \in \{1, 2, ..., N^2\} \\ 0 & \text{otherwise} \end{cases} \tag{A.43}$$

Elemental link betweenness (Elemental-B)

(Refer to Section 4.1.4.)
The partial grid network comprising all links that feature in SD_{ij} was defined by:

$$G^{\gamma} = (X^{\gamma}, E^{\gamma}) \tag{A.44}$$

with nodes

$$X^{\gamma} \subseteq X_2 \mid x_u^{\gamma} \in SD_{ij}, \qquad u \in \{1, 2, \ldots, N^{\gamma}\} \tag{A.45}$$

where

$$N^{\gamma} \equiv \text{number of unique nodes in } X^{\gamma} \tag{A.46}$$

and links

$$E^{\gamma} \subseteq E_2 \mid e_{uv}^{\gamma} \in SD_{ij} \qquad u, v \in \{1, 2, \ldots, N^{\gamma}\} \tag{A.47}$$

(A.48)

and link weights

$$w_{uv}^{\gamma} \equiv \text{ number of occurrences of } e_{uv}^{\gamma} \text{ in } SD_{ij} \tag{A.49}$$

Elemental-B was calculated from w_{uv}^{γ} as follows:

$$\text{Elemental-B}(e^2_{st}) = \begin{cases} \frac{w^{\gamma}_{uv}}{\sum_{i,j;i\neq j} P_{ij}} & \text{if } e^2_{st} \equiv e^{\gamma}_{uv} \text{ and thus } e^2_{st} \in E^{\gamma}; \\ & u, v \in \{1, 2, \ldots, N^{\gamma}\}, \\ & i, j \in \{1, 2, \ldots, N^{1K}\}, \text{ and} \\ & s, t \in \{1, 2, ..., N^2\} \\ 0 & \text{otherwise} \end{cases} \tag{A.50}$$

Link salience (Overall-S)

(Refer to Section 4.1.5.)
Consensus scores were calculated by:

$$c_{st} = \sum_{i,j;i\neq j} c_{st}(i,j) | s, t \in \{1, 2, ...N^2\}, i, j \in \{1, 2, ..., N^{1K}\} \tag{A.51}$$

where

$$c_{st}(i,j) = \begin{cases} 1 & \text{if } e^2_{st} \in \mathcal{S}_{ij} \\ 0 & \text{otherwise} \end{cases} \tag{A.52}$$

In a network where all node-pairs were still connected, the number of shortest path sets $\|\mathcal{C}(\mathcal{S}_{ij})\|$ was equal to the number of node-pairs. However, we continued the targeted attack simulation beyond disconnection, therefore we divided the consensus score by the number of shortest path sets and not the number of node-pairs such that:

$$\text{Overall-S}(e^2_{st}) = \frac{c_{st}}{\|\mathcal{C}(\mathcal{S}_{ij})\|} \text{ where } i, j \in \{1, 2, ..., N^{1K}\} \tag{A.53}$$

A.2.5 Vulnerability metrics

Redundancy

(Refer to Section 5.3.)

The ordered set of P_{ij} was defined by $\boldsymbol{P}$. In the case of an even number of sets the median was calculated by:

$$\tilde{P} = (P_{\text{mid}} + P_{\text{mid}+1})/2 \tag{A.54}$$

where $P_{\text{mid}} \in \boldsymbol{P}$, and

$$\text{mid} = (N^{1K}(N^{1K} - 1) + 1)/2 - 0.5 \tag{A.55}$$

with N^{1K} the number of nodes in G^{1K}. In the case of uneven case the median was calculated by:

$$\tilde{P} = P_{\text{mid}} \tag{A.56}$$

where $P_{\text{mid}} \in \boldsymbol{P}$, and

$$\text{mid} = (N^{1K}(N^{1K} - 1) + 1)/2 \tag{A.57}$$

The set of values that fell within the 25$^{\text{th}}$ percentile was denoted by:

$$\boldsymbol{P}^{25\%} \subset \boldsymbol{P} \text{ such that } \boldsymbol{P}^{25\%} = \left\{P_1, P_2, \ldots, P_{\lceil \|\boldsymbol{P}\|/4 \rceil}\right\} \tag{A.58}$$

The mean of the elements of $\boldsymbol{P}^{25\%}$ was defined as:

$$\tilde{P}^{25\%} = \frac{\sum\limits_{P_n \in \boldsymbol{P}^{25\%}} P_n}{\|\boldsymbol{P}^{25^{\text{th}}}\|} \tag{A.59}$$

The vulnerability metrics related to redundancy are summarised in Table A.6.

Table A.6: Summary of redundancy metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\tilde{P}$(All)	(A.54)(A.56) $(P_{ij} \in \mathcal{C}(\mathcal{S}_{ij}))$
	SD_{ij}	$\tilde{P}$(Dir)	(A.54)(A.56)$(P_{ij} \in SD_{ij})$
Left-tail centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\tilde{P}^{25\%}$(All)	(A.59) $(P_{ij} \in \mathcal{C}(\mathcal{S}_{ij}))$
	SD_{ij}	$\tilde{P}^{25\%}$(Dir)	(A.59) $(P_{ij} \in SD_{ij})$

Overlap

(Refer to Section 5.4.)

$$\boldsymbol{B}_{overall} \equiv \text{set of Overall-B}(e_{st}^2) \text{ in descending order} \tag{A.60}$$
$$\boldsymbol{B}_{elemental} \equiv \text{set of Elemental-B}(e_{st}^2) \text{ in descending order} \tag{A.61}$$

$$\overline{B}_{\text{overall}} = \frac{\sum\limits_{B_n \in \boldsymbol{B}_{overall}} B_n}{\|\boldsymbol{B}_{overall}\|} \tag{A.62}$$

and

$$\overline{B}_{\text{elemental}} = \frac{\sum\limits_{B_n \in \boldsymbol{B}_{elemental}} B_n}{\|\boldsymbol{B}_{elemental}\|} \tag{A.63}$$

The value of the 75$^{\text{th}}$ percentile was denoted by:

$$B_{overall}^{75\%} = B_{\lfloor \|\boldsymbol{B}_{overall}\|/4 \rfloor} \tag{A.64}$$

and

$$B^{75\%}_{elemental} = B_{\lfloor \|\boldsymbol{B}_{elemental}\|/4 \rfloor} \tag{A.65}$$

The range was then defined by:

$$R(B_{\text{overall}})^{75\%} = \max(\boldsymbol{B}_{overall}) - B^{75\%}_{overall} \tag{A.66}$$

and

$$R(B_{\text{elemental}})^{75\%} = \max(\boldsymbol{B}_{elemental}) - B^{75\%}_{elemental} \tag{A.67}$$

The vulnerability metrics related to overlap are summarised in Table A.7.

Table A.7: Summary of overlap metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$\overline{B}_{\text{overall}}$	(A.62)
	SD_{ij}	$\overline{B}_{\text{elemental}}$	(A.63)
Right-tail range	$\mathcal{C}(\mathcal{S}_{ij})$	$R(B_{\text{overall}})^{75\%}$	(A.66)
	SD_{ij}	$R(B_{\text{elemental}})^{75\%}$	(A.67)

Efficiency step-change

(Refer to Section 5.5.)
The step-change in the shortest path length of a node-pair (x_i^{1K}, x_j^{1K}) was:

$$\Delta L_{ij}(t; t+z) = L_{ij}(t) - L_{ij}(t+z) \tag{A.68}$$

where t was some defined point in time before a disruption occurred and z was some time after a disruption of G^2 occurred. The relative change was then:

$$Rel\Delta L_{ij}(t; t+z) = \frac{\Delta L_{ij}(t; t+z)}{L_{ij}(t)} \tag{A.69}$$

The aggregate of the relative step-changes across all node-pairs were:

$$Rel\overline{\Delta L}(t; t+z) = \frac{\sum\limits_{i,j;i\neq j} Rel\Delta L_{ij}(t; t+z)}{N^{1K}(N^{1K}-1)} \text{ where } i,j \in \{1,2,\dots N^{1K}\} \tag{A.70}$$

The vulnerability metrics related to efficiency step-change are summarised in Table A.8.

Table A.8: Summary of efficiency step-change metrics.

Aspect	**Scope**	**Metric**	**Equation**
Centrality	$\mathcal{C}(\mathcal{S}_{ij})$	$Rel\overline{\Delta L}(t;t+z)$	(A.70)

Appendix B

KS-test results

The tables in this appendix arrange the p-values for the Kolmogorov-Smirnov tests conducted in Section 6.1.2. The KS-test was used to test the following hypothesis:

- H_0: The sample distributions of $v_i(X)$ and $v_i(Y)$ are drawn from the same theoretical distribution.
- H_A: The sample distributions of $v_i(X)$ and $v_i(Y)$ are not drawn from the same theoretical distribution.

(Where v_i is a vulnerability metric from Table 6.4.)

The test statistic (D) quantifies the distance between the Empirical Distribution Function (EDF) of the vulnerability metric as measured from the sample of surviving instances X and the EDF of the vulnerability metric as measured from the sample of non-surviving instances Y. The null hypothesis that the sample distributions were drawn from the same theoretical distribution was rejected when the p-value was lower than the chosen significance level, which in this case was 0.05.

The KS-test was conducted for each vulnerability metric after each disruption, provided that the samples X and Y had more than 15 observations each. In the case of *efficiency loss*, the sample size of X for the FC was smaller than 15 after each disruption. Therefore, no tests were performed in this regard. The p-values for all the other tests are tabulated hereunder.

Table B.1: p-values from the KS-tests comparing the EDFs of the vulnerability metrics and robustness for the FC network. If $p < 0.05$ (highlighted in green) then the null hypothesis that the EDFs are from the same distribution was rejected. Blank cells indicate sample sizes ≤ 15.

			Robustness									
Percentage grid links remaining			95%	90%	85%	80%	75%	70%	65%	60%	55%	50%
Redundancy	$\tilde{P}$(All)	% Change	–	–	0.90	0.82	0.24	0.16	0.57	–	–	–
Overlap	$\overline{B}_{\text{overall}}$	% Change	–	–	0.89	0.23	0.86	0.06	0.58	–	–	–
		Final value	–	–	0.29	0.01	0.14	0.04	0.1	–	–	–
Efficiency Step-Change	$Rel\overline{\Delta L}(t; t+k)$	Average	–	–	0.01	0.02	<0.01	0.37	0.39	–	–	–
		Maximum	–	–	0.03	0.05	0.01	0.58	0.54	–	–	–

Table B.2: p-values from the KS-tests comparing the EDFs of the vulnerability metrics and efficiency loss for the SH network. If $p < 0.05$ (highlighted in green) then the null hypothesis that the EDFs are from the same distribution was rejected. Blank cells indicate sample sizes ≤ 15.

			Efficiency Loss									
Percentage grid links remaining			95%	90%	85%	80%	75%	70%	65%	60%	55%	50%
Redundancy	$\tilde{P}$(Dir)	Final value	0.78	0.35	0.37	0.77	0.54	–	–	–	–	–
Overlap	$\overline{B}_{\text{overall}}$	% Change	0.01	0.06	0.43	0.31	0.05	–	–	–	–	–
		Final value	0.69	0.18	0.14	0.50	0.12	–	–	–	–	–
	$\overline{B}_{\text{elemental}}$	% Change	0.04	0.17	0.20	0.31	<0.01	–	–	–	–	–
		Final value	0.68	0.47	0.06	0.53	<0.01	–	–	–	–	–
Efficiency Step-Change	$Rel\overline{\Delta L}(t; t+k)$	Average	–	0.93	0.69	0.34	0.48	–	–	–	–	–
		Maximum	–	0.94	0.15	0.30	0.45	–	–	–	–	–

Table B.3: p-values from the KS-tests comparing the EDFs of the vulnerability metrics and robustness for the SH network. If $p < 0.05$ (highlighted in green) then the null hypothesis that the EDFs are from the same distribution was rejected. Blank cells indicate sample sizes ≤ 15.

			Robustness									
Percentage grid links remaining			95%	90%	85%	80%	75%	70%	65%	60%	55%	50%
Redundancy	$\tilde{P}$(All)	Final value	–	–	0.76	0.01	0.99	0.81	1.00	–	–	–
	$\tilde{P}$(Dir)	Final value	–	–	0.59	0.13	0.15	0.99	0.94	–	–	–
Overlap	$\overline{B}_{\text{overall}}$	% Change	–	–	0.11	0.30	0.15	<0.01	0.36	–	–	–
		Final value	–	–	0.78	0.67	0.82	0.02	0.20	–	–	–
	$\overline{B}_{\text{elemental}}$	% Change	–	–	0.26	0.68	0.41	<0.01	0.15	–	–	–
		Final value	–	–	0.43	0.76	0.67	0.06	0.13	–	–	–
Efficiency Step-Change	$Rel\overline{\Delta L}(t; t+k)$	Maximum	–	–	0.57	0.73	0.97	<0.01	0.45	–	–	–

Table B.4: p-values from the KS-tests comparing the EDFs of the vulnerability metrics and efficiency loss for the DH network. If $p < 0.05$ (highlighted in green) then the null hypothesis that the EDFs are from the same distribution was rejected. Blank cells indicate sample sizes ≤ 15.

			Efficiency Loss									
Percentage grid links remaining			95%	90%	85%	80%	75%	70%	65%	60%	55%	50%
Redundancy	$\tilde{P}$(All)	% Change	0.02	0.68	0.27	0.99	0.15	0.80	–	–	–	–
	$\tilde{P}$(Dir)	Final value	0.07	0.72	0.80	1.00	0.97	0.43	–	–	–	–
Overlap	$\overline{B}_{\text{overall}}$	% Change	0.93	0.90	0.21	0.43	0.59	0.72	–	–	–	–
		Final value	0.23	0.93	0.39	0.22	0.38	0.99	–	–	–	–
	$\overline{B}_{\text{elemental}}$	% Change	0.62	0.83	0.49	0.41	0.16	0.78	–	–	–	–
		Final value	0.14	0.19	0.14	0.69	0.50	0.40	–	–	–	–
Efficiency Step-Change	$Rel\overline{\Delta L}(t; t+k)$	Average	–	0.45	0.13	0.39	0.51	0.61	–	–	–	–
		Maximum	–	0.29	0.02	0.14	0.65	0.18	–	–	–	–

Table B.5: p-values from the KS-tests comparing the EDFs of the vulnerability metrics and robustness for the DH network. If $p < 0.05$ (highlighted in green) then the null hypothesis that the EDFs are from the same distribution was rejected. Blank cells indicate sample sizes ≤ 15.

			Robustness									
Percentage grid links remaining			95%	90%	85%	80%	75%	70%	65%	60%	55%	50%
Redundancy	$\tilde{P}$(All)	Final value	–	0.82	<0.01	0.80	0.14	0.70	1.00	–	–	–
	$\tilde{P}$(Dir)	Final value	–	1.00	0.17	0.89	0.93	0.99	0.76	–	–	–
Overlap	$\overline{B}_{\text{overall}}$	% Change	0.65	0.60	0.23	0.41	0.05	0.51	–	–	–	
		Final value	0.85	0.78	0.39	0.60	0.04	0.39	–	–	–	
	$\overline{B}_{\text{elemental}}$	% Change	–	0.41	0.82	0.14	0.59	0.06	0.45	–	–	–
		Final value	–	0.90	0.97	1.00	0.31	0.02	0.81	–	–	–
Efficiency Step-Change	$Rel\overline{\Delta L}(t; t+k)$	Maximum	–	0.89	0.21	0.20	0.25	<0.01	0.22	–	–	–