Purdue University
Purdue e-Pubs

Open Access Theses

Theses and Dissertations

January 2016

Graph Theoretical Analysis of the Dynamic Lines of Collaboration Model for Disruption Response

Arfinandi Ferialdy
Purdue University

Graduate School Form
30 Updated 12/26/2015

PURDUE UNIVERSITY
GRADUATE SCHOOL
Thesis/Dissertation Acceptance

This is to certify that the thesis/dissertation prepared

By ARFINANDI FERIALDY

Entitled

GRAPH THEORETICAL ANALYSIS OF THE DYNAMIC LINES OF COLLABORATION MODEL FOR DISRUPTION RESPONSE

For the degree of Master of Science in Industrial Engineering

Is approved by the final examining committee:

Shimon Y. Nof
Chair

Vaneet Aggarwal

Seokcheon Lee

To the best of my knowledge and as understood by the student in the Thesis/Dissertation Agreement, Publication Delay, and Certification Disclaimer (Graduate School Form 32), this thesis/dissertation adheres to the provisions of Purdue University's "Policy of Integrity in Research" and the use of copyright material.

Approved by Major Professor(s): Shimon Y. Nof

Approved by: Abhijit Deshmukh 10/14/2016

Head of the Departmental Graduate Program Date

GRAPH THEORETICAL ANALYSIS OF THE DYNAMIC LINES OF COLLABORATION MODEL FOR DISRUPTION RESPONSE

A Thesis

Submitted to the Faculty

of

Purdue University

by

Arfinandi Ferialdy

In Partial Fulfillment of the

Requirements for the Degree

of

Master of Science in Industrial Engineering

December 2016

Purdue University

West Lafayette, Indiana

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

N2N	Network-to-Network
CPI	Cyber-Physical Infrastructure
DLOC	Dynamic Lines of Collaboration
CCT	Collaborative Control Theory
TIE	Teamwork Integration Evaluator
LOC	Line of Collaboration
CDR	Collaborative Disruption Response
ER	Erdos-Renyi Random Network
BA	Barabasi-Albert Scale-Free Networks
WS	Watts-Strogatz Small-World model
PW	USA Western States Power Grid Network
ACRP	Asynchronous Collaboration Requirement Planning
CBA	Centrality-based depot allocation
DBA	Degree-based depot allocation
BBA	Bridge-based depot allocation
ABP	activity-based priority
AL	Auxiliary Links

FCFS	First-Come-First-Serve
CE	Conflicts and human/machine errors
CEPD	Conflicts and Errors Prognostics and Diagnostics

LIST OF SYMBOLS

Symbol		Page
G	Client Network	1
N	Client Nodes	1
E	Edges of Client Network	1
S	Server Network	1
A	Service Agents Nodes	1
P	Edges of Server Network (Collaboration Capability)	1
A_G	Adjacency matrix of client network G	9
P_G	Edge distribution (degree distribution) of client network G	9
M_G	Incidence matrix of client network G	9
$state_t(i,j)$	state of edge (i,j) at time t	9
$state_t(i)$	state of node (i) at time t	9
F	Set of total failures	9
τ_i	Initial time stamp of disruption i	10
v_i	location of disruption i	10
γ_i	failure reference of disruption i	10
W	Set of Colors	11

Symbol		Page
D	Depots of disruption responder	11
φ	Cascade Threshold	11
c_{ij}	Weighted edges of virtual graph	12
$t_D(i,j)$	Travel time of agents between node i and j	12
$t_R(j)$	Time required to remove disruption at node j	12
$d(i,j)$	Distance function between node i and j	12
$deg(v)$	degree of node v of the client network	13
z	DLOC Total latency metric	15
σ_i	Time stamp when disruption i was repaired	15
F_0	Set of initiating failure	15
x_{ij}	Binary decision variable to determine an agent will to from node i to j	15
X	The set of all x_{ij}	15
$\overline{deg_G}$	Average degree of client network G	31
p_{F_0}	Failure Initiator Fraction	31
$P_{recover}$	DLOC Recoverability Metrics	31
$\lvert F_t \rvert$	Dynamic size of cascading failure at time t	32
S_t	Subset of client nodes in cascading cluster	32
$deg_s(A)$	degree of each node in server network	33

Symbol		Page
$P_{vulnerable}$	fraction of nodes in client network vulnerable to cascade	34
p_k	The probability that a node has degree k	34
$p_{v(k)}$	Binary variable to denote the vulnerability of a node with degree k	34
B_i	Betweenness centrality value of node i	42
ρ_{st}	The number shortest paths from node s to t	42
$\rho_{st}(i)$	the number of shortest paths from node s to t that passes through i	42
Dg_i	Degree Centrality Value of node i	42
Br_i	Bridge Centrality Value of node i	43
$C_{Br(i)}$	Bridging Coefficient of node i	43
R_i	set of neighbors of node i	43
$\lvert F_{(\alpha)} \rvert$	The total number of failures triggered by α centrality	46
S	Fraction of Failed elements to total	48
Δ	Total distance traveled by responder agent in DLOC	55
$\tilde{F}$	Set of failed elements without disruption response mechanism	55
P_{rev}	DLOC Preventability Metrics	55
t_{repair}	Disruption repair timespan	58
u	Failure Propagation speed	59
v	Agent’s travelling speed	59

Symbol		Page
n_R	Number of replication for simulation scenario	59
t_{max}	Simulation length for each scenario	59

ABSTRACT

Ferialdy, Arfinandi. M.S.I.E., Purdue University, August 2016. Graph Theoretical Analysis of The Dynamic Lines of Collaboration Model for Disruption Response. Major Professor: Shimon Y. Nof.

The Dynamic Lines of Collaboration (DLOC) model was developed to address the Network-to-Network (N2N) service challenge found in e-Work networks with pervasive connectivity. A variant of the N2N service challenge found in emerging Cyber-Physical Infrastructures (CPI) networks is the collaborative disruption response (CDR) operation under cascading failures. The DLOC model has been validated as an appropriate modelling tool to aid the design of disruption responders in CPIs by eliciting the dynamic relation among the service team when handling service requests from clients in the CPI network. The DLOC model for CDR operation is conceptually an abstraction of the CPI network into two interdependent networks of client and service networks. No preliminary design guidelines have been devised for DLOC-CDR from a network perspective using graph properties. Previous results of graph theoretical analysis for network behaviors under disruption may also not apply to DLOC-CDR due to the intrinsic nature of the N2N service challenge. Previous research works in DLOC-CDR have also not taken into consideration in protecting vulnerable CPI network elements which can cause system collapse by a single failure. Based on these observations, this research is guided by the following

questions: (1) What graph properties to be viable predictors for evaluating the reliability of N2N service designs in DLOC and (2) Where should the disruption responders (resource) to ensure timely disruption mitigation with regards to protection of vulnerable nodes?

To answer question (1), it is found that resiliency of a CPI network, as measured in DLOC by the Recoverability metric ($P_{recover}$), can be approximated by the proportion of vulnerable nodes ($P_{vulnerable}$) as a function of average degree and cascade threshold (φ). $P_{recover}$ measures the probability of a network to fully recover from a cascading disruption. It is found that there lies a certain regime where $P_{recover}$ =1 as approximated by $P_{vulnerable}$. By means of graph property analysis, we initially proposed two heuristics, $\overline{deg_G} > \frac{1}{\varphi}$ and $\sum_k p_{v(k)} p_k < 0.70$, to mark the regime where $P_{recover}$ is strictly 1. From numerical experiments, it was found that $\overline{deg_G} > \frac{1}{\varphi}$ is over conservative, while the latter applies to all tested networks. This experiment result also supports the conclusion that the existence of a "small-world" phenomenon in networks can either inhibit or accelerate cascade, depending on the complexity of the propagation.

To answer question (2), two heuristics protocols based on network centrality measures were initially proposed, namely the Bridge-Based Allocation (BBA) and Degree-Based Allocation (DBA). We initially hypothesized that the BBA would perform better in terms of preventing failures but with considerable trade-off in total response time compared to CBA, the existing resource allocation protocol of DLOC-CDR, in networks with high modularity. However, it was found that based on numerical experiments we concede that the BBA is not suitable to be applied in DLOC. The main advantage of the BBA is its ability

to identify bridging elements which its removal will make the network disconnected. The current DLOC, on the other hand, does not take into consideration of the connectivity state of the CPI network. Thus, rendering the BBA to become less effective than CBA and DBA. We also found that both CBA and DBA can be used interchangeably. Given a simple propagation, DBA constantly performs better on networks displaying high affinity towards power-law degree distribution compared to CBA. This is due to the high correlation between both centralities in these networks. For small-world networks, the performance increment from CBA to DBA has a decreasing trend with increasing network size. CBA's performance is relatively constant and outperforms DBA in large network size.

CHAPTER 1. INTRODUCTION

1.1 Motivation: N2N Service Challenge – Disruption Response

Pervasive connectivity in e-Work networks (Nof, 2003) has brought forth a new type of service challenge in production systems. These challenges are characterized by increasingly interdependent service requirements and concurrent collaboration among service providers. A novel method to approach this challenge was proposed recently by (Zhong & Nof, 2015; Zhong, 2016) in the form of Dynamic Lines of Collaboration (DLOC) model. Inspired by the network-to-network interface ubiquitous in telecommunication networks, the DLOC model presents a novel abstraction of production systems as two interdependent networks, namely the client and the server network. The service challenge is then defined by the DLOC model as how to efficiently provide effective Network-to-Network (N2N) services in e-Work systems. Let the client network be defined as graph $G = (N, E)$ where the nodes N and edges E represent the elements of Cyber-Physical Infrastructure (CPI) and the interdependency between them, respectively. The server (service team) network is defined as graph $S = (A, P)$ where the nodes A and edges P denote service agents and collaboration compatibility between them, respectively. An abstract representation of the N2N services in DLOC model can be seen in Fig 1.1

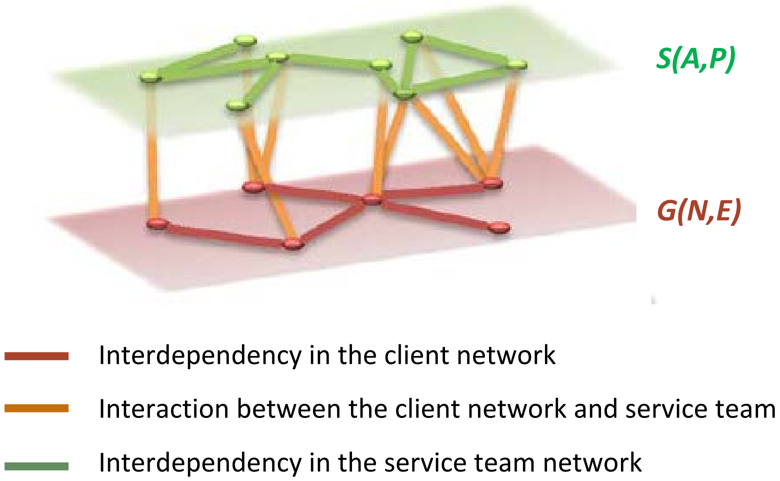

S(A,P)
G(N,E)
🟥 Interdependency in the client network
🟧 Interaction between the client network and service team
🟩 Interdependency in the service team network

Figure 1.1 Abstract representation of N2N services of e-Work in DLOC (Zhong, 2016)

Cascading failure is the mechanism by which failures propagate to cause large-scale catastrophes in complex systems. Cascading failure leads to the point of explaining large-scale blackout phenomenon which occurs in power-grid systems (Dobson et. al., 2007). Apart from power-grid systems, cascading failure behavior are also imminent in other man-made CPIs such as in Smart Water Distribution Networks (WDN) (Shuang et. al, 2015). Other researchers have also found that most man-made CPI network exhibits modularity (Newman, 2006) and have power-law degree distribution which can be closely modeled by scale-free networks (Barabasi & Albert 2002). The implication of this is that networks with power-law degree distribution are less resilient to targeted disruptions as compared to random networks (Motter & Lai, 2002).

The design of an effective disruption response team (service team) and operation in CPIs have to take into considerations of the interdependent service requirements between network elements. The DLOC model provides an appropriate modeling tool to aid the design of disruption responders by eliciting the dynamic relation and collaboration among service agents when handling service requests from clients in the CPI network. Control protocols have been developed and validated under the DLOC model to improve the performance of N2N services for Collaborative Disruption Response (DLOC-CDR) operations. This research will be focused on enhancing the DLOC model usability for designing and evaluating the performance of N2N services as well as investigating new protocols to improve DLOC-CDR operations.

1.2 Research Problem

Existing works on graph theoretical analysis of network behaviors under disruption have not yet analyzed the novel N2N formation of the DLOC model, especially for DLOC-CDR operations. Due to the intrinsic nature of N2N service challenge, the results of past analysis may not be applicable as a guideline for designing N2N service teams in e-Work. The existing resource allocation strategy for DLOC-CDR has also not taken into consideration of protecting vulnerable nodes in the client network, which can cause system collapse by a single node failure (Motter & Lai, 2002) or through global cascades (Watts, 2002; Singh et. al., 2013)

The research problem for this thesis can be defined as follows: To identify a set of graph theoretical properties for a preliminary guideline of N2N service designs and investigate new control protocols for resource allocation of service agents for protection of vulnerable nodes.

1.3 Research Questions

Based on the aforementioned research problem, the following research questions should be answered:

1. Research Question 1 (**RQ1**): **Graph Theoretical Analysis of DLOC Networks.** What graph properties that can be viable predictors for evaluating the reliability of N2N service designs? How good are these predictors?
2. Research Question 2 (**RQ2**): **Protection of Vulnerable Parts of the Network**. Which clients to connect with the service network to ensure timely disruption mitigation with regards to network topology and protection of vulnerable nodes?

1.4 Overview of Proposed Methods

This thesis proposes a development of one of the emerging principles of the Collaborative Control Theory (Nof, 2007) namely the DLOC Model (Zhong & Nof, 2015, Zhong, 2016) attributing to several network design methods and graph theoretical analysis. The DLOC model has been established to facilitate the modelling of various CPI systems and develop solutions for their respective N2N service challenges (Zhong & Nof, 2015, Zhong, 2016). Graph theoretical analysis will be conducted to find the relation among several basic network structural properties with the behavior and performance of N2N service for DLOC-CDR operations. The conclusion made from this analysis is set to establish preliminary guidelines for N2N service design in e-Work networks.

The Centrality-based depot allocation (CBA) protocol for service agent resource allocation on the previous version of DLOC only takes into consideration of the network's betweenness centrality. In this research, we will attempt to improve the performance of DLOC-CDR by approaching the development of new resource allocation protocol from a network vulnerability point of view on cascading failure. Tools and methods provided from graph theory will be used in guiding the development of hypothesis and verification & explanation of experiment results.

Finally, we will validate the proposed method of RQ1 and RQ2 using the Teamwork Integration Evaluator (TIE) – TIE/DLOC on several conceptual and real-world networks, namely electricity power grids.

1.5 Assumptions

The studies conducted in this research are built upon these following assumptions:

1. Propagation of services (failures)

   Client network of the DLOC model request services which are to be fulfilled by the server network. In terms of DLOC-CDR, the client networks are CPIs while the server networks are disruption responder agents. The client network requests services to the server networks whenever it experiences a failure. The propagation of failure or service request among the elements of client network will be modeled based on adapted Watts Threshold model (Watts, 2002).

2. Service team (Disruption responders)

The disruption responder agents travel within the client network according to the client's network topology. It is assumed that each edge in $G(N, E)$ represents one-unit length and the paths for agents to move are invulnerable to disruption, i.e. even though an edge or node is in failure state, an agent can still traverse through.

3. Collaboration of service provider

   Disruption responder agents are required to collaborate with each other to fulfill certain types of service requests, i.e. to recover an edge in client network requires the collaboration of two agents. The collaboration compatibility among responder agents is denoted by an edge between them in the server network.

4. Scheduling protocol

   The order in which a service is served within a queue follows the neuro-plasticity inspired protocols (Zhong, 2016). By this protocol, it is also assumed that the client network has an ability to add auxiliary edges to increase the robustness of local area to cascading failure.

CHAPTER 2. LITERATURE SURVEY

In this chapter, we will first review the DLOC model and its current development. Afterward, we will survey several literatures related to Graph theory application and analysis on complex networks. We will also be covering other topics related to complex networks are diffusion process, cascading failures, and disruption mitigation.

2.1 The Dynamic Lines of Collaboration Model (DLOC)

2.1.1 General Description

The Line of Collaboration (LOC) principles under Collaborative Control Theory (CCT) (Nof, 2007) addresses the relation among system elements (agents) to perform tasks in e-Work systems. Dynamic team structures are essential in e-collaboration, especially under emergent situations ((Velasquez, Yoon, & Nof, 2010). This statement is supported by the fact that different teams need to be formed to satisfy various task requirements and changing team structure is critical for the sustainability of the entire organization (Velasquez, Yoon, & Nof, 2010). Thus, the LOC are constantly updating inside and between teams.

The pervasive connectivity in emergent e-Work networks also administers collaboration requirement between interdependent networks, namely client and server networks (Zhong & Nof, 2015). For example, in a Cyber-Physical Systems (CPS), where elements are interdependent through cyber and physical links, service requirement from a client in a CPS will influence the other services in the same network, which dynamically affects the LOC within the server networks to provide their services. The concurrent collaboration of multiple servers is required to ensure that the service is provided promptly and prevents to become complicated by time.

The DLOC model was developed to address the emerging Network-to-Network (N2N) service challenge in e-Work systems. It captures all details of the dynamic interactions between the client and the server networks: A networked service team whose collaborative operations are dependent on the team structure and the requirements from the client network. The dynamic service request from the client network are dependent on existing services as well as the interdependency within the clients (Zhong & Nof, 2015).

2.1.2 Formulation of DLOC-CDR (Zhong & Nof, 2015)

The main formulation of DLOC consists of four major building blocks: client network, service network, service propagation and prevention of failure by service agents. The mathematical models for each building blocks are subject to the nature of the e-Work system under consideration. Therefore, the DLOC model is flexible to be implemented on various systems with subject to different control protocols and network structure.

2.1.2.1 Client Network

A client network is defined as $G = (N, E)$ with link distribution P_G, where N is set of nodes connected by edge (link) E and $|N|$ & $|E|$ represent the number of node and edges in the network, respectively. A edge (i, j) represent dependency the between two adjacent nodes $(i, j \epsilon N)$. The incidence matrix $\boldsymbol{M_G} = (m_{ne})$ maps out the incidence relationship between nodes $(n \epsilon N)$ and edges $(e \epsilon E)$ in client network G and adjacency matrix $\boldsymbol{A_G} = (a_{ij})$ contains the number of edges connecting nodes i and j, $(i, j \epsilon N)$. For simplicity we assume that the client network is a simple graph; no parallel edges or loops, such that $a_{ij} \leq 1, \forall_{ij}, i, j \epsilon N$. Each nodes or edge in G has two states at any time step: 0 or 1. In '0', the element is active – not requesting service, while '1' represent the element is disrupted – requesting service. Disruptions that occur at time *t* are categorized into two types: edge failure and node failure.

Client network edge failure $state_t(i, j) = 1$ and $state_{t\_}(i, j) = 0$ (2.1)

Client network node failure: $state_t(i) = 1$ and $state_{t\_}(i) = 0$ (2.2)

2.1.2.2 Failure and Service Propagation

Let *F* be the set of disruptions which occur in the client network. The mapping from node failures and link ruptures to *F* is as follow: a node failure is defined as a single disruption $(i \in F)$, and an edge failure is represented as two coupled disruptions $(i, j \in F)$. Each disruption is uniquely defined by a 3-tuple $(i = < \tau_i, \upsilon_i, \gamma_i >, i \in F)$, where τ_i is the initial timestamp of this disruption ($\tau_i \geq 0$); υ_i represents the location of the disruption

($v_i \in N$). Note: here i is *not* the index of nodes in *G* but an element of *F*; γ_i is a failure reference defined in Eq. (2.3).

Edge failure reference: $$\gamma_i \begin{cases} j, \ if\ state_{\tau_i}(v_i, v_j) = 1\ and\ state_{\tau_i^-}(v_i, v_j) = 0 \\ \emptyset, \ if\ state_{\tau_i}(v_i) = 1\ and\ state_{\tau_i^-}(v_i) = 0 \end{cases} \quad (2.3)$$

Eq. (2.3) shows that if two disruptions (i and j) are used to represent an edge failure together, the rupture references of them point to each other. If a disruption is used to represent a node failure, the reference is an empty pointer.

The service propagation is modeled using an adapted Watts Threshold Cascade model (Watts, 2002). Different from the original version, propagation of failure on edges is also incorporated in our adaptation. In the beginning, all network elements (node and edges) are in state “0”. When a node or edge experiences a disruption (failure), it switches to state “1”. The cascading failure is modeled as a sequence of state changes. If a node is failed at time t, all connected edges will also adopt state “1”. At each time step, a node will assess its connected edges to determine it state. If fraction of edges in state “1” is at least equal to the cascade threshold φ, the node adopts state “1”; thus, propagating the failure. Otherwise it retains its current state, vice versa ($0 \leq \varphi \leq 1$). Edges and nodes remain in state “1” unless repaired by responders.

2.1.2.3 Service Team

External to the client network (G), the service agents belong to another network defined as the service team. This network is a *k-coloring* of a graph $S = (A, P)$ where A is the set of service agents that can restore failed nodes and edges in the client network. P is set of weighted edges which connects them. The number of service agents in the team is denoted by $|A|$. There is a mapping of the node set of the graph, such that $c: A \rightarrow W$, where W is a set of *k colors* assignment to the nodes of S. Each colors of set W represents the different capability (skills, responsibility, and workflow) of each service agents, such no nodes of the same color are adjacent in the network (Bond & Murty, 2007). The node coloring of graph S is intended to model collaborative constraint of the service team to accomplish a certain task, such that for given task that requires collaboration, a given service agent (node) can only collaborate with another service agent which is adjacent to itself and has the required node coloring (in the case there many node colors which resembles different expertise). For the service team in this research, we will only apply *2-coloring* to the graph S.

Each agent has two states: 0 for idle and 1 for working. Initially, each agent has an inter-edge (E) with one node in G (i.e., the depot). If a node is failed, an agent will be assigned to repair, and thus transits to state 1. The working agent disconnects itself from its depot (i) or current location and connects with the failed node (j). It is assumed that the travel of an agent will not be affected by the failures in the client network. Therefore, agents can connect to failed nodes even if some links or nodes between i and j are failed. This assumption is reasonable for client networks, because the response team should have backup means to complete the response task.

The service agents are initially located at depots (D) and each agent has its own initial depot. We define a many-to-one correspondence relationship between sets of D and N as $(i \in D \rightarrow i \in N)$. This relationship denotes that a node in the client network can be represented as the depot for more than one agent. For example, $i, j \in D \rightarrow \ 'node\ y' \in N$. In other words, even though some agents are deployed at the same physical locations of the client networks, they can be distinguishable by different notations of set D. By using this formulation, we can define that all depots uniform capacity of 1 and only binary decision variables are required.

To model the relation between service agents and service requests (disruptions), we establish a virtual graph $(D \cup F, \{c_{ij}, i, j \in D \cup F\})$ - i, j here are not the index of N. This virtual graph has a directed and weighted edge denoted by $(\{c_{ij}\})$. The service agents travel on $\{c_{ij}\}$ to handle disruptions. Each edge of the virtual graph has two parts: the traveling time from two locations (nodes) and the repair timespan used to remove disruptions.

Weights of virtual graph: $c_{ij} = t_D(i, j) + t_R(j) = \frac{d(i,j)}{v} + t_R(j)$ (2.4)

The traveling time and the distance function between the represented nodes in the client network are denoted by $t_D(i, j)$ and $d(i, j)$, respectively. For example, if $i \in D$ and $i = \ 'node\ k_1' \in N$, the represented node for i is $node\ k_1$. Same goes if $j \in F$ and $v_j = \ 'node\ k_2' \in N$, the represented node for j is $node\ k_2$. Thus, $d(i, j)$ is the distance between k_1 and k_2. The distance function is application specific, where it can be denoted as Euclidean distance, shortest paths, etc. depending on the network application. v is the velocity that the agents travel within the same space of d. $t_R(j)$ is the time required remove disruption at node v_j,if $j \in F$. Otherwise, if $j \in D$, $t_R(j) = 0$. If j is part of an edge rupture, $(\gamma_j = k)$, $t_R(j) = t_R(k)$, which shows the current collaboration lasts the same for the repair at two different sites. Collaboration ability of service agents is specifically designed for N2N services. An edge (i, j) can only be repaired, If:

1. Two agents are connected to each ends (nodes) of the edge
2. The two agents can collaborate within the service team (S) as defined earlier.

2.1.2.4 Prevention of Failure by Service Team

The presence of disruption responders in the node of the client networks enables them to prevent errors from propagating or ever occurring to the supervised node. Supervision, node failure prevention, and edge failure prevention are defined as follows.

A node i, $i \in N$, is supervised if and only if a disruption responder agent is located at the node: there is an inter-edge between this node and the responder.

$$\text{CPS node supervision } (i) = \begin{cases} 1 \ \exists (i, a) \in E \\ 0 \ otherwise \end{cases}, i \in N, a \in A \tag{2.5}$$

An edge (i, j) is supervised if and only if two node incident to the edge are supervised and the supervising responder are able to collaborate.

$$\text{CPS node supervision } (i, j) = \begin{cases} 1 \ \exists (i, a) \in E, (j, b) \in E, (a, b) \in C \\ \qquad 0 \ otherwise \end{cases}, i, j \in N, a, b \in A \tag{2.6}$$

As modeled by the Watts Cascade Threshold model, a node will fail if at least φ fraction of its edges are failed. If an edge is supervised, this edge does not as a failed edge to propagate the failure. Thus, the definition of node failure is updated as follows.

Client node failure:

$$\text{state } (i) = \begin{cases} 1 \ \frac{\sum_{(i,j) \in E} state(i,j) - \sum_{state(i,j)=1} supervision(i,j)}{deg_G(i)} \geq \varphi \\ \qquad 0 \ otherwise \end{cases}, i, j \in N \tag{2.7}$$

A link failure is defined as follows: if only one of its incident nodes is failed, the edge will fail as the other node is not supervised. If both nodes are failed, the edge will fail unless it is supervised (Eq. (2.8)).

Client edge failure: state $(i, j) =$

$$\begin{cases} state(i)x\big(1 - supervision(i)\big) + state(j)x\big(1 - supervision(j)\big) & state(i) \neq state(j) \\ supervision(i, j) & state(i) = state(j) = 1 \\ 0 & state(i) = state(j) = 0 \end{cases}$$

$, i, j \in N$ (2.9)

2.1.2.5 Objective Function

Given all of the DLOC major components have been explained, the mathematical formulation of the objective function is presented below:

DLOC-CDR objective function: $\min z = \sum_{i,j \in F} (\sigma_i - \tau_i)$ (2.10)

s.t.

Response sequence constraint: $\sigma_i \geq \sum_{i \in D \cup F} x_{ij} c_{ij} + \sum_{i \in F} x_{ij} \sigma_{ij} \; for \; j \in F$ (2.11)

Depot visit constraint: $\sum_{j \in F} x_{ij} = \sum_{j \in F} x_{ji} \leq 1, for \; i \in D$ (2.12)

Disruption visit constraint: $\sum_{j \in D \cup F} x_{ij} = \sum_{j \in D \cup F} x_{ji} = 1 \, , for \; i \in F$ (2.13)

Depot-wander elimination constraint: $\sum_{i,j \in D} x_{ij} = 0$ (2.14)

Sub-tour elimination constraint: $\sum_{i.j} x_{ij} \leq |B| - 1, for \; B \subset F$ (2.15)

Repair constraint: $\sigma_i \geq \tau_i, for \; i \in F$ (2.16)

Concurrent collaboration requirement: $\sigma_i = \sigma_i$ for $\gamma_i = j, \gamma_j = i,$ and $i, j \in F$ (2.17)

Decision variables for repair tours: $X = \{x_{ij} = \{0,1\}, i, j \in D \cup F\}$ (2.18)

Cascading function of disruptions: $F = CASCAD(F_{0,}G, X)$ (2.19)

Where z is the total latency for disruption recovery; σ_i is the timestamp when disruption i is recovered; τ_i is the timestamp when disruption i started; x_{ij} is a binary decision variable indicating if a repair-agent should go from node i to node j (i, j are either a depot or disruption) and X is the set of all x_{ij}; B is the subset of any disruption; F_0 is the set of initial disruption at time 0; and $CASCAD$ is a model that generates the entire set of disruption F based on initial disruptions (see section 2.1.2.2).

2.1.3 TIE-DLOC Simulator

A software tool has been developed to facilitate the modelling of various systems with the DLOC model and evaluate different protocols called the TIE/DLOC (Zhong & Nof, 2015; Zhong, 2016). It is based on the previous concepts of Teamwork Integration Evaluator (TIE) for various systems developed at PRISM Center, Purdue University. The TIE tools apply parallel computers to simulate distributed e-Work enterprises, decision makers, agents, or sensors, which are communication and collaborating for a set of tasks. The input/output diagram of TIE/DLOC is illustrated in Fig. 2.1.

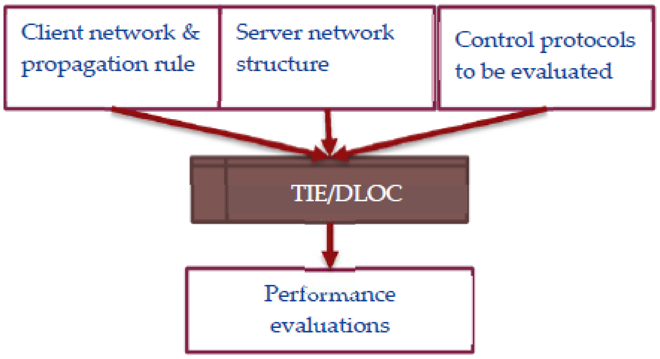

Client network & propagation rule
Server network structure
Control protocols to be evaluated
TIE/DLOC
Performance evaluations

Figure 2.1 Input/Output Diagram of TIE/DLOC (From Zhong, 2016)

2.1.4 Past Developments and Applications

The TIE/DLOC has been previously applied to solve several N2N services challenges listed as below:

1. Collaborative Disruption Response (Zhong & Nof, 2015)

   A collaborative disruption response (CDR) scenario in CPS was simulated on numerical experiment using three different networks models namely, ER: Erdos-Renyi random graph (Erdos & Renyi, 1959); BA: Barabasi-Albert scale-free networks (Barabasi & Albert, 1999; and WS: Watts-Strogatz small-world model (Watts & Strogatz, 1998). The client networks are modelled as people, systems, or other agents having two states: 0 for not requesting service and 1 for requesting service. They are linked into a network through their dependencies or interactions. It is assumed that the client network has a uniformly undirected and unweighted link. The client network will initiate a service request whenever there is a disruption within the network. Disruptions are in a form of link or node rupture within the client network.

   The service request will propagate throughout a network if not handled. In this work, a structure-based model is adopted to imitate the service request propagation. It is assumed that failed elements (nodes/links) will cause structurally connected elements to fail under the percolation theory (Watts, 2002; Bashan *et al.,* 2011).

   The server network is external to the modeled network, but has inter-connected edges with client network, which can be interpreted physically as repairman depot. A collaboration relation is defined in the form of links between the server agents if they have collaborative ability to execute a certain task. If a service depot (server) is positioned (supervising) at a certain node within the network, then the server has error prevention capability attributed to the node. Link rupture prevention can only be achieved if both ends (nodes) of the links are supervised by a server; both nodes have interconnected edges to the server network.

The N2N service challenge defined by this work is optimal resource allocation of the server/responder to manage the propagating service requests. This challenge was further breakdown in three specific questions: (1) How the team configuration should be? (2) How should the service depots be allocated? (3) How to schedule service operations to maximize quality of service?

An Asynchronous Collaboration Requirement Planning (ACRP) framework is established for the construction of reconfigurable service team to provide flexible services to the client (to answer (1)). Depot allocation decisions (2) are implemented by using the Centrality-based depot allocation (CBA) method. In this method, service agents are initially positioned at nodes that have high betweenness centrality (Freeman, 1977). This is given the assumptions that the repair agents can only traverse through the client network using the existing links (regardless of ruptured or not), e.g. repair agents traveling to restore electricity grids. Neuroplasticity-inspired protocols where applied to determine the schedule of providing service requested from the client network (3). This protocol consists of two main components; the activity-based priority (ABP) protocol (the main assignment protocol) and Auxiliary Links (ALs) addition procedure to the client network to improve local efficiency in the recovery operation.

Experiments were conducted on both conceptual complex network models, namely Erdos-Renyi random graph (Erdos & Renyi, 1959), Barbasi-Albert scale free networks (Barabasi & Albert,1999), and Watts-Strogatz small-world model (Watts & Strogatz, 1998), and a realistic case study of a water distribution system. The results of these experiments concluded that the small-world phenomenon (Milgram, 1967; Watts,1998) attributed in the difficulty of removing cascading failures by service agents because disruptions are more difficult to be removed if they propagate to interlinked clusters. A better and more efficient collaboration between the service teams (through training, etc.) can improve the improve response performance to certain upper boundary dictated by the availability of service team resources. Meanwhile, the neuroplasticity-inspired protocol was able significantly reduce the total number of failures, distance travelled by repair agents, and latency compared to First-Come-First-Serve (FCFS) scheduling protocols. The ABP can be applied to any client network, while ALs can only be applied to reconfigurable client networks with acceptable reconfiguration costs.

2. Furniture Manufacturing (Candranegara, Zhong, and Nof, 2015)

   Conflicts and human/machine errors (CE) between operations can propagate and leads inferior products in a furniture manufacturing systems. The N2N challenge in this system is how to efficiently allocate CE detection agents along the production line. Since these resources requires collaboration and services are interdepended, this problem can be solved by a DLOC model. Allocation of inspection resources (the server network) to the manufacturing stations (client network) for efficient detection, prevention, and recovery of CEs were conducted by the measuring eigenvector centrality of each stations (network nodes) with respect to historical occurrence of CEs and the influence the station to other stations. Two scenarios are simulated: inspection by humans or by emerging autonomous systems. Experiment shows that the developed method increases CEPD performance with statistically significance by reducing the time to completion compared with the decentralized method (allocating resources to every stations), and increasing the preventability and reliability while reducing the rectification cost compared with the centralized method (allocating resources at the end of process). This case study validates the DLOC model in a manufacturing client network.

2.2 Graph Theory and Network Science Problems

2.2.1 Review of Graph Theoretical Analysis in Complex Networks

The studies of network to model real-world problems came as early in the 17th century by Leonard Euler. His mathematical description vertices (nodes) and edges that builds network of Konigsberg bridge laid foundation to graph theory. The study of Graph theory does not directly translate into the study of network science and complex network, where the latter focuses on more of the applied application of graph theory tools to model and analyze real-world networks. Initially, real-world networks were thought to be random in their topology, as exemplified by the Erdos-Renyi random graph (Erdos & Renyi, 1959, 1960). In the late 1990s, further advances in network science have found that real-world networks have power-law degree distribution (Barabasi & Albert, 2002), which is not true in random networks that have Poisson degree distribution. Other attributes that were observed were the existence of short links connecting distant part of the network, but still maintain relative local interconnection (clustering). This attribute was later dubbed as the “small-world” property (Watts & Strogatz, 1998). Both of these attributes became canonical for network case-studies where they represented by the scale-free network (Barabasi & Albert, 2002) and small-world network (Watts & Strogatz, 1998).

The study of robustness of networks against perturbation has become attractive lately, mostly related to complex system designs. The common properties measured in studying the robustness of a network are average geodesic length and size of giant component. It has been found that scale-free networks have high degree of robustness against random failure, but extremely vulnerable to failure in its hubs (Albert *et al.*, 2000). The prevalence of community and modularity also effects networks robustness (Newman, 2006; Tran & Kwon, 2013). Modularity is negatively correlated with network robustness. The study of network robustness has also motivated in the development of new network centrality measures (Freeman, 1977). One particularly new centrality measure is the bridge centrality (Hwang et al., 2006). The Bridge centrality has interesting potential of effectively detecting bridging edges & nodes which connects different network modules.

2.2.2 Prior Research Works Related to Cascading Failures in Complex Networks and Vulnerability

It has been observed that real-world CPIs exhibits the characteristics of a cascading failure under perturbation. A profound example of this is the blackout of power grids which happened recently in the USA (Kadloor & Santh; 2010). The significance of this phenomena is that due to cascading failure, a single element failure may result into catastrophic consequence.

The analysis and modelling of cascading failure has only been available in the recent decade. There are two main school of thoughts for modelling cascades in complex networks: the load-based and the threshold-based. The former, load-based method, is based on the concept of dynamical redistribution of flow in networks (Crucitti et. al., 2004). This model commonly used to analyze the cascading dynamics in networks where the elements are subjected to loads, i.e. power grid, water distribution networks (Lv et. al.,2014; Shuang et. al., 2015; Shuang et. al., 2014). However, we will not focus our research in this method since it is not being used by the DLOC. The threshold-based model is developed based on considering the diffusion of information or propaganda in social networks, where each individual's state depends on its neighbors (Granovetter, 1978; Watts, 2002). Under this model, global cascade or full-size cascade occurs according to "cascade window" which varies according to the average degree and threshold value. In the initial model, cascade is triggered by a single failing nodes and the rest will also fail if the threshold is met. Further application of this model has resulted in a number of generalization in different networks. In social contagion (viruses, information, innovation), a generalized model of has been made by integrating interdependent interaction models (Dodds & Watts, 2004). The model has also been generalized using analytical approach in modular networks (Gleeson, 2008), degree-correlated networks (Dodds et. al., 2009), and networks with adjustable clustering (2011). It has also been studied that different failure initiator selection and number would influence the final size of cascade (Singh et al., 2013). This mode threshold lends some similarity with other social contagion models such as the SIS model, where both models consider the fraction of "failure" neighbors to determine a node's probability of becoming fault (Dodds & Watts, 2004). In information diffusion theory, sociologists have long argued that "bridges" between disjoint community clusters promotes the diffusion of information or diseases (Granovetter, 1973). This was further confirmed by the small-world network model (Watts & Strogatz, 1998) where links between otherwise distant nodes are created by rewiring that of a regular graph. It was found that disease infection spreads much easier and quickly in this network. Nevertheless, the type of cascade assumed in the aforementioned studies were *simple propagation*; one "fail" neighbor is sufficient to transmit information or tilt the status of its neighbor into "fail" as well. The other type of cascade is the *complex propagation* where it takes a minimum threshold of neighbors in "fail" status allow a given adopting the same "fail" status as well (Granovetter, 1978). Both of these cascade types are addressed in the Watts Global Cascade Threshold model by adjusting the cascade threshold (Watts, 2002). Complex propagation typically unfolds in clustered networks or within cluster modules of networks. This was exemplified in the studies of recruitment patterns for social movements; they are typically effective in locally dense network of relationship (McAdam, 1986; McAdam *et al.*, 1993). For *complex propagation*, it was shown “bridges” or random edges connecting node clusters can actually inhibit the cascade growth process (Centola *et al*., 2007). In fact, another studies proved that the occurrence of connected clusters in networks are the only obstacles to cascades (Easley & Kleinberg, 2010). Specifically, given the cascade threshold is q , a failure/disease/information cannot propagate into a different node cluster if given the next node cluster has a clustering density (coefficient) greater than $1 - q$.

Despite there have been many advances in study of cascading failures in complex networks, there hasn’t been any research with the integration of DLOC model. The DLOC is a new class of research problems in the emerging e-Work systems; past findings and research regarding cascade behavior can be applied to DLOC-CDR to aid in developing quick design guidelines for disruption responders in CPI networks.

2.2.3 Prior Research Works Related to Disruption Mitigation and Control

Previous works in network disruption mitigation and control can be broadly grouped into two general directions. The first direction discusses about designing robust network through identifying the critical elements of a network to maintain connectivity or connectivity reliability; maintain single component connected graph topology. This approach can also be implied as pre-disruption mitigation approach. The second approach, although not completely exclusive of the first, discusses about post-disruption mitigation. Prior works in identifying critical elements connectivity have mainly focused on the application of design of sensor and radio communication networks. The minimum number of neighbors needed to maintain connectivity in a random radio network was previously investigated through simulation (Ni & Chandler, 1994). This resulted the “magic number” of minimum be neighbors to be three to eight. For wireless sensor networks, the number is estimated as a logarithmic function of the total number of nodes (Xue & Kumar, 2004). Subsequent research works develop these model by taking into account of various conditions. For example, (Dong, et al., 2007) found the lower bound probability of a wireless sensor network being connected under Rayleigh Fading as a function of minimum node density. Connectivity properties was also studied large scale sensor networks as a mean to optimize multi-path routing (Pishro-Nik, et al., 2009). Another approach for pre-disruption mitigation is by providing redundancy. (Chen and Nof, 2000; Chen, 2002) investigated genetic algorithms to be used in modelling low-cost fault tolerant structure of Multi-Enterprise Networks. Inspired by the Fault Tolerance by Teaming principle of CCT (Nof, 2007), a new design paradigm called Resilience by Teaming (Reyes Levalle & Nof, 2015) has been validated on several supply networks to provide better network resiliency under disruptions.

Connectivity reliability analysis has also been done on several other real-world networks. In transportation networks, a research has been conducted to aid post-disaster road network recovery decisions (Bin, et al., 2009). The road network was modeled as a weighted flow graph, whereas the flow represents time-varying traffics, and by assessing connectivity reliability of different recovery scenarios an optimal decision can be found to minimize total travel time cost between each pairs of nodes. Optimal resource allocation (cost) for partial road network recovery has also been investigated using Lagrangian based heuristic algorithm (Liu & Qi, 2014).

Still under the category of pre-disruption mitigation strategies, other works have investigated network design for self-healing telecommunication networks by utilizing spare capacity planning. The network in these works are modeled as bi-directional weighted networks having multiple commodities between different source and sink nodes (multi-commodity flow problem). There are two main basic methods devised by these works, namely line restoration and path restoration. Link restoration allocates spare capacity to the links so that a faulty link's flow can be rerouted through an alternate path using the spare capacities of the links in the network (Veerasamy & Venkatesan, 1995). Other works under these methods have mainly focused on developing algorithms and heuristics to compute optimal rerouting policy (Krishnamurthy, et al., 2003; Grover, et al., 1991; Sakauchi, 1990). Path restoration on the other hand, considers each path disrupted by the link failure separately and rerouted over an alternate path between the source and sink nodes (Murakami & Kim, 1998; Doshi, et al., 1991; Grover, 2000). This method, although requires more computation power, results in a more efficient spare capacity planning. Research developments in pre-disruption mitigation strategies have equipped network planner with better insight on the design of a more resilient client network, especially in weighted networks (i.e. telecommunication networks). However, the N2N challenges that DLOC addresses requires the properties of Online Service and Cascading Failures (Zhong & Nof, 2015; Zhong, 2016). These properties have not been addressed in the aforementioned works.

Post-disruption mitigation strategies research works are commonly found in protection of vital infrastructure networks. In Water Distribution Networks (WDN), an emergency model was developed to redistribute water pressure and flow to prevent cascading failure due to overload (Shuang, et al., 2014). Nodes in the network represents water reservoirs, consumers, and tanks, while edges represent pipes, pumps, and valves. The loads (flows) are assumed to be dynamic, such that it will cause edge failure if exceeds the flow capacity. In this model, external emergency resource exists to fix failed elements (nodes/edges) and are triggered by a certain threshold. However, external resources here assumed to be unlimited. Power grids utilizes a load shedding strategy to balance overall demands with electricity availability. In this case, the power grid network is abstracted as a graph, where the nodes represent buses (loads and generators) and edges represent electricity lines (Xu & Girgis, 2001; Aponte & Nelson, 2006; Bevrani, et al., 2010). Another method is proposed where the removing of "insignificant nodes" that a contribute more load to the network than they handle is removed to reduce the size of cascading failure (Motter, 2004). There are still several other works not mentioned here which also discusses about post-disruption mitigation strategy. However, for the interest of conciseness, the literature survey presented here have fairly represented the general types of work previously done.

In summary, the post-disruption mitigation strategies presented gives a robust method to mitigate disruptions, especially for client network control during disruption. However, they have not yet included the external service network aspect of the DLOC.

2.2.4 Prior Research Works Related to Service Resource Allocation and Network Component Protection Priority

Service resource allocation problems found in previous works are commonly related to facility location problem or k-center problem in graph theory. Given a weighted network, this problem is concerned with optimal placement of facilities to minimize transportation cost across the network. There has been copious amount of research in this problem and exact and approximate algorithms have been found to fine the optimal placement (Current, et al., 1990). A subset of this problem deals not only with minimizing cost, but also maximizing coverage; the common name for these set problem the maximum coverage/shortest path problem (Current, et al., 1985). Location and covering problems in undirected and directed flow networks have been studied in (Tamura, et al., 1990; Tamura, et al., 1992), the optimal solution of both of these problems can be obtained in polynomial time. Resource allocation (facility location) for post-disaster management also requires the facility to have maximum coverage on the affected area with respect to minimum routing cost (Viswanath & Peeta, 2002). One of the key element in this work was the resource constraint on the number of resources to be deployed vs minimum cost routing which was solved using integer programming. A sensor location problem in traffic networks have also been investigated to find the minimum number of sensors such that information on flow volume in specific path can be obtained (Gentili & Mirchandani, 2005). The maximal coverage/shortest path problem aligns with resource allocation problems applied in DLOC, where server have to be initially positioned in nodes that minimizes overall expected routing cost as well having maximum coverage of the network in terms of group allocation. However, the approach employed on this problem cannot be directly adopted to the collaborative disruption response problem in DLOC. As stated in (Zhong & Nof, 2015; Zhong, 2016) the positioning of server agents in client networks also functions as error prevention on the supervised nodes/links. Thus, there is another objective of maximally positioning the server agents to protect important nodes.

Due to cascade of overload failure, the highly heterogeneous distribution of *loads* of real-world network makes them vulnerable to attacks such that an avalanche of failure nodes (cascading) may occur by disabling a single (or several) key nodes (Motter & Lai, 2002). A better protection strategy of client network can be developed by also taking into account this fact, i.e. priority protection on vulnerable nodes. Several research works have dedicated to investigate network survivability under the failure of these set of nodes and identify them for priority protection. (Cruciti, et al., 2004) showed that in weighted networks, the vulnerable nodes are the ones with the largest load. In fiber infrastructure network, a polynomial time algorithm has been developed to simulate several node failures (single or set) to identify the vulnerable nodes (Neumayer, et al., 2011). Vulnerable nodes in directed water distribution network where identified by assessing the ratio between discrepancy of failures cascade & direct failure and total number of initial node (Shuang, et al., 2014). The previously aforementioned works have investigated node vulnerability under the consideration of cascading overload failures, typically in flow networks. Notwithstanding the importance of flow continuum in networks, vulnerable nodes can also be identified by analyzing the network topological structure. Several measures graph (network) centrality measures have been developed, including degree centrality and betweenness centrality (Freeman,1977; Borgatti, 2005; Newman, 2001). Nevertheless, these previous centrality measures are dominated by elements' degree due. A newly developed measure, named bridging centrality measures, aims to identify the most important component in the networks by exploiting graph properties of cut edges/vertices and clustering (Hwang et al., 2006). Cut vertices/edges denotes the elements of a graph (network) where if removed would increase the number of connected components. This essential in real-world networks which mainly exhibits a modularity structure (Newman, 2006).

In summary, previous research works related service resource allocation and network component protection priority have given significant contribution in understanding the failure dynamics and protection of complex network. Nevertheless, these two areas are still disaggregated in approaching collaborative disruption response – although they both hold important findings to improve collaborative disruption response. Thus, in the following chapter we will try to improve the service allocation protocol in DLOC by taking into account both the optimal service allocation as well as differentiated network component protection priority.

CHAPTER 3. GRAPH –THEORETIC PROPERTY ANALYSIS OF DLOC

This chapter is dedicated for RQ1. We will revisit the results of previous research works on DLOC (Zhong & Nof, 2015; Zhong, 2016) and use tools from graph theory to develop inferences and quick guidelines for the design of disruption response operation in CPI network. The main contribution of this chapter is to find a general pattern of graph (network) properties which can approximate and explain the results of the DLOC experiments on conceptual networks and validate it on real-world CPI networks. The conceptual networks that will be used in this research are Erdos-Renyi Random Network (ER), Barabasi-Albert Scale-free Network (BA), and Watts-Strogaz Small-world Network (WS).

3.1 Phase Transition for Probability of Recoverability ($P_{recover}$)

The threshold cascade model of the DLOC has been observed to display a phase transition phenomenon governed by the critical value of cascade threshold (φ), average degree ($\overline{deg_G}$), and failure initiator fraction (p_{F_0}) at which beyond the critical mark, a global cascade first commence (Watts, 2002; Singh, 2013). Global cascade is defined as cascade size covering >90% of the networks elements. Motivated by this phenomena, we hypothesize that there's exist a certain regime where the probability of a CPI network (under the DLOC model) to fully recover from a cascading disruption is strictly less than 1, $P_{recover}$<1, and vice versa. Our initial analysis presumes that it is governed by the disruption responder team size ($|A|$), average degree ($\overline{deg_G}$), and cascade threshold (φ) – while keeping the failure initiator fraction (p_{F_0}) fixed at certain value. We will be validating this hypothesis numerically by using the TIE-DLOC simulator on conceptual networks and a power grid network.

3.2 Probability of Recovery from Cascade ($P_{recover}$) in DLOC

In the presence of cascading failures, the client network, according to the DLOC model, will request service to the server network (disruption responders) to start the recovery process from failures. The resiliency of client CPI network to recover from impending disruption is addressed in DLOC by the Recoverability $P_{recover}$ metric (Zhong & Nof, 2015; Zhong, 2016). It can be formulated as:

$$P_{recover} = P(|F_t| = 0, t > 0) \tag{3.1}$$

Where $|F_t|$ denotes the dynamic cascading failure size at time t. The realization of $P_{recover}$ from the TIE-DLOC simulator is an average value over the number of replications.

3.3 Analysis and Assumptions

Analytical models have been developed predict the average cascade size for threshold model in different networks. (Watts, 2002; Dodds & Watts, 2004; Gleeson, 2008; Hackett et al., 2011). Let S_t denote the subset of nodes in $G(N, E)$ which has already been infected by the cascading failure (cascade size) at time t, $S_t \subset N$. By construction of the Watts threshold model, $|S_{t+1}| > |S_t|$ holds only if there exist a node $v, v \notin S_t$ adjacent with at least one node in S_t and has a degree $(deg(v)) \leq \left\lfloor \frac{1}{\varphi} \right\rfloor$. We proceed to call these nodes which has potential to propagate failures as *vulnerable* nodes.

Several additional assumptions are applied to this study to limit the scope of analysis:

1. Disruption responder team formation

   We assumed a fixed the number of disruption responder team $|A|$ to be 0.07*$|N|$. Furthermore, it is assumed that each agent has the ability to collaborate with 1/3 of the disruption responder team, or $deg_s(A) = \frac{|A|}{3}$.

2. Edges of client network

   Although the adopted Watts threshold Cascade Model in DLOC takes into account of failures involving the edges of the network, we concede that this will not be taken into account. It can be argued that this will not influence the analysis significantly since the failure state of the edges depend solely on its adjacent nodes.

3. Failure initiator (F_0)

   For all network tested, we assume that initial failures occur randomly on 5 client network nodes.

4. Disruption Responder Allocation

The disruption responder agents will initially be allocated to the nodes of client network with the highest betweennnes centrality value. This method is known as the Centralit-Based Depot Allocation, CBA (Zhong & Nof, 2015; Zhong, 2016).

We proceed by presenting two hypotheses that can approximate the phase transition of $P_{recover}$:

Hypothesis 3.1

$$\overline{deg_G} > \frac{1}{\varphi} \rightarrow P_{recoverability} = 1 \tag{3.2}$$

By the construction of the Watts threshold cascade model, we hypothesized that the expected degree of node should be more than the reciprocal threshold cascade φ.

Hypothesis 3.2

$$P_{vulnerable} = \sum_k p_{v(k)} p_k < 0.70 \rightarrow P_{recover} = 1 \tag{3.3}$$

$$p_{v(k)} = 1 \; if \; k \leq \left\lfloor \frac{1}{\varphi} \right\rfloor, \text{0 otherwise} \tag{3.4}$$

We denote that if the fraction vulnerable nodes ($P_{vulnerable}$) is less than 0.70, then the network has a probability of 1 to recover from disruption. p_k denotes the probability that a node has a degree k, while $p_{v(k)}$ is binary variable conditioned on whether a node with a degree k is vulnerable based on the threshold cascade. This relation is not bidirectional.

3.4 Experiment: TIE-DLOC Simulation

We validate Hypothesis 3.1 and 3.2 through numerical simulation using TIE-DLOC simulation software. The data set is divided into two groups, namely conceptual networks and real-world networks. The latter consist of three different conceptual networks: Erdos-Renyi Random Network (ER), Barabasi-Albert Scale-free (BA), and Watts-Strogatz Small world (WS). Each network will have 2 combinations of $|N|$ and $\overline{deg_G}$ summarized in Table 3.1. Each network will also be tested on varying cascade threshold (φ): 0.25 and 0.4.

Table 3.1 Combination of |N| and $\overline{deg_G}$ for conceptual network

Combination	Number of Nodes $\lvert N \rvert$	Average Degree ($\overline{deg_G}$)
1	500	4
2	1000	6

The real-world network will be represented by a network model of the USA Western States Power Grid (PW). There are 4941 nodes in this network connected by 6594 edges. The average degree of this network is 2.67 (Watts & Storgatz, 1998). For the real-world network, the threshold cascade will be set at 0.25 and 0.4.

The simulation result is summarized in Table 3.2. A general pattern can be inferred from this result is that $P_{vulnerable}$ correlates negatively with $P_{recover}$ with a strong correlation. This is justifiable by the fact that $P_{vulnerable}$ correlates positively with total cascade size F.

Hypothesis 3.1 is negated by the fact that it does not applies on all experiment results. Specifically, on experiments no. (1) and (5), the average degree $\overline{deg_G}$ is equal to $\frac{1}{\varphi}$, i.e. $\overline{deg_G}$ =4 for $\varphi = 0.25$, but $P_{recover}$ is 1. We also conclude that the approximation proposed in Hypothesis 3.1 is risk-averse due to the fact that it underestimated the result of experiments (1) and (5). Nevertheless, this approximation can be a quick guideline for decision makers and designers to know the resiliency of their network without too much of analysis beforehand because it only requires the input of average degree and expected threshold cascade.

Hypothesis 3.2 is found to be valid in all set of experiments. All networks that have $P_{vulnerable} < 0.70$ has a probability 1 of recovering from a cascading failure. One important note is that by construction of the hypothesis we impose a logical relation of $P_{vulnerable} < 0.70$ would produce $P_{recover}$ = 1, but not the other way around (not bidirectional). Thus, experiment (5) still applies to this hypothesis.

A caveat of this experiment is the interesting pattern found in small-world networks (WS). Results past experiment of DLOC have found that the WS to be the most vulnerable compared to ER and BA (Zhong & Nof, 2015; Zhong, 2016). The same setup of this experiment was used for experiment no (1), (5), and (9). However, it can be seen when φ is changed to 0.4, WS has the lowest $P_{vulnerable}$ among ER and BA. This implies that the vulnerability statute has changed. This result also complies with the findings that “bridges” or “shortcuts” found in small-world networks actually inhibits the spread of complex propagation (Centola et al., 2006) – as $\varphi \gg$, propagation becomes more complex because it requires more neighbors to influence change of state. Finally, we also conject that exact numerical relation between $P_{recover}$ and $P_{vulnerabiltiy}$, i.e. relation in regards to exact numerical value, is dependent on the different network topology. This can be seen by comparing experiment no (5) and (9) where the former has higher $P_{vulnerabiltiy}$ than experiment (9) but higher $P_{recover}$ as well. Both network have different threshold of $P_{vulnerabiltiy}$ to allow $P_{recover}$ to phase into 1.

Table 3.2 Experiment Result for $P_{recoverability}$ based on average of 50 replications

No	Network	Cascade Threshold	Number of Nodes	Fraction of Vulnerable ($P_{vulnerable}$)	Average Degree ($\overline{deg_G}$)	Probability of Recovery ($P_{recoverability}$)
1	ER	0.25	500	0.632	4	1
2	ER	0.25	1000	0.292	6	1
3	ER	0.40	500	0.252	4	1
4	ER	0.40	1000	0.065	6	1
5	BA	0.25	500	0.782	4	1
6	BA	0.25	1000	0.596	6	1
7	BA	0.40	500	0.488	4	1
8	BA	0.40	1000	0	6	1
9	WS	0.25	500	0.72	4	0.96
10	WS	0.25	1000	0	6	1

11	WS	0.40	500	0.058	4	1
12	WS	0.40	1000	0	6	1
13	PG	0.25	4941	0.88	2.67	0.9
14	PG	0.4	4941	0.58	2.67	1

CHAPTER 4. RESOURCE ALLOCATION POLICY FOR PROTECTION OF VULNERABLE NETWORK ELEMENTS

Advances in network science have found that networks found in real world have varying properties which cannot be modeled by a classical random graph (Watts & Strogatz, 1998; Barabasi & Albert, 1999). Some of these properties that are prominent in man-made CPIs networks includes power-law degree distribution due to evolutionary networks (growing networks) with preferential attachments, small-world properties which elucidates the "six-degree of separation" phenomena and shorter average paths (Barabasi & Albert, 1999; Watts & Strogatz, 1998; Dorogovtsev & Mendes; 2002). These properties have an inherent impact on the heterogeneity of the network elements which also has an effect to how network respond to disruption (tolerance). For example, it is found that CPI networks having scale-free behavior has higher disruption tolerance against random errors but vulnerable against disruption to the network centralities which plays important roles in maintaining connectivity. The DLOC-CDR further elucidates this by assuming that CPI networks maintain cascading failure behavior disruption occurrence. Thus, if a failure happens to one of the vulnerability points, the results would be even more catastrophic. This chapter's focus will start from addressing the vulnerable points fact to create a better resource allocation policy for DLOC-CDR operations in CPIs under the DLOC model.

Compared to previous research in DLOC (Zhong & Nof, 2015; Zhong, 2016), we hypothesize that disruption responders must be allocated to the most vulnerable parts of the networks, such that preventing them ever to fail initially, as opposed to allocating based on reachability or shortest paths. An analysis of error vulnerability, therefore, will be conducted based on several targeted failures selection heuristics. The result of this analysis is expected to provide initial insights and further reinforce the hypothesis above to guide the development of the of a new and better resource allocation policy.

4.1 Vulnerable Points of Network

The objective of this study is to elucidate the effect of targeted failures in CPI networks using conceptual complex networks models and gain insight on vulnerability points of network. As mentioned before, we will use the insight gained from this study as a guideline in developing a better resource allocation policy which takes into consideration of prioritizing protection of vulnerable network elements (nodes & edges). The conclusion of this study will highlight the effectiveness of proposed network centrality measures in identifying vulnerable network elements; measured by average cascade size and speed with respect to differing network properties.

4.1.1 Design of Study

Vulnerability denotes the decrease of network performance due to random or selected removals of nodes and edges (Barabasi & Albert, 1999). In this research, we define vulnerability points of the network as a set of nodes (and edges) whose removal will maximize cascading failure effect within the network, although not necessarily result in a global cascade (Watts, 2002). Ideally we would require detail knowledge of the whole network topology to locate and pinpoint each vulnerable set of network elements. This is a tedious and time-consuming process. A more manageable choice is to look upon network centrality heuristics, which will also be the base of our newly improved resource allocation policy.

Network centrality reveals the importance of an element (node or edge) within the whole network, and many centrality measures have been developed based on structural information of the network (Borgati, 2005; Lee, 2012). For this research, we will compare three different centrality measures as explained below:

1. Betweenness Centrality (Freeman, 1977)

   This centrality measures the share of times that all shortest paths pass through the node being measured. The betweenness centrality value of node $i \in N$ can be expressed as:

   $$B_i = \sum_{\substack{s \neq i \neq t \\ s,t,i \in N}} \frac{\rho_{st}(i)}{\rho_{st}} \tag{4.1}$$

   Where ρ_{st} is the number of shortest paths from node s to t and $\rho_{st}(i)$ is the number of shortest paths from node s to t that passes through i. Previous research on DLOC (Zhong & Nof, 2015; Zhong, 2016) have applied this centrality measure as a heuristic for resource allocation policy under the CBA policy. It has also been proven to yield a better result on all dimensions of DLOC performance metrics (Time, failure, distance, and preventability) compared to random allocation.

2. Degree Centrality (Freeman, 1978; Newman, 2004)

The degree centrality measures the share number of edges connected to the node being measured. For a given node $i \in N$, the degree centrality can be mathematically expressed as:

$$Dg_i = \frac{\deg(i)}{|E|} \tag{4.2}$$

Where $\deg(i)$ is the degree of node i and $|E|$ is the total number of edges of graph G.

3. Bridge Centrality (Hwang et. al, 2006; Nanda & Kotz, 2012)

A bridge node or edge is a network component which connects two modular structure (connected components) in a graph network. They are commonly known as cut vertices or edges (Bondy & Murty, 2008). If a bridge node or edges fails, it will have higher probability to reduce the connected network into a disconnected network and increase the number of connected components. This arguably causes higher damage to the central network with more nodes losing connection to each other. Bridging centrality of a node and edge are defined in Eq. (4.3) below:

$$Br_i = B_i . C_{Br(i)} \tag{4.3}$$

Where Br_i denotes the bridging centrality of node $i \in N$ of client network G. The bridging centrality of node i then is regarded as the product of its betweenness centrality (B_i) and bridging coefficient ($C_{Br(i)}$). The bridging coefficient of a node determines the extent how well node is located between high degree nodes. It is defined as below:

$$C_{Br(i)} = \frac{\deg(i)^{-1}}{\sum_{j \in R_i} \frac{1}{\deg(j)}} \tag{4.3}$$

Where R_i is the set of neighbors of node i and $i, j \in N$.

Three conceptual network models are selected in this research to represent different characteristics of observable network properties, they are Erdos-Renyi Random Network (ER), Barabasi-Albert Scale-free Network (BA) and Watt-Strogatz Small-world Network (WS). The unique characteristics of each model is the result of their differing generating mechanism. The ER model is generated by connecting nodes randomly. Each pair of nodes in the graph has a probability p to be connected with an edge and independent with every other pair. By probabilistic methods, the expected number of edges of an ER network $G(N, E)$ is $\binom{N}{2}p$. The resulting degree distribution of this network will be a Gaussian bell-shaped curve, implying the ER network has low heterogeneity in node degree. This degree distribution tends to be unrealistic when modelling real-world networks (Albert & Barabasi, 2002). The ER network also has average geodesic length and clustering coefficient which are shorter and lower, respectively, compared to most real-world networks. Regardless of its shortcoming, the ER model is widely used a benchmark for comparison with other network model.

The other two network models, BA and WS, were developed to fill in the network property gap that the ER model had compared to real-world networks. The BA model generates network with the power law degree distribution (“scale-free”) by using a preferential attachment mechanism. Many real-world networks were observed to have power-law degree distribution and thus this model was established (Barabasi & Albert, 1999). The BA model has low clustering coefficient which scales to zero as $N \rightarrow \infty$. Finally, the WS model generates a network with high clustering coefficient and low average geodesic length. This is achieved by rewiring the nodes of a regular ring lattice with a specified probability. The shortcoming of this models is it tends to produce an unrealistic degree distribution due to its ring lattice structure. Nevertheless, this model successfully replicates the "small-world effect" in which the average geodesic length of the network scales proportionally with the logarithm of nodes (Watts & Strogatz, 1998). This effect is known to be prevalent in many social networks. Both the BA and WS model also maintain a hub-spoke architecture, although WS has a higher modularity.

As mentioned in the beginning of this research, the error propagation in this study will be modeled using an adopted Watts Threshold model (Watts, 2002). All nodes and edges – except a handful of initiators - will initially be in normal state "0" and will convert to failure state "1" if the threshold fraction, φ, of neighbors in state "1" is reached. Based on this model, we will evaluate the number of network components (nodes and edge) that are induced to state "1" by error propagation from initiator nodes over time (hours) and without repair.

Complex networks display a high degree of error tolerance against random attack, but are prone to targeted attack on their important nodes. (Albert et al., 2000; Cruciti et al., 2002). Furthermore, a recent study on threshold-limited error spread (propagation) to ER networks has found that selecting initiator node based on degree centrality will yield a higher average error propagation compared to random selection (Singh et al., 2013). Therefore, we claim the random selection method of initiator node will not perform better than the three centrality methods (Degree, Betweenness and Bridge) that has been proposed earlier in this study. By conjecture, this claim is also more apparent in BA and WS networks since the former has a power law degree distribution – a small number of nodes are connected to most of the other nodes - and the latter owes to its high clustering coefficient; both properties implies some component within the network is more important than the others. We further establish several hypotheses that are to be verified in this study.

Hypothesis 4.1

Failures initiated on a set of nodes with high degree centrality would yield a higher average and total cascading failure size compared to failures initiated on the same number of nodes with high betweenness centrality. Let $|F_{(\alpha)}|$ and $|F_{t(\alpha)}|$ be the total number of failures on the network at the end of cascade process and dynamic size of cascading failure at time t, respectively, triggered by failures of nodes having the highest α centrality. This hypothesis can be expressed as:

$$|F_{(degree)}| \geq |F_{(betweenness)}| \tag{4.4}$$

$$|F_{(degree)}| \geq |F_{(bridge)}| \tag{4.5}$$

With Eq. (4.4) & (4.5) having possibility of maintaining equality depending average degree of the network. This observation is supported by past studies which found that within the network same network size (number of nodes) and type, varying average degree affects the cascading behavior of different centrality-based initiator (Singh et al., 2013).

This hypothesis can be theoretically verified by the construction of the Watts Threshold Cascade model. A node will change its state to “1” if the fraction of its neighbor adopting state “1” exceeds the fraction φ. Thus, a node’s degree (number of neighbors assuming no looping edges) plays an essential role in determining cascade failure tolerance of a given node. As a consequence, high-degree node (high degree centrality nodes) are harder to influence by cascading failures. The degree centrality heuristics will select high-degree centrality nodes as error initiator for the cascading failure process. Based on the previous arguments, these nodes are the ones harder to influence by normal propagation and, by the virtue of high degree, are capable of influencing larger number of nodes with lower average degree - more prone to cascading failure.

<u>Hypothesis 4.2</u>

The WS network model would be more vulnerable by error initiated on high bridge centrality nodes compared to betweenness centrality node. This hypothesis can be expressed as:

$$|F_{(bridge)}| \geq |F_{(betweenness)}| \tag{4.6}$$

$$|F_{t(bridge)}| \geq |F_{t(betweenness)}|, \forall t > 0 \tag{4.7}$$

The same argument on equality for Eq. (4.4) also applies for Eq. (4.6). On the contrary to BA, the WS has high clustering coefficient which by construction we deduce is suitable for the bridge centrality to find bridging nodes between clusters.

4.1.2 Experiment: Simulation Result

The objective of this experiment is to compare the effectiveness of the three selected network centrality heuristics on pin-pointing vulnerable parts of the network with respect to the conceptual network models. The metrics that will be evaluated in this experiment are the total number of failures, $|F|$, and dynamic size cascading failure over time t, $|F_t|$. We assume that the three conceptual networks have the same number nodes, $|N| = 500$. The average degree, $\overline{\deg_G}$, of the client network $G(N, E)$ will be varied within certain range to overcome bias of differing network formation due to degree variation. The cascading threshold φ is set to 0.2 and cascading failure process is initiated by selecting and turning 5 nodes into failure state “1” based on the centrality heuristics. We limit the observations of the experiment up to 30 time units (hours).

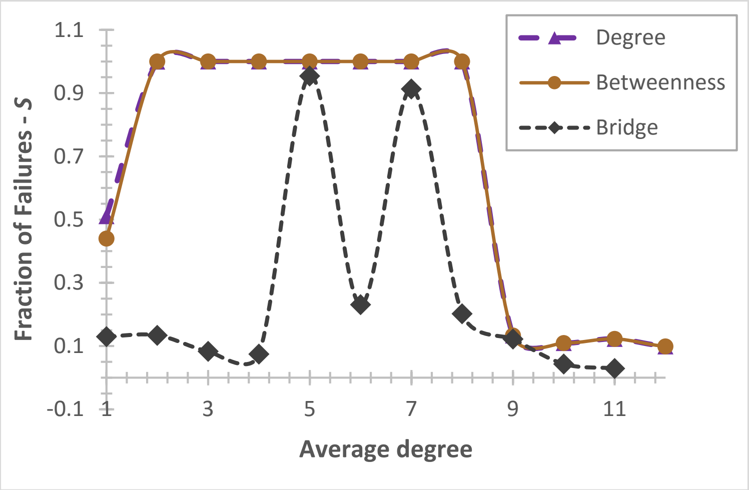
Figure 4.1 summarizes the experiment result measured by Fraction of failures S which is obtained by dividing total number of failures $|F|$ with the total number of nodes $|N|$ and edge $|E|$. In general, all conceptual network with degree centrality failure exhibits a pattern of “global cascade window” (Watts, 2002; Watts, 2007). The “global cascade window” is an intermediate range of $\overline{\deg_G}$ where global cascades are realized.

Hypothesis 4.1 generally applies to all of the tested networks. For BA scale-free network, the degree centrality heuristic outperforms the betweenness centrality heuristics when the average degree is small but equalizes as the average degree scales, which implies nodes of both centrality overlap each other in that region. This overlapping phenomenon between centrality and betweenness heuristics can be explained by several attributes of the BA scale-free network: (1) High-degree hub nodes, existent due to power-law distribution, connects a large number of small-degree node which allows short path length between these nodes by traversing through the hubs. The hub nodes, consequently, have high fraction of shortest paths going through them. This becomes more profound as the hubs are connected to more nodes as a function of the average degree (Albert & Barabasi, 2002; Cohen & Havlin, 2003) (2) The BA scale-free networks and networks with scale-free degree distributions in general shows high correlation between their degree and betweenness centrality values; the higher degree, the higher the betweennesss (Holme et al., 2002). The aforementioned arguments may underline the superiority of the degree centrality compared to the betweenness centrality with respect to the objective of this study. However, it can also be seen on Figure 4.2(a) that the degree centrality failure propagates faster throughout the network compared to the betweennes centrality in within the low degree range, albeit the same final cascade size. Real-world CPIs networks tend to have low average degrees, e.g. Power grid (Watts and Strogatz, 1998). Thus, this low average degree range is crucial for applications in DLOC-CDR. On the other hand, the low performance that the bridge centrality yields in BA networks may be attributed to its construction. The bridging coefficient $C_{Br(i)}$ is a multiplier to the betweeenness centrality value, which values highly nodes that are within the intersection of high degree nodes. This is, however, counter-intuitive with regards to the network topology of the BA network for cascading failure. Cascading failure initiated from these bridge nodes have low probability of influencing (fulfilling the threshold fraction φ) of the much higher degree hubs to change state, i.e. other nodes are miniscule in terms of degree compared to the hub. BA network also tend to have low clustering coefficient which causes “bottlenecks” during failure propagation to other members of the hub cluster, i.e. minimum/no edge-disjoint paths between bridge centrality and hub cluster (Figure 4.3). In conclusion, these arguments lead to the fact that among the network centralities tested the degree centrality most effectively identify the vulnerability points in BA scale-free networks.

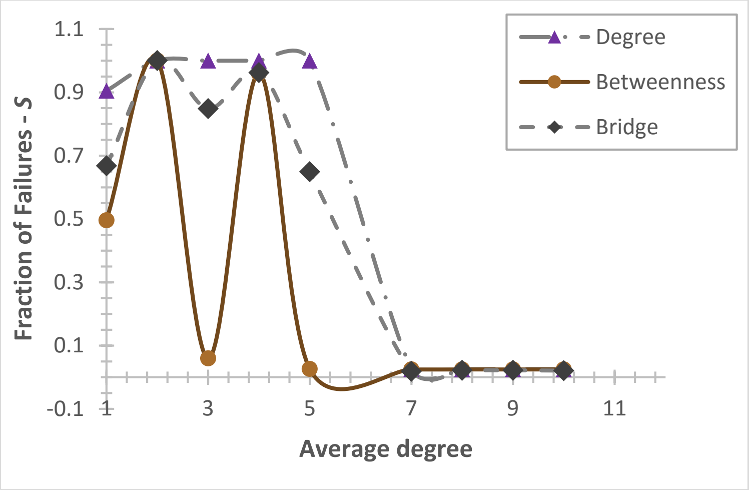
Both hypothesis 4.1 and 4.2 applies for the result of WS small world network (Figure 4.1(a)). The degree centrality displays its superiority compared to the two others centrality measures, which verifies the theoretical proof of hypothesis 4.1. The performance of betweenness centrality and bridge centrality seem to follow a same pattern of performance within a given range of average degree. Nevertheless, the bridge centrality performs equally or better on all of tested average degree range compared to betweenness centrality, which also verifies hypothesis 4.2. It is concluded that the degree centrality and bridge centrality more effectively identify the vulnerability points in WS small world networks compared to betweenness centrality.

For the ER random network, both heuristics (between and degree) perform equally well when average degree is small. This is because the constructed network only consists of small clusters without a giant component (Bollobas, 1984). Thus, spread is localized only to those small clusters and the rest follows the “global cascade window” phenomenon (Watts, 2002; Watts, 2007). The experimental result on this network also verifies

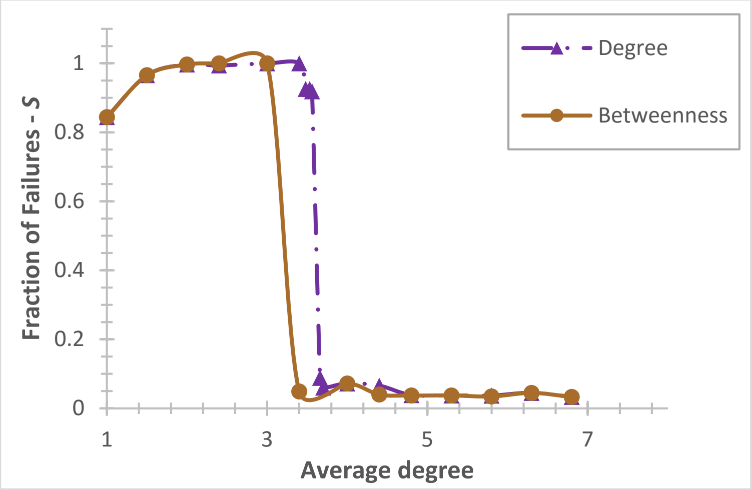
Hypothesis 4.1 in terms of the superiority of the degree centrality compared to the others (see Figure 4.1(c) and 4.2(b)).



(a) BA Scale-free

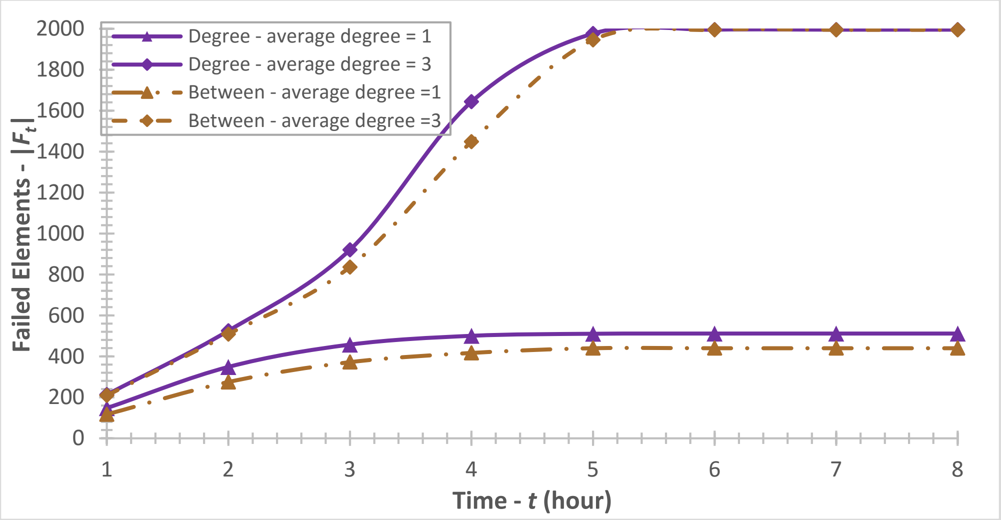


(b) WS Small World

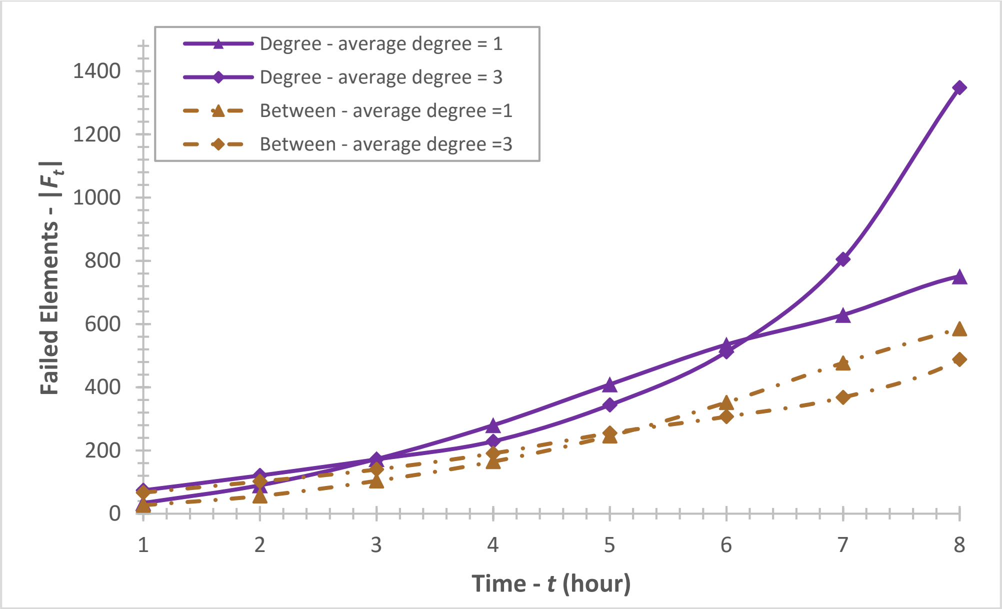


(c) ER Random Network

Figure 4.1 Cascade Failure Size S as a function of average degree for different selection heuristics



(a) BA Scale-free



(b) ER Random Network

Figure 4.2 Dynamic Cascade Failure size $|F_t|$as a function of time t for different selection heuristics and average degree 1 and 3

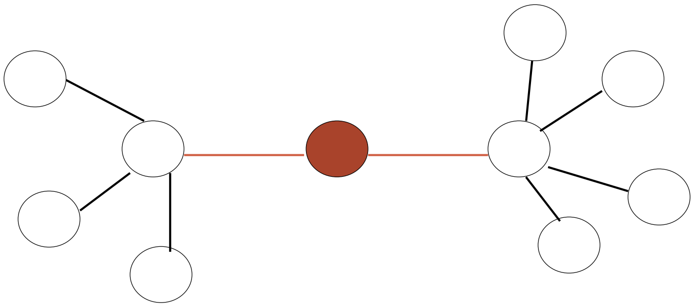


Figure 4.3 Failure propagation from bridge centrality node (colored red) reaches bottleneck to neighboring clusters in BA

4.2 DLOC Model Simulation: Conceptual Networks

In this section, we will use the insights gained from section 4.1 on vulnerability point of network to analyze and develop a better resource allocation policy for collaborative disruption response (CDR) operation under the DLOC model. In this regard, we will set the current DLOC-CDR resource allocation policy as a baseline (will be explained later) and failures will be initiated randomly throughout the simulation replication. It is assumed that the CPI network will have 5 nodes as failure initiators. The TIE-DLOC software (Zhong & Nof, 2015; Zhong, 2016) will be used to simulate the DLOC-CDR operation in this experiment to numerically verify the results of the newly developed resource allocation policy

The current resource allocation policy that DLOC-CDR uses, CBA, is based on the betweenness centrality measure. The formulation of CBA is defined as below:

$$CBA\ policy: \ i = v, B_v \geq \max(B_a)\,, i \in D, v \ \in N, a \in N\ and\ a \not\subset D \tag{4.8}$$

$$i \neq j,\ \forall i, j \in D \tag{4.9}$$

Where B_v is the betweenness centrality of node v and D is the set nodes in $G(N, E)$ which are set as depots (initial locations) for disruption responders (Zhong & Nof, 2015; Zhong, 2016). CBA policy is a greedy heuristic to minimize response time by positioning the service team depot nodes which are intersections of shortest paths. The performance metric that will be measured by the TIE-DLOC simulator are:

1. Total response time (z) :

   This metrics indicates response effectiveness of the responder team of $S(A, P)$. Let F be the set of disruption occurred within the network and $\sigma_i \& \tau_i$ denotes the timestamp when disruption i is repaired and initial timestamp of disruption i, respectively. The total response time z can be formulated as:

$$z = \sum_{i \epsilon F} (\sigma_i - \tau_i) \tag{4.10}$$

2. Total Failures ($|F|$):

   This metric indicates the total set of failed elements during the cascading failure process.

3. Total distance travelled by responders (Δ):

   The total distance relates to the disruption responder agents' movement within the network to rectify failure/disruption; agents' movement follows the network's topology, i.e. edge-node connection.

4. Preventability (P_{rev}):

   Preventability measures the impact of applying disruption response mechanism to the network. Let $\tilde{F}$ denote the set of disruption that would have occurred without a disruption response mechanism. Preventability can be formulated as:

$$P_{rev} = \frac{|\tilde{F}| - |F|}{|\tilde{F}|} \tag{4.11}$$

The aforementioned performance metrics will help verify the effectiveness of the newly developed resource allocation policy.

The analysis and development of the new resource allocation policy will be divided into three phases: (1) Experiment with static disruption responder team size, (2) Experiment with varying disruption responder team size (3) Experiment with varying initial failure size. We will compare the performance of three depot allocation heuristics (based on centrality): Betweenness centrality (CBA), Degree Centrality (DBA), and Bridge Centrality (BBA). The heuristics will be applied on the three conceptual networks previously tested: ER, BA, and WS. These conceptual networks can approximate current emerging CPIs with certain accuracies (Surana, et al., 2005; Yagan, et al., 2012; Chen & Nof, 2012). The second and third phase will build upon the result of the first phase; only two allocation policies will be tested per network, thus the first phase will act as a screening for the subsequent phases.

4.2.1 Static Disruption Responder Team Size

From section 4.1, it can be concluded that different networks have differing set of vulnerability points against cascading failures, as it was identified by their respective centrality measures. Based on the selection of conceptual networks that were tested, it can also be inferred that these discrepancies arise because of their inherent network properties (e.g. clustering coefficient, degree distribution, path length, etc) that each of these networks have – each conceptual network models a unique property that real-world networks have (Albert & Barabasi, 2002). Based on these findings, two hypothesis regarding DLOC-CDR performance are put forward to be verified by the TIE-DLOC simulator.

Hypothesis 4.3:

The DBA policy will yield an improvement by decreasing maximum cascading failure F_{max} and increasing preventability P_{rev}, but with a trade-off of increase in total response time z or latency and total distance travelled responders Δ compared to CBA. Let $z(\alpha)$, $P_{rev}(\alpha)$ and $\Delta(\alpha)$ denote the total latency, preventability, and total distance travelled performance metrics result of TIE-DLOC simulation using α resource allocation policy. This hypothesis can be expressed as:

$$z(DBA) - z(CBA) > 0 \tag{4.12}$$

$$|F_{(DBA)}| - |F_{(CBA)}| < 0 \tag{4.13}$$

$$P_{rev}(DBA) - P_{rev}(CBA) < 0 \tag{4.14}$$

$$\Delta(DBA) - \Delta(CBA) > 0 \tag{4.15}$$

Theoretically, the DBA policy will allocate depot to high degree nodes and based on the vulnerability analysis from section 4.1 all of the tested conceptual networks generally have their highest degree node as the vulnerable point of the network. Thus, by protecting these nodes it is presumed that the cascading failure can be minimized throughout the simulation replication if failures are initiated or propagates to these nodes. However, the trade-off is the agents (depot) are not initially located within the track of shortest path to most of the element within the network; as opposed to the CBA.

Hypothesis 4.4:

The BBA policy in the WS small world network will yield improvements by decreasing maximum cascading failure F_{max} and increasing preventability P_{rev} , but with a considerable trade-off of increase in total response time z or latency and total distance travelled responders Δ compared to CBA. This hypothesis can be expressed as:

$$z(BBA) - z(CBA) > 0 \tag{4.16}$$

$$z(BBA) - z(DBA) < 0 \tag{4.17}$$

$$\left|F_{(BBA)}\right| - \left|F_{(CBA)}\right| < 0 \tag{4.18}$$

$$P_{rev}(BBA) - P_{rev}(CBA) < 0 \tag{4.19}$$

$$\Delta(BBA) - \Delta(CBA) > 0 \tag{4.20}$$

$$\Delta(BBA) - \Delta(DBA) < 0 \tag{4.21}$$

As it was shown in section 4.1, the bridge centrality (BBA) was able to identify vulnerability points more effectively compared to betweenness centrality (CBA) in WS small world network. Thus, we concede that this should apply as well in the DLOC model. On the other hand, the BBA by construction also takes into consideration the betweenness centrality. Therefore, it should have better performance in terms of latency and total travelling distance compared to DBA.

The simulation parameters for this experiment parameters are summarized in Table 4.1. In order for the results to be comparable, all three conceptual CPI networks have the same number of nodes $|N|$ and average degree $\overline{deg_G}$. The disruption responder team consist of 35 agents where each of them are able to collaborate with 4 other agents (degree of each agent). It is assumed that failures spread at the speed of 1 edge/hour, while the agent's travel speed v is 1 edge/hour and the repair time t_{repair} is 1 hour. For each conceptual network, there will be three scenarios based on the allocation policies (CBA, BBA, DBA) and each scenario runs for 40 hours with 100 replications. The cascading is initiated by selecting 5 random nodes to be switched to failure state "1".

Table 4.1 Parameters for TIE-DLOC Simulation on Fixed Number of $|A|$

Parameter	Description	Value
$\lvert A\rvert$	Number of disruption responder agents	35
$deg_S(A)$	Degree for each agent	4
v	Agent's traveling speed	1 edge/hour
$\lvert N\rvert$	Number of nodes in Client Network	500
$\overline{deg_G}$	Average degree of Client Network	4
t_{repair}	Disruption repair timespan	1 hour
φ	Cascade threshold	0.25
u	Spread speed of propagating failures	1 edge/hour
$\lvert F_0\rvert$	Number of initial failures	5
n_R	Number of replication for simulation scenario	100
t_{max}	Simulation length for each scenario	40 hours

Table 4.2 and Table 4.3 summarizes the measured performance metrics result of the experiment while Figure 4.4 shows the disruption response behavior during the initial 40 hours. It can be seen that both hypothesis (4.3 & 4.4) have failed to explain the DLOC-CDR simulation results. In terms of total failures $|F|$, the DBA policy outperforms CBA policy on all tested networks: (1) ER decreases by -10.05%, (2) BA decreases by -3.61%, and (3) WS decreases by -14.71%, with an 80% statistical significance (based on two-sample t-test). Increment percentage is calculated by subtracting CBA result from DBA and divided by CBA result. This method will be used herein to denote performance metrics increment percentage. These results also appear to be in correlation with the vulnerability study in section 4.1. In all of the tested networks, the Degree centrality is most effective in identifying vulnerable elements which translates to effectiveness of DBA in reducing total failure. The performance gap between these two policies also follows a similar trend from the vulnerable study results, e.g. BA performance increment is marginal (<10%) where the vulnerability difference between Degree and Betweenness centrality method is also marginal (see Figure 4.2.a).

DBA policy also outperforms CBA in terms of total latency z : (1) ER decreases by -14.58%, (2) BA decreases by -5.41%, and (3) WS decreases by -18.02% with an 80% statistical significance (based on two-sample t-test). Again, the performance increment among networks also follows the same pattern from result of the vulnerability study (section 4.1). This result negates Hypothesis 4.3 by showing that total latency z is also improved along with the total failure $|F|$. This also implies that there is a positive correlation between these two metrics. Contrary to what has been researched earlier (Zhong & Nof, 2015;

Zhong, 2016), this result also suggests the objective function of DLOC-CDR can be adjusted to include total failure consideration.

The BBA policy performs the worst among the allocation policies in all tested networks. Despite its breakthrough method in identifying bridging nodes and edges, we have to concede that the it may not be suitable due to the performance metrics and failure model (Watts threshold cascading model) used by the DLOC. It has been validated that the Bridge Centrality is able to identify bridging elements which its removal will make the network disconnected (Hwang et al., 2006). In real worlds CPI network this is a catastrophic failure as it may halt flows or reachability to certain parts of the networks. In the DLOC model, however, the performance metric only measures the total failure as indicated by the state of elements (0 or 1), without regards to connectivity. This also shows the drawback of the DLOC model in its inability to detect the connectivity state of a network and maintain connectivity in the event of cascading disruption. This result also negates the validity of Hypothesis 4.4 with a 60% statistical significance.

In general, we can conclude from the experiment results that the DBA outperforms the CBA policy in all performance metrics on all of tested networks with an overall 80% statistical significance (see Table 4.4 and 4.5). The extent of when this conclusion is valid will be studied on the next section. Furthermore, we also conjecture two things.

Conjecture 4.1:

An a-priori vulnerability study on a network can foretell the performance of CBA and DBA allocation policy under the DLOC-CDR model.

Conjecture 4.2:

The effectiveness DBA (relative to CBA) to real-world CPI networks can be approximated by comparing the similarity of network structural features to that of the conceptual networks. For example, if a CPI network has an average geodesic length and clustering coefficient almost the same as an WS small world network with the same size (node and average degree) then effectiveness of each allocation policy (DBA and CBA) can be approximated by the WS small world network.

Table 4.2 Experiment Results for Static Team Size – Total Latency (z)

Network	**CBA**		**BBA**		**DBA**	
	Mean	STD.	Mean	STD.	Mean	STD.
ER	374.91	376.99	-	-	320.26	258.91
BA	253.48	182.69	837.54	1218.6	239.76	189.87
WS	1891.967	1729.29	2047.91	2082.1	1550.98	1443.21

Table 4.3 Experiment Results for Static Team Size – Total Failure ($|F|$)

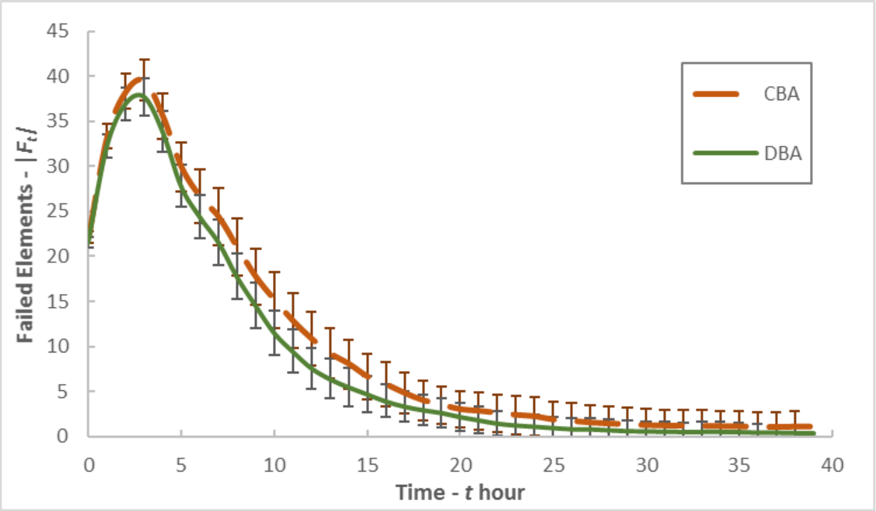
Network	**CBA**		**BBA**		**DBA**	
	Mean	STD.	Mean	STD.	Mean	STD.
ER	65.76	37.34	-	-	59.15	31.01
BA	54.35	27.65	100.55	85.54	52.39	28.25
WS	194.3	127.81	201.47	154.59	165.71	104.6

Table 4.4 Statistical Significance of Total Latency (z) Increment on two-sample t-test for Static Team Size

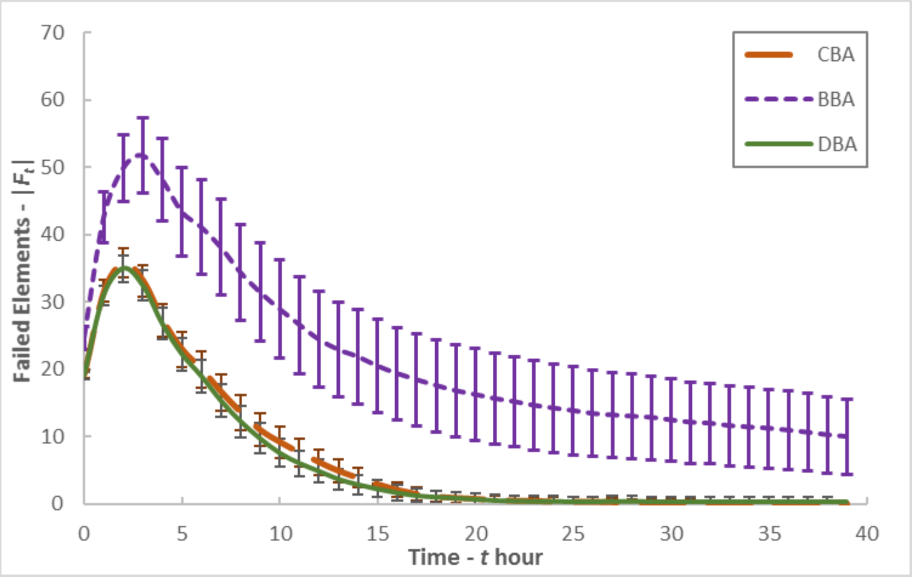
Network	**Increment from CBA to DBA**			**Increment CBA to BBA**		
	Mean	SEM	p-value	Mean	SEM	p-value
ER	-54.65	33.61	0.106	-	-	-
BA	-13.72	10.82	0.207	584.06	125.07	0.000
WS	-340.99	154.40	0.028	155.94	102.74	0.131

Table 4.5 Statistical Significance of Total Failure $(|F|)$ Increment on two-sample t-test for Static Team Size

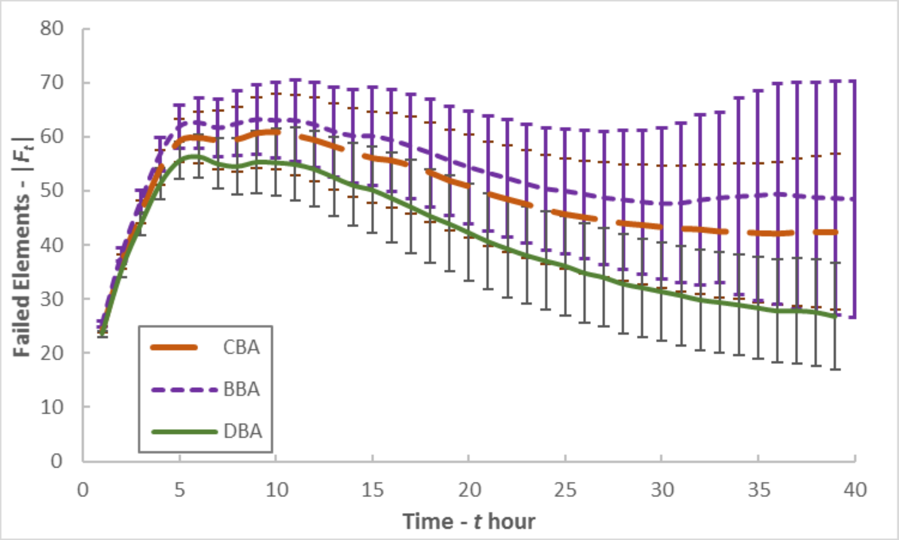
Network	**Increment CBA to DBA**			**Increment CBA to BBA**		
	Mean	SEM	p-value	Mean	SEM	p-value
ER	-6.60	3.12	0.036	-	-	-
BA	-1.96	1.49	0.191	46.2	9.111	0.000
WS	-28.59	10.86	0.009	7.17	8.70	0.411



(a) ER Random Network



(b) BA Scale-Free



(c) WS Small World

Figure 4.4 Dynamic Cascade Failure size $|F_t|$as a function of time t on different networks (with 0.95 confidence interval)

4.2.2 Varying Disruption Responder Team Size

Based on the results of the previous experiment, we will only be testing the CBA and DBA policy from herein. Only two conceptual networks will be used in this experiment, BA and WS, and with varying disruption responder team size $|A|$. The responder collaboration capability ($deg_s(A)$) is fixed at 4 because it has been verified in the previous research works (Zhong & Nof, 2015; Zhong, 2016) that changes in $deg_s(A)$ will have positive correlation on performance of resource allocation policy. The simulation parameter for this experiment is summarized in Table 4.6. A critical range of $|A|$ in BA and WS have also been observed in previous research works (Zhong & Nof, 2015; Zhong, 2016). Beyond this critical range, the responder team size is either too scarce or too abundant, such that different allocation policies have minor influence to the performance metrics. The critical range for BA and WS are $30 \leq |A| \leq 42$ and $56 \leq |A| \leq 61$, respectively.

Table 4.6 Parameters for TIE-DLOC Simulation on Varying Number of $|A|$

Parameter	Description	Value
$\|A\|$	Number of disruption responder agents:	
	- BA scale-free	$30 \leq \|A\| \leq 42$
	- WS small world	$56 \leq \|A\| \leq 61$
$deg_S(A)$	Degree for each agent	4
v	Agent's traveling speed	1 edge/hour
$\|N\|$	Number of nodes in Client Network	500
$\overline{deg_G}$	Average degree of Client Network	4
t_{repair}	Disruption repair timespan	1 hour
φ	Cascade threshold	0.25
u	Spread speed of propagating failures	1 edge/hour
$\|F_0\|$	Number of initial failures	5
n_R	Number of replication for simulation scenario	100
t_{max}	Simulation length for each scenario	40 hours

Table 4.7 and 4.8 summarizes the average performance metric increment percentage on BA and WS, respectively. Figure 4.5 and 4.6 shows the value of each performance metrics as function of $|A|$ on BA and WS, respectively, to illustrate the variation of performance increment.

As it can be seen, the performance metric increments on both networks are constant around the mean with low standard deviation. This result further adds to the conclusion from the previous section: For a fixed network size $|N|$ and initiator failure size ($|F_0|$), the DBA outperforms CBA with a constant performance increment around the mean with varying disruption responder team size. This result is also shown to be more statistically significant on WS network where null hypothesis is rejected at significance level 0.1, compared to BA network at significance level 0.2.

Table 4.7 Mean Performance Increment on BA with Varying $|A|$

Metric	Mean increment from CBA to DBA (%)*	Standard deviation of increment	p-value of one-sample t-test**
Total failure ($|F|$)	-3.75%	0.035	0.028
Total distance traveled by agents (Δ)	-3.46%	0.029	0.006
Total latency (z)	-6.30%	0.051	0.207
Preventability (P_{rev})	0.26%	0.006	0.175

*Calculated by subtracting CBA result from DBA and divided by CB

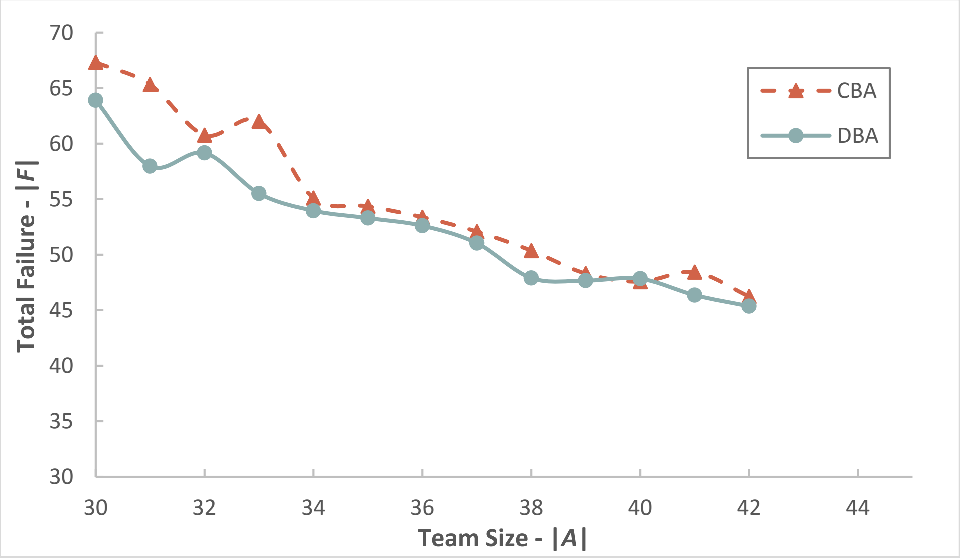
**Null hypothesis of no increment

Table 4.8 Mean Performance Increment on WS with Varying $|A|$

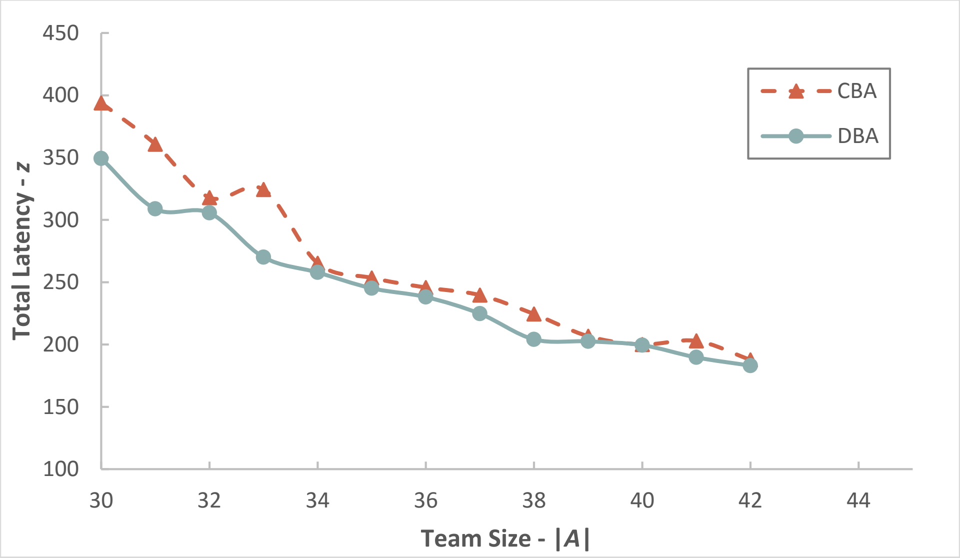
Metric	Mean increment from CBA to DBA (%)*	Standard deviation of increment	p-value of one-sample t-test**
Total failure ($|F|$)	-14.97%	0.040	0.000
Total distance traveled by agents (Δ)	-9.56%	0.034	0.09
Total latency (z)	-19.63%	0.050	0.000
Preventability (P_{rev})	0.12%	0.007	0.05

*Calculated by subtracting CBA result from DBA and divided by CB

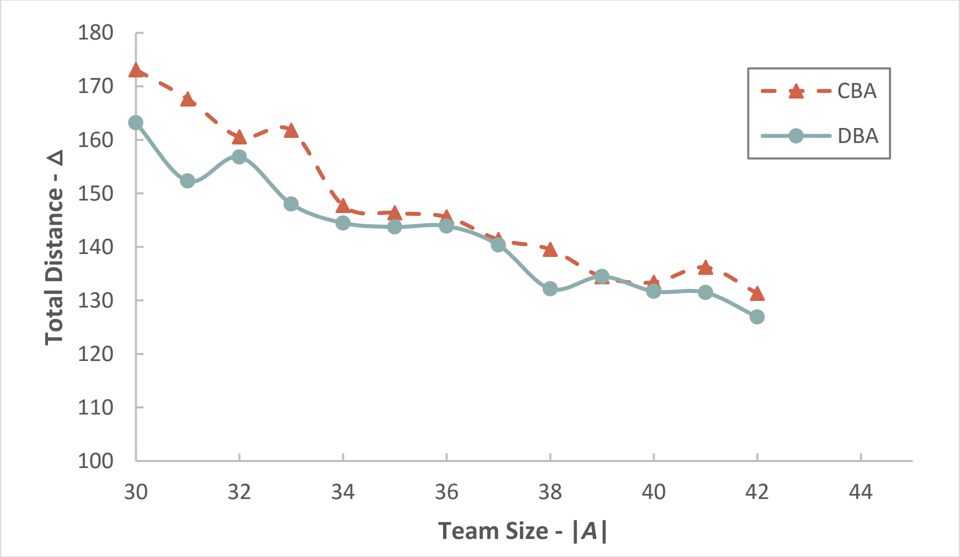
**Null hypothesis of no increment



(a) Total Failure ($|F|$)

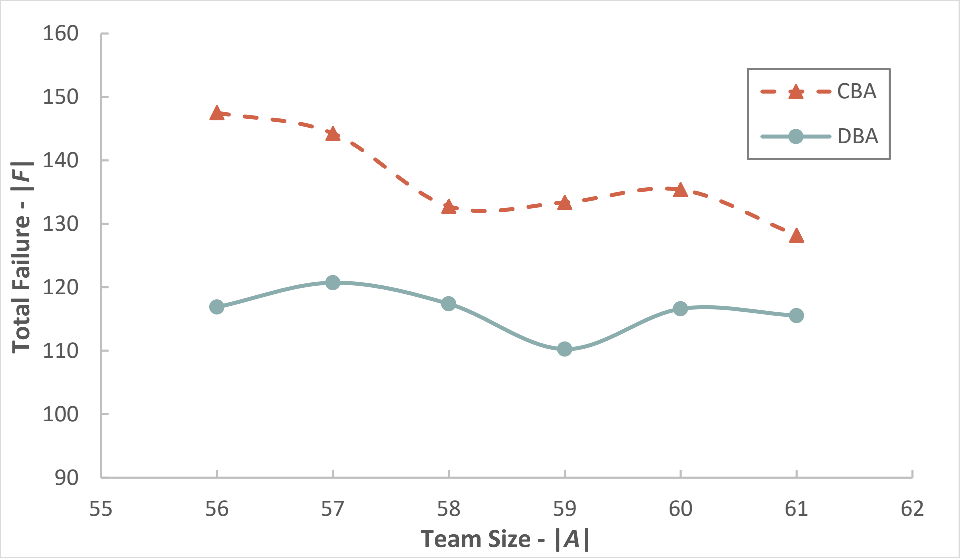


(b) Total Latency (z)

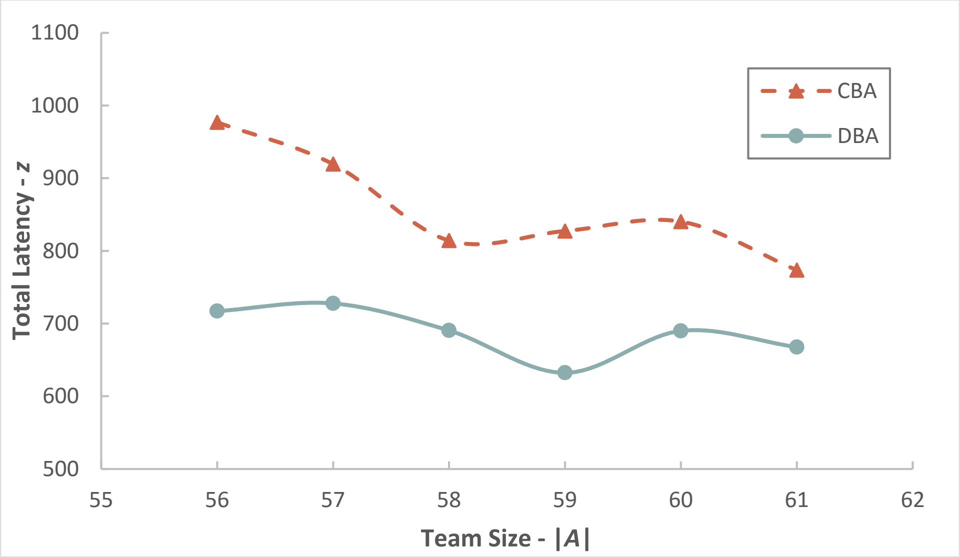


(c) Total Distance Travelled (Δ)

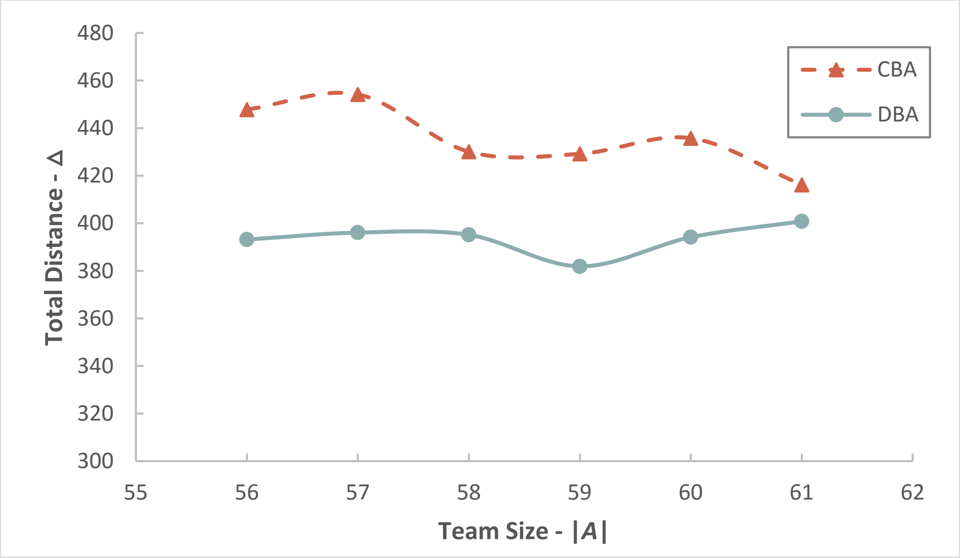
Figure 4.5 Experiment Results on BA with Varying $|A|$



(a) Total Failure ($|F|$)



(b) Total Latency (z)



(c) Total Distance Travelled (Δ)

Figure 4.6 Experiment Results on WS with Varying $|A|$

4.2.3 Varying Initiator Failure Size

The two previous experiments have validated the performance of the DBA policy relative to CBA on the tested conceptual networks from the perspective of the server network (i.e. responder team size). For this set of experiment, the performance difference between DBA and CBA is further investigated by varying the size of initial failure $|F_0|$. The simulation parameters for this experiment are listed in Table 4.9. The conceptual networks tested (BA and WS) will have fixed $|N|$, $|A|$, $\overline{deg_G,}$ & $deg_s(A)$ at 1000, 50, 4 and 4, respectively. Five variation of initial failure size $|F_0|$ will be tested: 5, 10, 20, 30, 40, and 50, which corresponds to 0.005, 0.01, 0.02, 0.03, 0.04 and 0.0.05 fraction of the total nodes $|N|$, respectively.

Table 4.9 Parameters for TIE-DLOC Simulation on Varying Number of $|F_0|$

Parameter	Description	Value
$\|A\|$	Number of disruption responder agents	50
$deg_S(A)$	Degree for each agent	4
v	Agent's traveling speed	1 edge/hour
$\|N\|$	Number of nodes in Client Network	1000
$\overline{deg_G}$	Average degree of Client Network	4
t_{repair}	Disruption repair timespan	1 hour
φ	Cascade threshold	0.25
u	Spread speed of propagating failures	1 edge/hour
$\|F_0\|$	Number of initial failures	5; 10; 20; 30; 40; 50
n_R	Number of replication for simulation scenario	100
t_{max}	Simulation length for each scenario	40 hours

Figure 4.7 and 4.8 shows the experiment results for total failure ($|F|$) and total latency (z) metrics increment (%) on WS and BA, respectively. The experiment result reveals that within a fixed client and responder network size, the performance of both network policy, relative to each other, varies greatly with varying size of initial failures. The DBA performs better within intermediate range of $|F_0|$, while the CBA triumphs in the low range ($|F_0| < 25$) and high range ($|F_0| > 40$) of $|F_0|$ in WS small world network. Both metrics, $|F|$ and

z, observes a similar pattern. On other hand, the DBA constantly outperforms CBA within the tested interval in BA scale-free network, albeit with a declining trend. This would imply that DBA performs better relative to DBA in low range of initial failure size within a fixed network size.

Based on this experiment, it can be concluded that the performance of CBA and DBA varies between different network structure and initial failure size. Among different network structure and within a fix range of failure size and number nodes $|N|$, the allocation policy behaves differently, e.g. in BA the DBA is constantly better than CBA while it is not true in WS (Figure 4.7 and 4.8). On the other hand, within the same network and fixed responder team size, the relation between CBA and DBA, in terms of performance, is not strict; one can outperform the other depending on the initial failure size. This is further supported by the statistical properties gathered from the TIE-DLOC simulator presented in Table 4.10 and Table 4.11. On average, the DBA policy shows improvements on all performance metrics compared to CBA on both networks. However, these results are not statistically significant as indicated by the *p*-values of one-sample *t*-test on null hypothesis of zero increment. The BA network had a maximum *p*-value of 0.382, while the WS network had 0.865.

Table 4.10 Mean Performance Increment on BA with Varying $|F_0|$*

Metric	Mean increment from CBA to DBA	Standard deviation of increment	p-value of one-sample t-test
Total failure ($|F|$)	-13.06	14.937	0.382
Total distance traveled by agents (Δ)	-0.01	4.377	0.998
Total latency (z)	-364.44	332.778	0.274
Preventability (P_{rev})	0.001	0.007	0.874

* Result on $|F_0|$ = 40 omitted for consistency

**Null hypothesis of no increment

Table 4.11 Mean Performance Increment on WS with Varying $|F_0|$

Metric	Mean increment from CBA to DBA	Standard deviation of increment	p-value of one-sample t-test*
Total failure ($|F|$)	-10.81	63.962	0.865
Total distance traveled by agents (Δ)	-9.56%	0.034	0.09
Total latency (z)	-274.45	1393.282	0.844
Preventability (P_{rev})	0.000	0.000	0.096

**Null hypothesis of no increment

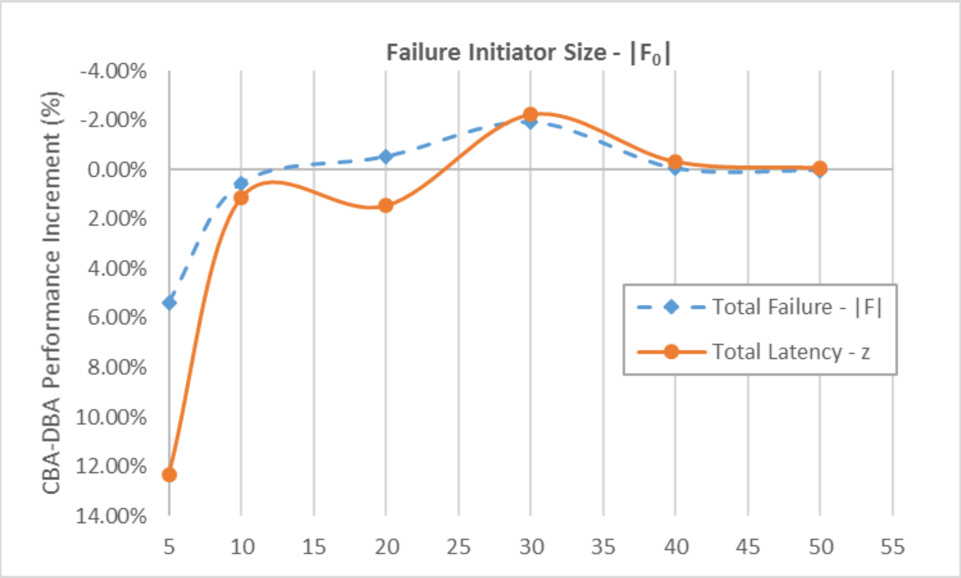


Figure 4.7 Experiment Results on WS with Varying $|F_0|$

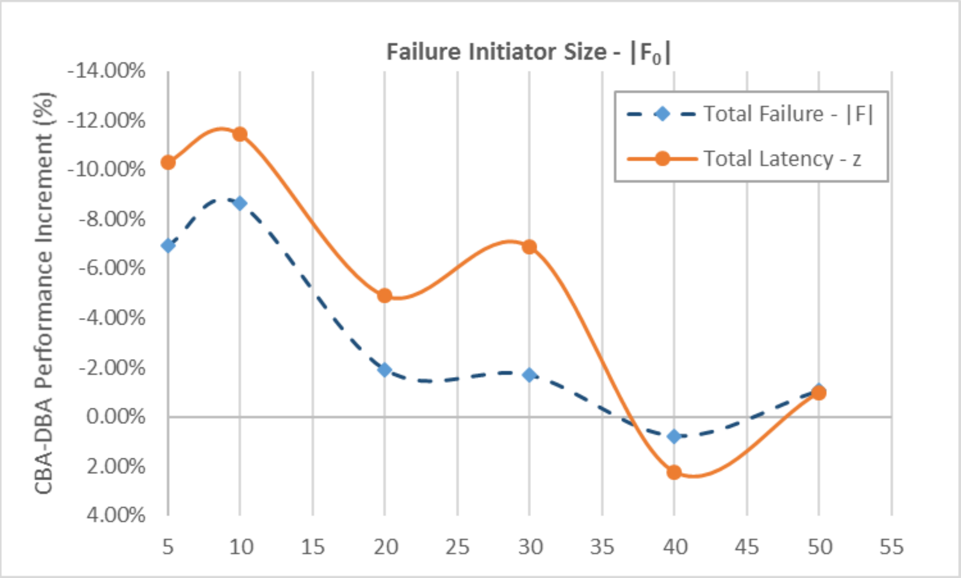


Figure 4.8 Experiment Results on BA with Varying $|F_0|$

4.3 DLOC Model Simulation: USA Western Power Grid

In this section, a real-world power grid system (PG) and its DLOC-CDR operation will be simulated in TIE-DLOC. The objective of this final experiment is to test the validity of the insights and conclusions gained from previous experiments on conceptual networks to real-world networks. The client network will simulate a network based on the USA Western Power Grid. The nodes of the client represent generators, transformers, and substations and the edges represent transmission line (Watts & Strogatz, 1998). There are 4941 nodes in this network connected by 6594 edges. The average degree of this network is 2.67.

Past research works have analyzed the structural properties of the PG network (Watts & Strogatz, 1998). PG has an average path length almost similar to that of ER with the same size: 18.7 ~ 12.4 (PG and ER, respectively), but significantly higher clustering coefficient: 0.080 ~ 0.005 (PG and ER, respectively). The fact that the PG network maintains low average path and high clustering coefficient entitles it to having the "small-world network behavior"; most of the nodes can by another with small number of steps. Past research results in DLOC model using the CBA allocation policy concluded that this property makes the disruption responder less effective in recovering the client network as measured by P_{rev} Zhong & Nof, 2015; Zhong, 2016. The experiment in section 4.2.1 used the same datasets and setup as (Zhong & Nof, 2015; Zhong, 2016) and has shown that improvement in all of the performance metrics can be made by using the DBA allocation policy in CPI with small-world networks (WS). Nevertheless, the PG does not perfectly imitate all the properties that WS have; WS has a clustering coefficient of ~0.18 and a shorter right tail distribution (see Figure 4.9). Given the difference in clustering coefficient and degree distribution, it seems unreasonable to conclude the results of DLOC from WS model can fully predict the PG network.

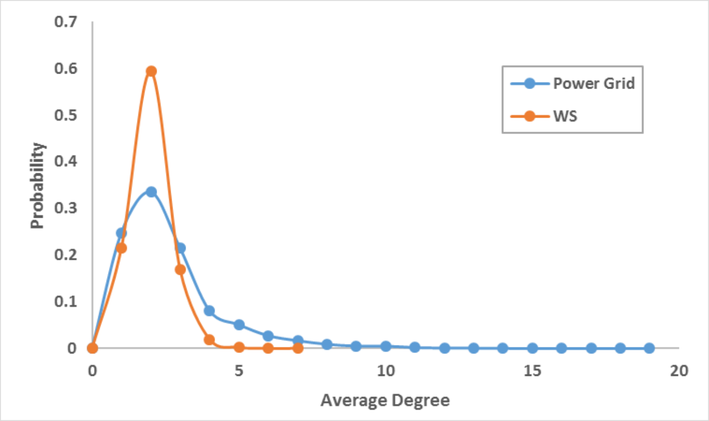


Figure 4.9 Degree Distribution of PG and WS

The simulation parameter for this experiment is shown in Table 4.12. In this experiment, the disruption responder team has 500 agents and each of them are able to collaborate with 300 others. The cascade threshold is fixed at 0.25 and failures propagates at 1 edge/hour. All four performance metrics used in section 4.2 will also be measured in this experiment. Based on the insights of the varying initial failure size experiment (section 4.2.3), we will use two initial failure size $|F_0|$, 5 and 49, to analyze the validity of CBA-DBA performance change. Each simulation scenario will run for 40 hours with 100 replications.

Table 4.12 Parameters for TIE-DLOC Simulation on USA Western Power Grid (PG)

Parameter	Description	Value
$\vert A\vert$	Number of disruption responder agents	500
$deg_S(A)$	Degree for each agent	300
v	Agent's traveling speed	1 edge/hour
$\vert N\vert$	Number of nodes in Client Network	4943
$\overline{deg_G}$	Average degree of Client Network	2.67
t_{repair}	Disruption repair timespan	1 hour
φ	Cascade threshold	0.25
u	Spread speed of propagating failures	1 edge/hour
$\vert F_0\vert$	Number of initial failures	5; 49
n_R	Number of replication for simulation scenario	100
t_{max}	Simulation length for each scenario	40 hours

Table 4.13 summarizes the performance metrics increment result from this experiment, while Table 4.14 and Table 4.15 shows the statistical significance of the increments for $|F_0| = 5$ and $|F_0| = 49$, respectively. Figure 4.10 shows the process of cascading failures with disruption response control after initial failures on PG.

The performance metrics increments show that the CBA policy performs better than DBA in the PG network (for the current simulation setting). For $|F_0|$ = 5, the CBA has significantly lower total latency z at 2170.24 as opposed DBA at 2595.39, which represents 19.59% increment reduction. CBA also performs better relative to DBA on the other metrics with $\approx 10\%$ increment. The disruption response over time (Figure 4.10.a) shows that the DBA actually performs better (in terms of $|F_t|$) than CBA during the initial hours of the cascading failures ($t < 9$), though it was later outperformed by CBA. This pattern can be explained by the following argument. By the vulnerability analysis on PG (not shown), the Degree centrality is in fact more effective in identifying vulnerable elements resulting in higher cascading failure size, S. This is supported by the degree distribution of PG which shows the top 5 highest degree node can reach 2% of the nodes in PG network, compared to WS 0.5% given the same size, which means these nodes are more significant in terms of vulnerability – albeit not as significant as in the BA with the same size reaching 8% of the nodes. However, this still doesn't lead to foretell that the DBA should perform better, as it has already been disapproved by the experiment result. Another driving factor is the small-world property that the PG maintains due to its low average path and high clustering coefficient. It has been studied that the small-world property enables diseases to spread much more easily and faster due (Watts & Strogatz, 1998). This insight is in the same analogy of the Watts threshold model used in DLOC and enables failures to propagate faster throughout the network, relative to the responder recovery speed. Finally, these pieces of facts together construct the argument that the DBA performs better during the initial hour because depots are better positioned to recover the vulnerable nodes if they are selected as initial failure throughout the simulation replication. The significance of DBA during this initial hour is also supported the PG's degree distribution as aforementioned before. During the latter hours, the small-world property enables the initial failures to propagate faster throughout the PG network, which are widely spread within 4941 nodes. In this case, the CBA policy has better positioning to reach to distant nodes since the responders are located between intersection of shortest paths. This explains the turn-around point observe at $t = 9$ (Figure 4.10.a) for CBA, while the DBA needs more time to reach the failing elements – WS-like networks have low correlation between degree and betweenness centrality (Holme *et al.,* 2002). The validity of these argument also hinges upon the assumption of propagation type we are modelling; in this case it is simple propagation as noted by the low cascade threshold φ.

Increasing the initial failure size $|F_0|$ further reduces the performance gap between CBA and DBA. As it can be seen, all performance metrics have increments <10%. This pattern confirms the study that was conducted in section 4.2.3 where the number of initial failures affects the performance of both policies relative to each other. The DBA also had much longer range in which before it was outperformed by CBA at $t = 11$ (Figure 4.10.b). Nevertheless, it can also be seen in Figure 4.10.b that both policies have marginal effect towards disruption response operation because the inadequate number of responder resource ($|A|$) as signified by the increasing trend of $|F_t|$ towards the end of the initial 40 hours.

In conclusion, TIE-DLOC simulation of USA Western Power Grid Network (PG) for disruption response operation has helped validate the conclusion gained from experiments in previous sections and bridge the gap between studies on conceptual networks and real-world CPI networks. Conceptual networks (ER, BA, and WS) models distinct properties of networks that are commonly found collectively in real-world networks. In this case, the PG has small-world property which is modeled by the WS network, but a degree distribution akin to the BA, though not purely power law distribution. As result, the response behavior mimics an intermediary between the two models. The DBA performs better initially due to power law-like degree, but later outperformed by CBA due to the fast propagation in small-world and less correlation between degree and betweenness centrality. By comparing the network properties of CPI networks with conceptual networks model, decision maker and designer can better predict the disruption response operation of the CPI network under different resource policies and failures. It was also shown that in real-world CPI network, the CBA and DBA can be used interchangeably depending on the network structure – each policies do not constantly outperform the other.

Table 4.13 Performance Increment of CBA-DBA on PG with varying $|F_0|$

Metric	Performance increment (%)* for $|F_0| = 5$	Performance increment (%)* for $|F_0| = 49$
Total failure ($|F|$)	9.62%	0.65%
Total distance traveled by agents (Δ)	9.55%	5.02%
Total latency (z)	19.59%	5.29%

*Calculated by subtracting CBA result from DBA and divided by CBA

Table 4.14 Statistical Significance of Performance Metric Increment on one-sample t-test for $|F_0| = 5$

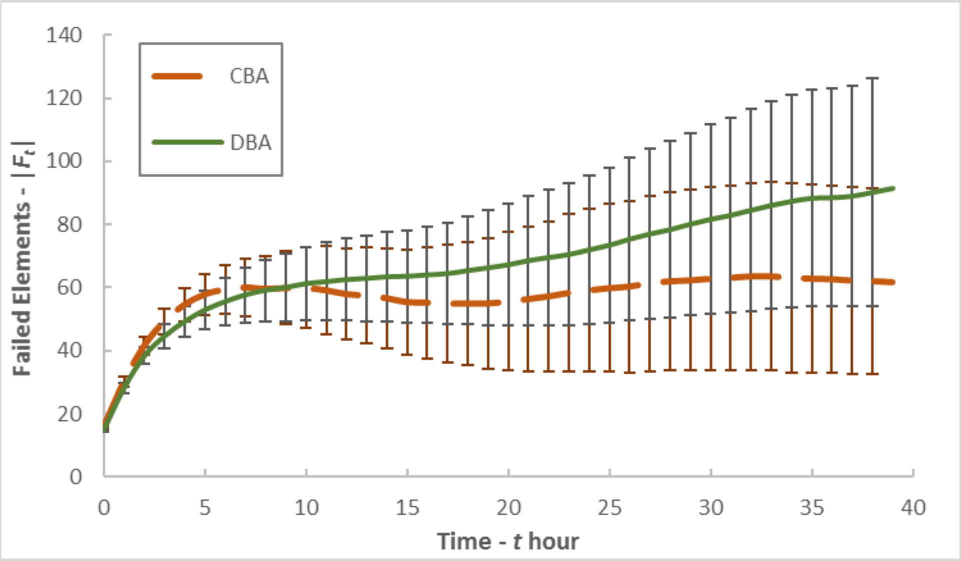
Metric	Mean increment from CBA to DBA	Standard deviation of increment	p-value of one-sample t-test*
Total failure ($|F|$)	28.28	33.442	0.399
Total distance traveled by agents (Δ)	90.74	105.231	0.391
Total latency (z)	425.15	327.720	0.197

*Null hypothesis of no increment

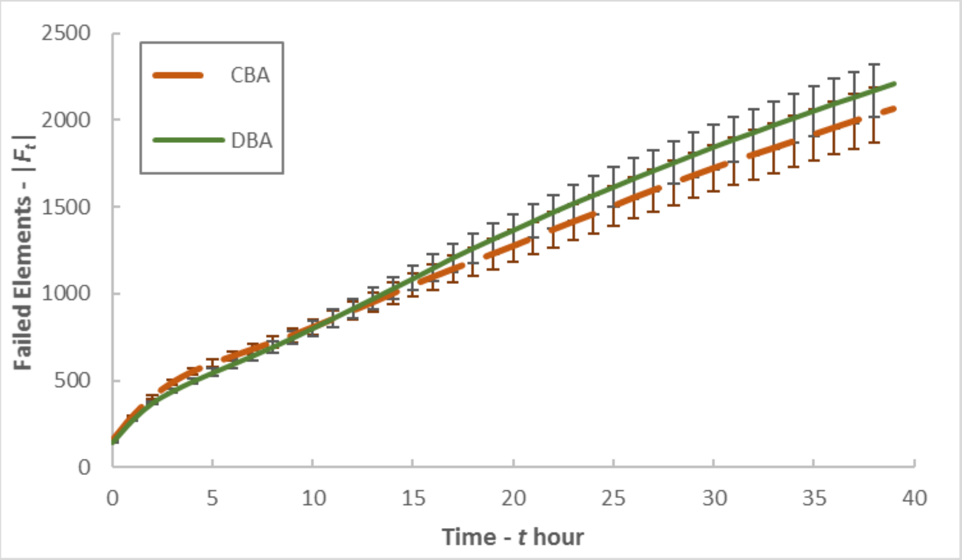
Table 4.15 Statistical Significance of Performance Metric Increment on one-sample *t*-test for $|F_0| = 49$

Metric	Mean increment from CBA to DBA	Standard deviation of increment	p-value of one-sample *t*-test*
Total failure ($|F|$)	29.62	98.413	0.764
Total distance traveled by agents (Δ)	485.77	119.759	0.000
Total latency (z)	2540.08	1366.21	0.065

*Null hypothesis of no increment



(a) $|F_0|$=5



(b) $|F_0|$=49

Figure 4.10 Dynamic Cascade Failure size $|F_t|$as a function of time t on PG (with 0.95 confidence interval)

4.4 Summary

Incorporating vulnerability consideration to the of DLOC-CDR has yielded many useful insights for future work in designing resource allocation policy for disruption responder teams in CPI networks. From a network vulnerability point of view, the degree centrality appears to be most effective heuristic to point out vulnerable elements of a network under cascading failure. We initially conject that the results of vulnerability analysis to be a good indicator of the performance of the respective CBA and DBA policy (Conjecture 4.1). However, Conjecture 4.1 is rejected based on the experiments results of 4.2.3. Thus, network designer should not solely use network vulnerability indicator to select which resource allocation policy to use.

Two new heuristic resource allocation policy heuristics were implemented in this research, namely the Degree-based Allocation (DBA) and Bridge-based Allocation (BBA) based on degree centrality and bridge centrality, respectively. It was initially hypothesized that the BBA would perform better than CBA because of its ability to detect and protect bridging node which can increase the number of connected components in networks. However, the simulation conducted proved otherwise. This is because the TIE-DLOC simulator does not measures the connectivity state of the client CPI network. In reality, some CPI network weights connectivity to be more important as opposed to number of failure itself, e.g. water distribution network. This can be a future direction for the development of the TIE-DLOC simulator to also measure network connectivity state.

Finally, it is found that two competing resource allocation policy heuristics, DBA and CBA can be used interchangeably depending on the network properties a CPI network has. From our experiment on conceptual networks (ER, BA and WS), we made a conjecture (Conjecture 4.2) that by comparing the similarity of the network graph structural features (average geodesic length, degree distribution, clustering coefficient) to that of the corresponding conceptual network, we can make an inference on how well the CBA-DBA policy will perform with respect to each other – given the same network size. Albeit a general guideline is left for future work, some constant patterns have been found:

1. Varying disruption responder team size, while fixing agent's degree, will yield an expected increment value which approaches a mean with marginal STD. Varying fixed degree will improve performance by increasing collaboration capability (Zhong & Nof, 2015; Zhong, 2016).

2. The performance increment from CBA to DBA, within a given network size, is sensitive to the size of initial failure $|F_0|$. Thus, network design must assess what is the expected number failures that may appear simultaneously to trigger a cascading failure.
3. The DBA policy constantly performs better compared to CBA on networks with power-law degree distribution or similar to BA Scale-free. This assertion is supported by experiments from section 4.2.1 and 4.2.3 and theoretically due to the fact the DBA and CBA have high correlation in terms of their centrality (Holmes, 2002).
4. In networks with small world property, the performance increment from CBA to DBA has a decreasing trend with increasing network size. For large network size, the CBA will outperform DBA. This assertion is supported by the experiment results in section 4.2.1 and 4.2.2 where in the latter the WS client network has more nodes, $|N|$, compared to the former and the DBA policy performs worse than CBA. The same behavior is also prevalent in the PG network. Theoretically, this can also be explained by the fact that: (1) the degree centrality and betweenness centrality have lower correlation in networks with increasing clustering coefficient (Holmes, 2002) and (2) Small-world properties allows failures to propagates faster throughout the network. Due to fact (1), responder agents under CBA policy are better positioned to travel quickly throughout the network to recover propagating failures.

CHAPTER 5. CONCLUSION AND FUTURE WORK

The findings on this research can be concluded according to the RQ as following

1. RQ1: Graph Theoretical Analysis of DLOC

   The DLOC models measures resiliency of networks under cascading disruption by the Recoverability metric $P_{recover}$ which denotes the probability of a network to fully recover from a cascading disruption. The behavior of cascade failure in DLOC is driven by the underlying model of Watts Threshold Cascade. Using percolation theory in graphs, past research works have been able to create an analytical model to predict average cascade size in different networks (Watts, 2002). Motivated by this, we hypothesize there exist a regime where $P_{recover} = 1$ for all different networks and is driven by its degree distribution and threshold cascade φ. We found that this is true. More specifically, we found that if the fraction of vulnerable nodes of the network, $P_{vulnerability}$, is below 0.7, $P_{recover} = 1$. This hypothesis was tested on three different conceptual networks with varying size to represent the changes in topology and one real-world CPI network, the USA Western Power Grid (PG). All experiment complies with the aforementioned hypothesis. Furthermore, the experiment result complements findings from past DLOC research works in terms of the "small-world" effect of on disruption response.

Finally, this result can also be used a quick guideline for designers of e-Work systems on evaluating the reliability of their disruption responder teams. Similar analysis can also be conducted with different responder team size to produce a more robust guideline.

2. RQ2: Protection of Vulnerable Parts of the Network.

   Based on vulnerability analysis, two new heuristic resource allocation policy heuristics were implemented in this research, namely the Degree-based Allocation (DBA) and Bridge-based Allocation (BBA) based on degree centrality and bridge centrality, respectively. However, BBA was later proven to be inferior in performance compared to the two others.

   The DBA and CBA can be used interchangeably depending on the network properties that a CPI network have. From our experiment on conceptual networks (ER, BA and WS), we made a conjecture (Conjecture 4.2) that by comparing the similarity of the network graph structural features (average geodesic length, degree distribution, clustering coefficient) to that of the corresponding conceptual network, we can make an inference on how well the CBA-DBA policy will perform with respect to each other – given the same network size. Albeit a general guideline is left for future work, some constant patterns have been found:

   a. Varying disruption responder team size, while fixing agent’s degree, will yield an expected increment value which approaches a mean with marginal STD.

   b. The performance increment of DBA and CBA, within a given network size, sensitive to the size of initial failure $|F_0|$. Thus, network design must assess what is the expected number failures that may appear simultaneously to trigger a cascading failure.

c. The DBA policy constantly performs better to CBA better on networks with power-law degree distribution or similar to BA Scale-free.

d. In networks with small world property, the performance increment from CBA to DBA has a decreasing trend with increasing network size.

Based on the results of this research, we have identified selected topics that have big potential for the development of DLOC model for future works:

1. Development analytical methods to approximate DLOC-CDR performance

   The results and analysis from RQ1 implies that future works can be directed to construct analytical methods in reviewing DLOC design, as opposed to the current empirical methods. Analytical methods will further improve DLOC's modelling capability and versatility in aiding network designers by providing quick guidelines and accurate approximation. Some references worthwhile reviewing is regarding percolation theory and universal behavior of cascade and contagion behavior.

2. Measure Connectivity State of Client Network

   Connectivity state is an important issue in complex infrastructure systems, such as transportation network, water distribution network, etc. While the DLOC has rigorously modeled failure states of network elements, it is still lacking the ability to measure connectivity state. The inclusion of connectivity state may alter the design guidelines produce by the current DLOC, as a new objective may be added to minimize disconnection in network. This argument can be implied by the effectiveness of the BBA in the current DLOC (Chapter 4), which is inferior to CBA and DBA despite the advantage of detecting bridges in network. Bridges are important element networks where the severing of these bridges can increase the number of connected components and disconnect a network (Bondy & Murty, 2008).

LIST OF REFERENCES

LIST OF REFERENCES

Albert, R., & Barabasi, A.-L. (2002). Statistical mechanics of complex networks. *Reviews of Modern Physics, 74*.

Albert, R., Jeong, H., & Barabasi, A.-L. (2000). Error and Attack Tolerance of Complex Networks. *Nature, 406*(6794), 378-382.

Aldecoa, R., & Martin, I. (2013). Surprise Maximization Reveals the Community Structure of Complex Networks. *Scientific Report 3*, 1060.

Aponte, E. E., & Nelson, J. K. (2006). Time Optimal Load Shedding for Distributed Power Systems. *IEEE Transactions on Power Systems, 21*(1), 269-277.

Arrowsmith, D., Bernardo, M. D., & Sorrentino, F. (2005). Effects of variations of load distribution on network. *IEEE International Symposium on Circuits and Systems.* Kobe: IEEE.

Barabasi, A.-L., & Albert, R. (1999). Emergence of Scalling in Random Networks. *Science, 286*(5439), 509-512.

Barthelemy, M. (2004). Betweenness Centrality in Large Complex Networks. *Eur. Phys. Jour. B, 38*(163).

Bevrani, H., Tikdari, A. G., & Hiyama, T. (2010). An Intelligent Based Power System Load Shedding Design Using Voltage and Frequency Information. *Modelling, Identification and Control (ICMIC), The 2010 International Conference on* (pp. 545-549). Okayama: IEEE.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast Unfolding of Communities in Large Networks. *PhysicsSoc*.

Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). Fast Unfolding of Communities in Large Networks. *J. Stat Mech.*

Bollobas, B. (1984). The evolution of random graphs – giant components. *Transactions of The American Mathematical Society, 286*(1), 257-274.

Bondy, A., & Murty, U. S. (2008). *Graph Theory.* London: Springer-Verlag.

Borgatti, S. P. (2005). Centrality and network flow. *Soc. New., 27*(1), 55-71.

Candranegara, G., Zhong, H., & Nof, S. Y. (2015). Conflict and error management based on collaborative control theory: A case study in the furniture industry. *Proc. the 23rd International Conference on Production Research.*

Centola, D. (2009). Failure in Complex Social Networks. *Journal of Mathematical Sociology, 33*, 64-68.

Centola, D., Eguiluz, V. M., & Macy, M. W. (2007). Cascade Dynamics of Complex Propagation. *Physica A*, 449-456.

Chen, J. (2002). *Modelling and Analysis of Coordination for Multienterprise Networks (Doctoral Dissertation).* School of Industrial Engineering, Purdue University. Retrieved from http://search.proquest.com/docview/305541775?accountid=13360

Chen, J., & Nof, S. Y. (2000). Multi-Enterprise Networking. *Proceedings of International Conference on Manufacturing Systems: Innovations for the 21st Century.* Ann Arbor, MI.

Chen, X. W., & Nof, S. Y. (2012). Agent-Based Error Prevention Algorithm. *Expert Systems with Applications, 39*(1), 280-287.

Chen, X. W., & Nof, S. Y. (2012). Conflict and Error Prevention and Detection in Complex Networks. *Automatica, 48*(5), 770-778.

Cohen, R., & Sholomo, H. (2003). Scale-Free Networks are Ultrasmall. *Phys. Rev. Lett, 90*.

Cohen, R., Erez, K., ben-Avraham, D., & Havlin, S. (2000). Resilience of the Internet to Random Breakdowns. *Physical Review Letters, 85*(21), 4626-4628.

Crucitti, P., Latora, V., & Marchiori, M. (2004). Model for Cascading Failures in Complex Networks. *69*(4).

Crucitti, P., Latora, V., & Marchori, M. (2004). Model for cascading failures in complex networks. *Physical Review E:Statistical, Nonlinear, and Soft Matter Physics, 69*(4).

Crucittit, P., Latora, V., Marchori, M., & Rapisarda, A. (2003). Efficiency of Scale-Free Networks: Error and Attack Tolerance. *Physica A, 320*.

Current, J. R., Re Velle, C. S., & Cohon, J. L. (1985). The maximum covering/shortest path problem: A multiobjective network design and routing formulation. *European Journal of Operational Research, 21*(2), 189-199.

Current, J., Min, H., & Schilling, D. (1990). Multiobjective analysis of facility location decisions. *European Journal of Operational Research, 49*(3), 295-307.

Dezo, Z., & Barabasi, A.-L. (2002). Halting Viruses in Scale-Free Networks. *Physical Review E, 65*.

Dodds, P. S., & Payne, J. L. (2009). Analysis of a threshold model of social contagion on degree-correlated networks. *Physical Review E*.

Dodds, P., & Watts, D. J. (2002). Universal Behavior in a Generalized Model of Contagion. *Physical Review Letters, 92*(21).

Dong, J., Chen, Q., & Niu, Z. (2007). Random graph theory based connectivity analysis in wireless sensor networks with rayleigh fading channels. *2007 Asia Pacific Conference on Communications* (p. 123). Bangkok: IEEE.

Dorogovtsev, S. N., & Mendes, J. F. (2002). Evolution of networks. *Advances in Physics, 51*(4), 1079.

Duan, W., Chen, Z., Liu, Z., & Jin, W. (2005). Efficient target strategies for contagion in scale-free networks. *Physical Review E, 72*.

Duan, W., Chen, Z., Liu, Z., & Jin, W. (2005). Efficient Target Strategies for Contagion in Scale-Free Networks. *Physical Review, 72*.

Dunn, S., & Wilkinson, S. M. (2016). Increasing the Resilience of Air Traffic Networks using a Network Graph Theory Approach. *Transportation Research Part E, 90*, 39-50.

Easley, D., & Kleinberg, J. (2010). *Cascading Behavior in Networks.* Cambridge University Press.

Erdős, P., & Rényi, A. (1959). On Random Graphs. I. *Publicationes Mathematicae, 6*, 290-297.

Erdős, P., & Rényi, A. (1960). On the evolution of random graphs. *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*.

Freeman, L. C. (1977). A set of measures of Centrality Based on Betweenness. *Sociometry, 40*(1), 35-41.

Freeman, L. C. (1977). A Set of Measures of Centrality Based on Betweenness. *Sociometry, 40*(1), 35-41.

Gentili, M., & Mirchandani, P. B. (2005). Locating Active Sensors on Traffic Networks. *Annals Operations Research, 136*(1), 229-257.

Gleeson, J. P. (2008). Cascades on correlated and modular random networks. *Physical Review E*.

Goh, K.-I., Kahng, D., & Kim, D. (2001). Universal Behavior of Load Distribution in Scale-free Networks. *Phys. Rev. Lett, 87*.

Goh, KI et al. (2002). Classification of scale-free networks. *PNAS, 99*(20).

Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology, 78*(6), 1360-1380.

Hackett, A., Melnik, S., & Gleeson, J. (2011). Cascades on a class of clustered random networks. *Physical Review E*.

Hagberg, A., Schult, D., & Swart, P. (2015). *NetworkX Reference.*

Hart, Michael G., et al. (2015). Graph Theory Analysis of Complex Brain Network: New Concepts in Brain Mapping Applied to Neurosurgery. *J Neurosurg*.

Hein, O., Schwind, M., & Konig, W. (2006). Scale-Free Networks. The Impact of Fat Tailed Degree Distribution on Diffusion and Communication Processes. *Wirtschaftsinformatik, 48*(4), 267-275.

Hines, P., Blumsack, S., Cotilla Sanchez, E., & Barrows, C. (2010). The Topological and Electrical Structure of Power Grids. *System Sciences (HICSS), 2010 43rd Hawaii International Conference* (pp. 1530-1605). Honolulu: IEEE.

Holme, P., Kim, B. J., Yoon, C. N., & Han, S. K. (2002). Attack vulnerability of complex networks. *Phys. Rev, 65*.

Hwang, W., Choe, Y.-R., Zhang, A., Ramanathan, & M. (2006). Bridging Centrality: Identifying Bridging Nodes In Scale-Free Networks. *KDD.*

Kadloor, S., & Santhi, N. (2010). *Understanding Cascading Failures in Power Grids.* Los Alamos National Laboratory.

Kawahigashi, H. (2005). Modeling ad hoc sensor networks using random graph theory. *Second IEEE Consumer Communications and Networking Conference, 2005*, (pp. 104-109).

Krisnamurthy, S., Chandrasekaran, R., Venkatesan, S., & Dawande, M. (2003). Highly Efficient Spare Capacity Planning for Generalized Link Restoration. *Proceedings of 12th International Conference on Computer Communications and Networks*, (pp. 47-52).

Li, B., Hu, X., & Xie, B. (2009). Transportation network reconstruction for natural disasters in the emergency phase based on connectivity reliability. *Proceedings of the 2nd International Conference on Transportation Engineering* (pp. 2963-2968). ASCE.

Li, S., Li, L., Jia, Y., Liu, X., & Yang, Y. (2013). Identifying Vulnerable Nodes of Complex Networks in Cascading Failures Induced by Node-Based Attacks. *Mathematical Problems in Engineering*.

Li, Y. (2014). Networked Analysis Approach of Supply Chain Network. *Journal of Networks, 9*(3), 777-784.

Liu, L., & Qi, X. (2014). Network disruption recovery for multiple pairs of shortest paths. *2014 11th International Conference on Service Systems and Service Management (ICSSSM).* Beijing, China: IEEE.

Lu, Q., Korniss, G., & Szymanski, B. K. (2006). Threshold-Controlled Global Cascading in Wireless Sensor Networks. *Proceeding of third International Conference on Networked Sensing Systems*, (pp. 164-171).

Morris, S. (2001). Political Correctness. *Journal of Political Economy, 109*(21).

Motter, A. E. (2004). Cascade Control and Defense in Complex Networks. *Phys. Rev. Let., 93*.

Motter, A. E., & Lai, Y.-C. (2002). Cascade-based Attacks on Complex Networks. *Physical Review E, 66*.

Nanda, S., & Kotz, D. (2012). Localized Bridging Centrality. In M. T. Thai, & P. M. Pardalos, *Handbook of Optimization in Complex Networks* (pp. 197-218). New York: Springer New York.

Neumayer, S., Zussman, G., Cohen, R., & Modiano, E. (2011). Assessing the Vulnerability of the Fiber Infrastructure to Disasters. *IEEE/ACM Transactions on Networking, 19*(6), 1610-1623.

Newman, M. E. (2001). Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. *Physical Review E, 64*(1), 8577-8582.

Newman, M. E. (2002). Random Graphs as Models of Networks. *SFI Working Paper*.

Newman, M. E. (2003). The structure and function of complex networks. *SIAM Rev, 45*(2), 167-256.

Newman, M. E. (2006). Modularity and community structure in networks. *PNAS, 103*(23).

Newman, M. E., Barabasi, A.-L., & Watts, D. J. (2006). *The Structure and Dynamics of Networks.* Princeton: Princeton University Press.

Newman, M. E., Strogatz, S. H., & Watts, D. J. (2001). Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E, 64*.

Newman, M., & Girvan, M. (2004). Finding and Evaluating Community Structure in Networks. *Physical Review E 69*.

Ni, J., & Chandler, S. (1994). Connectivity Properties of a Random Radio Network. *IEE Proceedings - Communications, 141*(4), 289-296.

Nof, S. Y. (2003). Design of Effective e-Work: review of models, tools, and emerging challenges. *Production Planning & Control, 14*(8), 681-703.

Nof, S. Y. (2007). Collaborative control theory for e-Work, e-Production, and e-Service. *Annual Reviews in Control, 31*, 281-292.

Pishro-Nik, H., Chan, K., & Fekri, F. (2009). Connectivity properties of large-scale sensor networks. *Wireless Networks, 15*(7), 945-964.

Reyes Levalle, R., & Nof, S. Y. (2015). A resilience by teaming framework for collaborative supply networks. *Computers & Industrial Engineering, 90*, 67-85.

Reyes Levalle, R., & Nof, S. Y. (2015). Resilience by teaming in supply network formation and re-configuration. *Internation Journal of Production Economics, 160*, 80-93.

Shuang, Q. et al. (2015). A Cascade-Based Emergency Model for Water Distribution Network. *Mathematical Problems in Engineering*.

Shuang, Q., Zhang, M., & Yuan, Y. (2014). Node vulnerability of water distribution networks under cascading failures. *Reliability Engineering & System Safety, 124*, 132-141.

Singh, P., Sreenivasan, S., Szymanski, B. K., & Korniss, G. (2013). Threshold-limited Spreading in Social Networks with Multiple Initiators. *Scientific Reports*.

Sun, J., Zhao, Y., & Lu, Q.-C. (2015). Vulnerability Analysis of Urban Rail Transit : A Case Study of Shanggai. *Sustainability, 7*, 6919-6936.

Surana, A., Kumara, S., Greaves, M., & Raghavan, U. N. (2005). Supply-chain networks: A complex adaptive system perspective. *Int. J. Prod. Res., 43*(20), 4235-4265.

Tamura, H., Sengoku, M., & Shinoda, S. (1990). Location Problems Undirected Flow Networks. *The Transaction of The IEICE, 73*(12).

Tamura, H., Sengoku, M., Shinoda, S., & Abe, T. (1992). Some Covering Problems in Location Theory on Flow Networks. *IECE Trans. Fundamentals, 75*(6).

Tran, T. D., & Kwon, Y. K. (2013). The relationship between modularity and robustness in signalling networks. *Journal of Royal Society Interface, 10*(88).

Velasquez, J. D., Yoon, S. W., & Nof, S. Y. (2010). Computer-based collaborative training for transportation security and emergency response. *Computers in Industry, 61*(4), 380-389.

Viswananth, K., & Peeta, S. (2002). The multicommodity maximal covering network design problem. *IEEE 5th International Conference on Intelligent Transportation Systems.* Singapore: IEEE.

Wagner, S. M., & Neshat, N. (2010). Assesing the Vulnerability of Supply Chains using Graph Theory. *Int J. Production Economics, 126*, 121-129.

Watts, D. J. (2002). A Simple Model of Global Cascades on Random Networks. *PNAS, 99*(9), 5766-5771.

Watts, D. J., & Dodds, P. (2007). Influentials, networks, and public opinion formation. *Journal of Consumer Research, 34*, 441-458.

Watts, D. J., & Strogatz, S. H. (1998). Collective dynamics of 'small-world' networks. *Nature, 393*(6684), 440-442.

Xu, D., & Girgis, A. A. (2001). Optimal load shedding strategy in power systems with distributed generation. *Power Engineering Society Winter Meeting* (pp. 788-793). Columbus: IEEE.

Xue, F., & Kumar, P. R. (2004). The Number of Neighbors Needed for Connectivity of Wireless Networks. *Wireless Networks, 10*(2), 169-181.

Yangqi, H., & Zhang, J. (2006). Reliability Evaluation of Main Electrical Connection of Substation based on Graph Theory. *Applied Mechanics and Materials, 527*, 273-276.

Zhang, L., Zhong, H., & Nof, S. Y. (2015). Adaptive Fuzzy Collaborative Task Assignment for Heterogeneous Multirobot Systems. *International Journal of Intelligent Systems, 30*(6), 731-762.

Zhong, H. (2016). *Dynamic Lines of Collaboration in E-Work Systems (Unpublished doctoral dissertation).* West Lafayette: School of Industrial Engineering, Purdue University.

Zhong, H., & Nof, S. Y. (2014). Dynamic lines of collaboration in CPS disruption response. *Proceedings of the 19th IFAC world congress.*

Zhong, H., & Nof, S. Y. (2015). The dynamic lines of collaboration model: Collaborative disruption response in cyber–physical systems. *Computers & Industrial Engineering, 87*, 370-382.