Modelling and Simulation of Interdependencies between the Communication and Information Technology Infrastructure and other Critical Infrastructures

by

Hafiz Md. Abdur Rahman

A THESIS SUBMITTED IN PARTIAL FULFILLMENT OF
THE REQUIREMENTS FOR THE DEGREE OF

DOCTOR OF PHILOSOPHY
in

The Faculty of Graduate Studies
(Electrical and Computer Engineering)

THE UNIVERSITY OF BRITISH COLUMBIA
(Vancouver)
September 2009

Abstract

Critical infrastructures are the lifelines of modern societies. The Communication and Information Technology Infrastructure (CITI) provides the basic mechanisms for sharing control and decision-making information among different critical infrastructures. Failures in CITI, either due to an accident or malicious action can propagate to other infrastructures and degrade or disrupt their functionality. Conversely, failures in other infrastructures can also propagate to CITI and hence disrupt the operation of many of the interconnected systems. For reliable and consistent operation of critical infrastructure networks, it is important to have tools and techniques to model and simulate CITI related interdependencies. This research is focusing on developing such methods and tools for CITI interdependency modelling and simulation. Our approach is based on system engineering techniques, where critical infrastructures are viewed as a system of systems. Interdependencies between different system components are captured using precise mathematical functions. As such, our approach goes beyond the limitations of agent-based modelling and simulation paradigms, where interdependencies are considered an emergent behavior. In this research, we have used predictive modelling techniques commonly used in power systems, data communication networks and information systems. The approach is based on results from real CITI interdependency related data. In our model, we used these data to identify the origin of different types of CITI failure and their impacts on critical infrastructures. Following that, we developed techniques to estimate interdependencies between CITI and other critical infrastructures. Finally, we developed techniques to simulate CITI interdependencies in a critical infrastructures simulator. The simulation results were validated against real-life failure cases. Our approach gives a comprehensive solution to CITI interdependency modelling and simulation problems and hence is an important step in the critical infrastructure related research. Even though our techniques are developed for CITI interdependency, they will be useful for other critical infrastructure networks as well.

Table of Contents

List of Tables

List of Figures

List of Acronyms

ACM	Association for Computing Machinery
BIDMC	Beth Israel Deaconess Medical Center
CAS	Complex Adaptive Systems
CI	Critical Infrastructures
CITI	Communications and Information Technology Infrastructure
DHS	Department of Homeland Security
EBNF	Extended Backus Naur Form
EMTP	Electromagnetic Transients Program
FCC	Federal Communications Commission
FFM	Fluid Flow Model
HRT	Human Readable Table
I2Sim	Infrastructure Interdependencies Simulator
IP	Internet Protocol
ICT	Information and Communication Technology (same as CITI)
JIIRP	Joint Infrastructures Interdependencies Research Program
LAN	Local Area Network
MATE	Multi-Area Thevenin Equivalent
NISAC	National Infrastructure Analysis and Simulation Center
OC	Optical Carrier
ODE	Ordinary Differential Equation
OSI	Open Systems Interconnection
OVNI	Object Virtual Network Integrator (simulator)
PSTN	Public Switched Telephone Network
SONET	Synchronous Optical Network
STP	Spanning Tree Protocol
TCP	Transmission Control Protocol

List of Notations

$[G]$	interdependency matrix
$[A]$	clusters' coefficient matrix
$[p]$	clusters' current (tokens) injection vectors
$[z]$	matrix of links' Thevenin impedances
$[h_A]$	vector of clusters' accumulated currents (tokens)
$[V_\alpha]$	vector of link Thevenin voltages
$[v_A]$	subsystems' nodal voltages
$[i_\alpha]$	vector of link currents (tokens)
$x(t)$	sending tokens
$y(t)$	receiving tokens
α	loss factor
τ	delay
M	maximum TCP window size
$W(t)$	expected window size at time t
$R(t)$	round-trip time
$\lambda_i(t)$	loss indication rate experienced by a flow class i
n_i	set of TCP classes
C	link capacity (bandwidth) of the queue
$A(t)$	packet arrival rate
$B(t)$	packet throughput

Acknowledgments

This will be a long list if I have to acknowledge all those wonderful people who have influenced and shaped my vision and ideas over the years. However, there are those who ought to be acknowledged for their help for this research and dissertation. First, I take this opportunity to thank my supervisor Prof. José R. Martí for his guidance and support during my PhD studies. He extended his hand at a crucial moment of my PhD program that I will never forget. Without his help at that time, I would never have been able to think of finishing my degree. While working under his guidance, I have been significantly enriched from his vast knowledge of electrical engineering. It was also a great source of insight and inspiration to follow Prof. Martí's ideas that he continuously shared in countless hours of discussions with his students. Regarding his vision of I2Sim, I always felt as English poet William Blake said:

> *To see a world in a grain of sand*
> *And a heaven in a wild flower,*
> *Hold infinity in the palm of your hand*
> *And eternity in an hour.*

Accordingly, we were entrusted to follow him to accomplish this vision. I was also amazed to note Prof. Martí's sharp mind, warm heart and wonderful sense of humor. It was also a valuable experience to learn how Prof. Martí and Prof. Srivastava have assembled and inspired a large group of talents to accomplish their vision. In this regard, I also acknowledge my other advisory committee members Prof. K. D. Srivastava, Prof. Juri Jatskevich, Prof. Lutz Lampe and Prof. Hermann Dommel for their interest and encouragement in my research. In particular, I would like to thank Prof. Srivastava and Prof. Jatskevich for their insightful comments and suggestions at different phases of my research. My special thanks go to Mazana (Dr. Mazana Armstrong) for her help to build the full working prototype of the I2Sim simulator in this research. She sacrificed lots of her time from her busy schedule and from family life for guiding the development of the

simulator. She has been a true friend and a great mentor. I also thank Prof. Konstantin Beznosov for guiding part of my work. I also thank the faculty and staff of the Department of Electrical and Computer Engineering at UBC for providing an excellent environment for research.

I also thank my Master's supervisor, Prof. Edward J. Coyle at Purdue University (now at Georgia Tech), for deeply influencing my thoughts and engineering knowledge. The approach and methods I learned from him have become an important contributor during my PhD research. Prof. Coyle has been my great example for all my professional work. I also would like to thank my Master's committee member at Purdue, Prof. Arif Ghafoor, and my supervisor during my software engineering job at Purdue, Ms. Julayne Moser. Both of them encouraged me to pursue the doctorate degree. I thank the external examiner of my work, Prof. Don Towsley of the University of Massachusetts for his many important suggestions and detail comments on my thesis. I also thank Prof. Yong Liu at Polytechnic Institute of NYU for providing me the source code of the Fluid Flow Model simulator, which was one of the important components of my hybrid model solution technique.

I also would like to thank my friends and colleagues at the Electric Power and Energy Systems Group. They include DeTao Mao, Marcelo Tomim, Tom De Rybel, Lu Liu (Lucy), Quanhong Han, Michael Wrinch and others. Each of them has been a great friend. Discussions of technical subjects with them have been very insightful; and non-technical subjects, such as sports, politics, economy and jobs were always a great pleasure as well. While doing my research, I was also greatly benefited from the Google Scholar search engine, Boost C++ libraries, Lyx text editor, JabRef reference manager and many other tools, systems and information. I am acknowledging the contribution of the creators of these resources, many of which they voluntarily dedicated freely to the advancement of scientific knowledge.

I would also like to thank my mother Farida Khatun for her constant well wishes. I also thank my father, late Engineer Abdul Gafur, who always had high expectations for his sons. The most important inspiration has been from my wife Dr. Nazmun Nahar. Her love and support have sustained me throughout my entire PhD studies. She has been my best friend to share my grief and joy of every moments of my research. Being a civil engineer, she did not hesitate to read my writings and to give her valuable suggestions to improve these. My dad and my wife are graduates of the same engineering school in Bangladesh and both of them influenced my life significantly. I also thank my daughter Marihah and son Ahmad for their patience and understanding during my busy hours. I express my deepest gratitude to all these family members and dedicate this dissertation to them.

Most surely in the creation of the heavens and the earth and the alternation of the night and the day there are signs for men who understand. Those who remember Allah standing and sitting and lying on their sides and reflect on the creation of the heavens and the earth: Our Lord! Thou hast not created this in vain! Glory be to Thee. (Al-Quran 3:190-191)

This work is dedicated to my father late Abdul Gafur, mother Farida Khatun, wife Nazmun Nahar, daughter Marihah and son Ahmad.

Chapter 1

Introduction

The technological developments in the last few decades have created large and complex networks for the delivery of critical services to the people of modern societies. These include electricity, telecommunication, water distribution, healthcare services and financial services, etc. Our everyday living largely depends on the smooth functioning of these critical infrastructures (CI). Even though we are not much aware of their contribution in our daily life, any disruption of these services may create catastrophic impacts. These critical infrastructure networks rely on the Communications and Information Technology Infrastructure (CITI) for their consistent output. CITI dependencies are typically related to control and decision-making services that run internally within the infrastructure facility or may come externally from other CITI service providers through electronic communication links. Interdependencies between critical infrastructures and CITI are also known as cyber interdependencies and have been categorized as one of the four principal classes of critical infrastructure interdependencies (the other three are physical, geographical and logical interdependencies) [1]. Cyber interdependencies have significant impact on the operation of healthcare, banking and financial services, air transportation services, etc.

An example of cyber interdependency for healthcare infrastructure can be given. On 13 November 2002, regular operation of Beth Israel Deaconess Medical Center at Harvard University was significantly disrupted due to a computer network outage [2–4]. Doctors could not order medication or lab reports electronically, no decision support software was available, and there were many other problems. The increasing dependency of the critical infrastructures on CITI services makes this kind of incidents more likely. However, could this failure be foreseen and be avoided? Should the network be upgraded to a newer archi-

tecture that could avoid this catastrophic failure? Or, could that add more vulnerabilities to the hospital facility? Were there other hidden infrastructure interdependencies that could make the hospital operation unstable? These are very reasonable questions for any critical facility's design and operation. For infrastructure operators and managers, it is important to have tools and techniques to answer these kinds of queries.

The more general research questions are: Which types of CITI failures have major impacts on critical infrastructures? What are the impacts of these failures in spatial and temporal dimensions? Do they have any implication on public safety? Is there a way to quantify cyber interdependencies for modelling and simulation? If so, how can we construct a functional relationship between CITI and other critical infrastructures? How can we simulate cyber interdependencies for vulnerability identification, reinforcement planning, adding or discarding redundancies in the infrastructure networks? These are important, but unanswered questions that form the basis of this scientific exploration. The aim of this research is to develop techniques to address these questions thoroughly, so that reasonable conclusions can be made.

Our approach is based on system engineering techniques, where infrastructures are viewed as a system of systems, and interdependencies between infrastructures are captured using mathematical functions. In our model, we used real CITI interdependency related data to identify the origin of different types of CITI failure and their impacts on critical infrastructures. Following that, we developed techniques to estimate interdependencies between CITI and other critical infrastructures. A set of CITI interdependency functions were estimated for different critical infrastructures using our procedure. Following that, a prototype I2Sim [5] critical infrastructure simulation framework was built. Finally, we developed a hybrid simulation technique to extend the capability I2Sim framework to simulate domain specific events. The CITI interdependency was simulated in I2Sim as a special case of hybrid technique. A real CITI failure case was simulated to show the usefulness of our approach. In this thesis, we have used cyber interdependency and CITI interdependency interchangeably. Also, in the research literature CITI is often written as ICT (Information and Communication Technology) infrastructure. However, we have used the term CITI in accordance with the critical infrastructures' definition by the Government of Canada [6]. This work has been carried out under the auspices of the Joint Infrastructures Interdependencies Research Program (JIIRP) [7] of the Government of Canada.

1.1 The I2Sim Infrastructure Modelling and Simulation Framework

Effective response during any large disaster scenario requires coordinated decision-making among different critical infrastructure operators. Over the years, the integration among critical infrastructures has become pervasive, extensive and complex. The CIIS (Complex Interdependent Integrated Systems) research group (also known as I2Sim research group [8]) at the University of British Columbia has developed the I2Sim modelling and simulation framework [5] for effective coordination and decision-making during disasters for multiple infrastructure operators while retaining their confidentiality. The framework helps infrastructure operators plan-response and recovery actions during large disasters. In the I2Sim framework, infrastructure entities are represented as a system of systems. As the internal detail of each system is not exposed to the outside entities, this allows multiple infrastructure operators to coordinate while maintaining their respective organization level security procedures and confidentiality requirements. This approach is particularly helpful for CITI, as different groups, such as end users, content providers, service providers, network operators, device manufacturers and application developers use the same network infrastructure for different business models and operational objectives.

While showing the utility of I2Sim framework, we have selected a hospital infrastructure as a simulation case. This is because a hospital is a very critical infrastructure in any disaster scenario. Understanding the servicing capacity of a hospital has great benefit for any disaster response community. As such, a hospital has received special attention from the beginning of I2Sim project [5, 9, 10].

1.2 General Background and Motivation

There have been several motivations behind this research. The principal motivation comes from the importance of the problem itself. Our study of CITI failure analysis [11, 12] shows that cyber interdependencies have large impact on many of the critical infrastructures, such as healthcare, banking and financial services, administration and public safety services and air transportation services, etc. Due to a growing use of CITI services for automation and control of infrastructure services, cyber interdependency is becoming pervasive in all infrastructures, and has a very large impact on many different aspects of our society.

Despite this considerable importance, to the best of our knowledge, there has not been much research done on modelling and simulating cyber interdependency among different critical infrastructures. Pederson et al. [13] have compiled a survey on contemporary research on critical infrastructure modelling and simulation. This is a comprehensive-study that shows a wide variety of new ideas proposed in the recent years. However, only a handful of frameworks have the capability to model and simulate cyber interdependencies. Accordingly, many questions mentioned in the previous section are not yet answered. As such, this research was initiated to develop methodologies and tools that will help researchers and practitioners to identify, model and simulate cyber interdependencies for any man-made or natural disaster scenario. In this section, we give an overview of those works that are related to CITI interdependencies and are available in the published literature.

The US Department of Homeland Security (DHS) started a program in 2004 named National Infrastructure Analysis and Simulation Center (NISAC) [14]. Sandia National Laboratories and Los Alamos National Laboratory are the main contractors for NISAC. The key objective of this program is to develop models and simulation tools for US critical infrastructures. Sandia National Laboratory, with the help of other organizations, has developed models and simulation tools for different critical infrastructures [14, 15]. Among these tools, Telecommunications Network Simulation Modelling and Analysis Tools (N-SMART) [16,17] have been developed jointly by Sandia National Laboratory and Lucent Technologies with the intention to model and simulate interdependencies between the telecommunications network and other critical infrastructures. Published literature shows that N-SMART is based on statistical models of telephone network's call servicing capacity [17,18]. It has been used to study the impact of hurricanes Katrina, Rita and Wilma on telephone networks [19]. There are two important limitations of the N-SMART model. First, it is solely based on the telephone network, where the objective is to identify how external failures affect voice communications. It does not address failures related to data communication networks. However, our study on publicly reported failure cases [11] shows that the vast majority (more than 85%) of the cyber interdependencies are related to failures of data communication networks. Second, N-SMART does not have any model for cross infrastructure interdependencies [20,21]. For instance, there is no way to know how failures in the telephone network may affect hospitals, fire services, or academic institutions. The

information we can get from this model is how the availability of voice communication (wire and wireless) is affected due to a natural or man-made disaster [17, 18].

In Europe, an agent-based model named Critical Infrastructure Simulation by Interdependent Agents (CISIA) was developed by Panzieri et al. [22]. This framework simulates critical infrastructures where the telecommunications network is a component of the overall architecture. However, no experimental results were presented showing how failures in the telecommunications network affect other critical infrastructures. A recent article on CISIA [23] gives more detail about the simulator's internal architecture and simulation capabilities. However, no description for cyber interdependency simulation was given in the article. There is also an electric power and communication simulator named EPOCHS [24]. The purpose of EPOCHS is to accommodate the role of data communication networks and electrical power networks in a single framework, so that both of their role in the electrical infrastructure can be simulated. For instance, it can simulate the role of TCP/IP based protection network within the electrical power networks. An agent-based architecture is used to integrate both electrical and data communication simulator components. However, EPOCHS has no generalized model to simulate cyber interdependencies between different critical infrastructures. Other than these frameworks, the survey by Pederson et al. [13] mentions three other frameworks, Athena, CIP/DSS and Fort Future, as having capabilities to simulate data communication networks related interdependencies. However, no detail was found in the published literature regarding these frameworks.

My personal motivation comes from my background in computer networks, power systems and software engineering. I have had education or work experience in all these three areas during my career. This has given me the opportunity to explore the MATE model [25], the Cell-Channel model [10], Latency techniques [26, 27], Bedell's method [28], Fluid Flow model [29, 30] and to be part of the team that has built the I2Sim critical infrastructures simulator [20]. The ideas of MATE and other models are pivotal to formalize the complex nature of cyber interdependency and the environment where they exist. This research will make these ideas more accessible to different groups, such as, information system researchers, telecommunication systems designers and public safety practitioners and may become common knowledge to understand this field.

Another motivation comes from the research objective of the I2Sim research group [8] at UBC, which is to understand, model, and simulate infrastructure interdependencies

during large disasters. The goal of the I2Sim group is to develop effective decision-making tools, which will be used by policy makers and infrastructure service providers to maximize the number of human lives saved during natural or man-made disaster scenarios. Such a humanitarian goal was an important motivation for this scientific exploration.

1.3 Research Approach and Contributions

The objective of this research is to model and simulate interdependencies between CITI and other critical infrastructures. A bottom-up approach was taken to achieve this target. First, we built our understanding of CITI failure sources and their impacts on different infrastructures from real-life failure cases. Following that, we built an interdependencies model between CITI and other infrastructures. Finally, we incorporated this interdependencies model into a critical infrastructures simulator for cyber interdependency simulation. More specifically, we divided our tasks into the following three objectives:

- To identify the origin of different types of CITI failure and their impacts on critical infrastructures.
- To develop methods to estimate interdependencies between CITI and other critical infrastructures.
- To develop techniques to simulate cyber interdependencies in a critical infrastructures simulator.

The first objective was important in order to understand the origin of CITI failures and their propagation patterns. To develop our knowledge from real-life CITI failure cases, we collected 347 infrastructure failure cases from the Association for Computing Machinery's RISKS forum that were reported from 1994 to 2005. We studied these reports [11, 12] to determine the causes of infrastructure failures and their impact on CITI and other critical infrastructures in a number of dimensions, such as origin of failures, impacts of failures in spatial and temporal dimensions, their effect on public safety and how failures propagate from one infrastructure to another. We also identified eight orthogonal fault classes and generic faults that belong to each of these classes.

To accomplish the second objective, we have developed a technique [21] to formalize cyber interdependencies for different critical infrastructures and for their representation we estimated a set of empirical functions. These functions establish a relationship between infrastructure output to the corresponding CITI service input. Our approach was to identify

important CITI services for each of these infrastructures and systematically rank these services according to their contribution to the infrastructure's output. The information used to construct these empirical functions was collected from contemporary research, interviews of the infrastructure operators, and from the CITI failure reports that were collected in our previous study.

In this project, we built an implementation of the critical infrastructures simulator I2Sim [5, 10]. Our I2Sim implementation [20] is based on the matrix partitioning technique "Multi-Area Thevenin Equivalent (MATE)" developed in [25, 31]. The accuracy of the I2Sim simulator has been confirmed through various benchmark tests. To address the third objective of simulating cyber interdependencies, we extend I2Sim's capability using a hybrid simulation approach. In the hybrid technique, we partitioned the infrastructure components into two mutually exclusive sets: long time-step system components and short time-step system components. Long time-step (e.g., $\Delta T = 5$ minutes) components were used to model regular infrastructure events and were simulated through I2Sim. The short time-step (e.g., $\Delta t = 5$ milliseconds) components captured fast changing events, such as events in data communication networks and were simulated using a Fluid Flow Model (FFM) [30] simulator. The results from both simulators were periodically combined using latency techniques [26, 27]. The hybrid approach was validated against real-life failure cases.

1.4 Outline of this Dissertation

This dissertation has six chapters. Chapters 2 and 3 are related to the first two research objectives. Chapters 4 and 5 are related to the third objective, as described below:

Chapter 2 is related to CITI fault identification and impact analysis. The chapter gives a background of the research problem, develops a methodology to collect and analyze real-life data and presents the results of the analysis.

Chapter 3 is related to estimating CITI interdependencies for different critical infrastructures. We developed a methodology to construct interdependency functions based on the importance and effectiveness of different CITI services for different critical infrastructures.

Chapter 4 is related to the design and implementation of a version of the I2Sim critical infrastructures simulator. A detail account of I2Sim's architecture is discussed and some simulation results are presented.

Chapter 5 presents an extension techniques to I2Sim to simulate CITI interdependencies. The idea of a hybrid simulation framework is introduced and implementation techniques are discussed. This is followed by simulating the results of a real-life scenario.

Chapter 6 summarizes the outcomes of this research. Since this is a promising new field of investigation, some suggestions and ideas for future research are discussed as well.

1.5 List of Publications

Work presented in Chapter 2 was first published in the CRIS 2006 conference and was given the best paper award from the Third International Conference on Critical Infrastructures, Alexandria, VA. CRIS 2006, September 2006. The paper was later published in a special issue of the International Journal of Critical Infrastructures.

- Hafiz Abdur Rahman, Konstantin Beznosov and José R. Martí, "Identification of Sources of Failures and their Propagation in Critical Infrastructures from 12 Years of Public Failure Reports", International Journal of Critical Infrastructures, Vol. 5, No. 3, 2009, pp. 220-244.

Work presented in Chapter 4 was published in the IEEE EPEC 2008 conference. A manuscript is now being prepared for a journal submission.

- Hafiz Abdur Rahman, Mazana Armstrong, DeTao Mao and José R. Martí, "I2Sim: A Matrix-partition based Framework for Critical Infrastructure Interdependencies Simulation," Proceedings of IEEE 8th Annual Electrical Power & Energy Conference 2008 (EPEC 2008), Vancouver, Canada, 6-7 October, 2008.

Works presented in Chapter 3 and Chapter 5 are now available in the following technical report format. These will be submitted for publication soon.

- Hafiz Abdur Rahman and José R. Martí, "Quantitative Estimates of Critical Infrastructure Interdependencies on Communication and Information Technology Infrastructure," University of British Columbia, Tech. Rep. I2SIM-TR-2008-02, 22 December 2008.

- Hafiz Abdur Rahman and José R. Martí, "A Hybrid Systems Model to Simulate Cyber Interdependencies between Critical Infrastructures," University of British Columbia, Tech. Rep. I2SIM-TR-2009-01, 18 May 2009.

Chapter 2

CITI Fault Identification and Impact Analysis

Understanding the origin of infrastructure failures and their propagation patterns in critical infrastructures is important for secure and reliable infrastructure design. Among the critical infrastructures, Communication and Information Technology Infrastructure (CITI) is crucial, as it provides the basic mechanism for sharing information among other critical infrastructures [32], such as electricity, water supply, oil and gas networks, transportation, financial services, etc. Over the years, integration of these infrastructures with CITI has become pervasive, extensive, and complex. Failure in CITI, either due to an accident or a malicious action, can propagate to other infrastructures and degrade or disrupt their functionalities. Conversely, failures in other infrastructures can also propagate to CITI and hence disrupt the operation of many of these interconnected systems. Such disruptions may lead to substantial disturbances in the public life of modern nation states. Volatile world situations increase these threats even further. As a result, there are enormous concerns for secure and reliable operation of different critical infrastructures [32, 33]. Smooth operation of these interconnected infrastructures requires an understanding of their interdependencies. By studying the origin of infrastructure related failures, their propagation patterns and their relation to the topology of scale-free networks [34], we can develop a better understanding of their interdependencies, which can be useful to decision makers and infrastructure operators in policy making and system design [7].

Since 1992, US telephone companies have been required to submit major failures information to the US Federal Communications Commission (FCC). Using FCC outage reports, a study was done on the failure patterns of Public Switched Telephone Networks

(PSTN) [35]. To the best of our knowledge, no other critical infrastructure providers in North America are obliged to disclose their failure-related information. However, such data could help the research community develop a better understanding of failure patterns and their interdependencies among different critical infrastructures. Data from infrastructure service providers is especially helpful, because they may give detailed information about system states, control parameters, input and output specifications, operating assumptions, backup facilities, management procedures and practices, and other physical and environmental constraints [36]. Unfortunately, both public and private infrastructure operators are reluctant to share this information with the research community [37]. The FCC outage report mentioned above is accessible only to the FCC officials and the US Department of Homeland Security (DHS) [36].

Given this reality, one possible alternative is to use public domain infrastructure failure reports, such as newspaper or other mass media reports to develop an understanding of infrastructure interdependencies. There are two major difficulties in this approach;

1. These failure reports normally provide only brief information, and
2. They do not have any formal structure.

Even though individual reports may not give much information about a specific failure, by studying a large number of cases we can trace common trends among similar classes of failures. To address the second obstacle, we have classified these reports based on their failure type, and extracted meaningful information through some critical attributes. We collected 347 cases of 12 years of failure data (1994 to 2005) from the Association for Computing Machinery's (ACM) RISKS forum [38], which is the largest known public repository of for this kind of reports. The RISKS forum was started more than 20 years ago and is still very active. Posting to this forum is moderated, which ensures a certain level of quality and reliability. Since the cases reported to this forum represent only a fraction of all actual events, there may be a concern about the usefulness of the statistics we derive from our analysis. However, the trend of reporting to this forum is related to the public perception of risks, and research shows that despite partial information public perception of risks is fairly accurate [39,40]. To ensure the authenticity of our selected cases, we have paid particular attention to the verifiability of our selected reports' sources. For instance, we gave more weight to newspaper reports than to reports by private individuals. Our methodology is discussed in detail in Section 2.2. A recently reported work (2007) on

infrastructure interdependencies by Chang et al. [41] also uses newspaper reports as data sources.

In this work, we have identified interdependencies among CITI and other infrastructures based on some key factors, such as origin of failures, their impact in spatial and temporal dimensions, their effect on public safety, and their propagation from CITI to other critical infrastructures and vice versa. More specifically, we would like to determine the main causes of infrastructure failures and the nature of their impact; the type of locality affected and their geographical location; how their degree of fatality changed over time; and how infrastructures are related to each other. In the absence of any formal model of interdependencies among CITI and other critical infrastructures, our findings should be useful to policy makers, practitioners, and researchers. In Section 2.1, we discuss previous efforts to classify and interpret infrastructure-related failures. In Section 2.2, we give a brief overview of our methodology. In Section 2.3, we describe our failure database. In Section 2.4, we summarize the results of our analysis. Finally, in Section 2.5, we discuss the contributions of this research as well as future research directions.

2.1 Related Work

There are three major approaches to classifying and interpreting infrastructure related failures. The first approach focuses on failures and their impacts in relation to CITI [35, 42, 43]. The second approach focuses on understanding failures in any computer-based system, and is not limited to CITI [44, 45]. The third approach classifies failures and interdependencies among the critical infrastructures in a general system independent way [1].

Richard Kuhn [35] analyzes public switched telephone network (PSTN) failure data based on the following six failure categories: human errors, acts of nature, hardware failures, software failures, overloads, and vandalism. Using this scheme, Kuhn analyzes two years of PSTN failure data (1992-1994) from the US Federal Communication Commission (FCC) and shows the impact of different types of failures on PSTN operation. John Howard [42] proposes a taxonomy based on malicious attacks in computer networks and uses that taxonomy to perform a frequency analysis of more than 4,000 security-related incidents reported to the Computer Emergency Readiness Team Coordination Center (CERT/CC). From the results of this analysis, he proposes a set of recommendations for government,

vendors, the CERT/CC, and individual users to improve security practices. Howard and Longstaff [46] further extend this taxonomy by incorporating additional terms, such as additional objects and attributes like site name, attack date, reporting time, etc. Chakrabarti and Manimaran [43] propose a taxonomy to classify Internet infrastructure security failures, based on a survey of different intrusion detection and prevention techniques. They classify Internet infrastructure failures into four categories: DNS hacking, routing-table poisoning, packet mistreatment, and denial of service attacks.

Peter Neumann initiated the Association for Computing Machinery's (ACM) RISKS forum [38] in 1985 to compile computer-related system mishaps that affect public life. In 1994, Neumann published the book *Computer-Related Risks* [44], in which he selectively compiled a large collection of RISKS forum reports based on problem sources. These reports included problems in requirement definition, system design, hardware implementation, software implementation, system use and operation, operating environment, etc. Through his analysis, Neumann draws attention to the safety and security issues associated with each type of failure. Avizienis et al. [45] propose another generalized taxonomy based on the reliability and security aspects of computer systems. Their approach is to compile a few key definitions as a set of generalized concepts and to extend those concepts with extended sets of definitions. The main objective of this taxonomy is to be able to use these concepts in a wide variety of cases.

Rinaldi et al. [1] address classification of failures and interdependencies among the critical infrastructures in a system independent way. Their taxonomy is based on six functional dimensions to determine cross-infrastructure interdependency issues. These are: type of interdependency (e.g., physical, cyber, logical); infrastructure environment (e.g., business, economic, healthcare); coupling and response behavior (e.g., adaptive, loose, tight); infrastructure characteristics (e.g., temporal, spatial, organizational); type of failure (e.g., common cause, cascading, escalating); and state of operations. (e.g., normal, stressed, repaired). However, their failure source classification is very restrictive (e.g., common cause, cascading, escalating) and gives a very limited number of options for analyzing the RISKS forum failure reports.

In our research, we used Kuhn's [35] approach for CITI and related critical infrastructure failure classification. However, we added the following two additional categories to Kuhn's original six: malicious logic fault and authorization violation fault. Even though these two faults are software related, due to their intentional and malicious nature, we

placed them in separate categories. In recent years, these two faults have become of increasing concern for critical infrastructure management. Their remedial methodologies are also different from those used for traditional software failures.

2.2 Approach and Methods

We have followed a four-step methodology in collecting and analyzing failure reports. We started by systematically collecting failure cases from the RISKS forum. We then categorized these reports based on their failure type, extracted useful information from them, and then performed an analysis of the extracted information. The following sections discuss these steps in detail.

2.2.1 Data Collection

Postings in the RISKS forum cover a wide range of computer related risk topics, including system failure reports, conference announcements, book reviews, etc. Collecting useful infrastructure failure reports from RISKS forum data (we scanned more then 10,000 records) was the most difficult, time consuming but important step. During our selection process, we selected only those reports where the failure originated from CITI and affected other critical infrastructures (including CITI), or the failure originated from some other critical infrastructure and affected CITI. Failure is defined as the inability of a system or component to perform its required functions within the specified performance requirements [47], and may be the result of one or many faults. A fault is defined as a defect in a hardware device or component, or an incorrect step, process, or data definition in a computer program [47]. In our study, failure is attributed to critical infrastructures. The following infrastructures are considered critical, based on a US Congress document on critical infrastructures identification [32]:

- IT Infrastructure
- Telecommunication Infrastructure
- Water Supply
- Electrical Power System
- Oil and Gas
- Road Transportation
- Railway Transportation

- Air Transportation
- Banking and Financial Services
- Public Safety Services
- Healthcare System
- Administration and Public Services

We collected a unique report for each incident. However, in cases of widespread failure, we collected unique reports from different affected sites. One example of such widespread failure is an Internet-wide worm attack. The apparently simple task of selecting appropriate sets of reports became quite complicated due to the subtleties associated with each of the reports; the following three examples reveal this intricacy. We identified these three reports using reference IDs as RISKS (i, j), where i is the volume number and j is the issue number within the volume as available from the RISKS forum website [38]. This is because the second and third reports are not included in our database. The first is an example of a report that we chose to include in our database:

> *On November 19, 1994, Iowa City's US West telephone system shut down at about 3:30 p.m., local time, and service was gradually restored between 7:30 and 9:30 p.m, affecting about 60,000 people. Analysis showed that a new switching system had been installed in July 1994. In removing the old system, an electrical grounding cable had been inadvertently removed. Iowa City Press Citizen, November 22, 1994, "Extended phone failure in Iowa City", RISKS (volume 16, issue 58)*

The above example clearly shows that a fault in the electrical power system due to human error caused a failure in the telecommunications infrastructure. The report has a clear reference to a newspaper source. In contrast to the above report, the following is an example of a report that we omitted:

> *I suppose I shouldn't be surprised, but the power went out for 17,000 here in our small town (38,000) last week. The local newspaper first reported that the power company didn't know why it went out, but that it "may be related to someone digging in their back yard". A week later they fixed the blame. A phone call (by the power company), supposedly to one substation, (completely automated judging by the tone of the article) went instead to a different substation (for unexplained reasons) and shut that substation down. It was down for 1.5 hours. "Make a Call, Turn Off the Power", RISKS (17, 4)*

In the above report, the failure in the electrical power system is not directly related to CITI. There is no clear reference to where it occurred and there is an undefined term 17,000 in

the report. Similarly, we avoid survey reports, as they are not attributed to any particular failure case. The following is an example of such a survey report:

> *The Federal Trade Commission says that complaints of Internet-related identity theft more than tripled last year, to 2,352 last year from the year before. Jay Foley of the Identity Theft Resource Center says, "Online fraud is becoming as big an issue for eBay and AOL as security is for Microsoft." Typically, eBay covers buyers or sellers for up to $200 (or $500 for some listings) if an item is not delivered or is in bad condition, though there is a $25 processing fee. USA Today, 24 Oct 2003, "Internet fraud update", RISKS (22, 98)*

2.2.2 Fault Classification

Our next step was to categorize collected reports based on the nature of the failure. As explained, we used eight fault classes, most of which were derived from the taxonomies proposed by Kuhn [35]. The major advantage of Kuhn's approach is that the failure sources are orthogonal and as such, can be dealt with independently. A similar approach was used by Chillarege [48] for software defect classification. For instance, the risk of hardware faults can be minimized by using a redundant physical channel, redundant backup power supply, etc. [49], which are independent of other types of fault consideration. Similarly, for malicious logic faults, different kinds of protection techniques can be used. These include secure routing protocols, secure domain name systems, firewalls and anti-virus tools [43]. Sometimes failure management can be infrastructure dependent, which requires more specialized tools and techniques. For instance, air transportation services require specialized hardware and software tools to ensure their systems' reliability [50]. Table 2.1 shows the different fault classes used in our study.

Table 2.1: Fault Classes Related to Critical Infrastructures

Fault Name	Description
Hardware Fault	All fault classes that affect hardware.
Software Fault	Fault caused by an error in the software system.
Human Error	Non-deliberate faults introduced by a mistake.
Natural Fault	Physical faults that are caused by natural phenomena without human participation.
Overload	Service demand exceeds the designed system capacity.
Vandalism	Sabotage or other intentional damage.
Malicious Logic Fault	These include Trojan horses, logic or timing bombs, viruses, worms, zombies or DoS attack.
Authorization Violation	Attempt by an unauthorized person to access or damage network resources, but does not exclude the possibility of authorized users who are exceeding their rights. This also includes unauthorized sharing of digital contents, like audio, video or software.

Faults that trigger infrastructure failure can be mapped to different generic fault types. These generic faults belong to one of the eight fault classes mentioned above. As we analyzed failure cases and identified the root cause of each failure, we tried to identify type of fault source for each of these root causes. These generic fault sources are similar to Kuhn's decomposition of fault classes into finer detail, such as failure of cable component or power supplies, software version mismatch, etc. Table 2.2 lists the generic faults and the fault classes they belong to.

Table 2.2: Generic Faults Related to each Fault Class

Fault Class	Generic Fault
Hardware Fault	Physical link failure.
	Hardware design or implementation flaw.
	Failure due to external operating environment exceeds predefined limit.
	Device failure due to lack of backup power supply.
	Device failure due to lack of proper maintenance.
	Device failure, origin unknown.
Software Fault	System failure due to software glitch.
	Software design or implementation flaw.
	System failure due to software configuration or update error.
	System failure due to weak encryption algorithm.
Human Error	Usability factors not considered in system design.
	Inadequate safety measures.
	Careless mistake.
	Data entry error.
	Lack of proper user training or documentation.
Natural Fault	Natural calamity.
	Resource unusable due to natural cause.
Overload	User request failed due to inadequate system capacity.
Vandalism	Intentional breakage of physical links or devices.
Malicious Logic Fault	System failure due to malicious logic.
	Misguiding using malicious logic.
Authorization Violation	Unauthorized access by an outsider.
	Access right violation by authorized user.
	Unauthorized use of technology for malicious intention.
	Identity theft through authorization violation.
	Unauthorized capture or sharing of digital contents.

2.2.3 Feature Extraction

Once we categorized a failure report to a particular failure class, we extracted key features from each of these reports using a set of key attributes, which were sometimes judgmental. For example, degree of impact is a feature that is intended to capture the severity of a failure. Reading the failure case, we tried to understand how many people or systems were affected and how that number affected the overall functionality of an organization. The degree of impact assigned a rating “High” indicates a massive effect on the functionality of

CITI and other critical infrastructures. Similarly, ratings of "Medium" and "Low" indicate moderate and low impacts, respectively. Clemen et al. [51] show that in the absence of detailed information, a judgmental approach can be followed to predict risk. The following three examples illustrate the assignment of degree of impact by subjective judgment. These reports are identified using 'Report ID #' that refers to an ID in our database.

> Degree of impact - High (Report ID # 5) - *On November 19, 1994, Iowa City's US West telephone system shut down at about 3:30 p.m., local time, and service was gradually restored between 7:30 and 9:30 p.m, affecting about 60,000 people. Analysis showed that a new switching system had been installed in July 1994. In removing the old system, an electrical grounding cable had been inadvertently removed.*
>
> Degree of impact - Medium (Report ID # 3) - *MCI's inbound Internet gateways were saturated during July 1994, resulting in days of delay in delivering e-mail to MCI customers. A fix was considered to be months in the offing.*
>
> Degree of impact - Low (Report ID # 8) - *A software glitch on March 10, 1995, caused Prodigy's e-mail system to send 473 e-mail messages to incorrect recipients and to lose 4,901other messages. The system had to be shut down for five hours.*

In the first report, the telephone services to 60,000 people were affected, which was an important consideration in designating it as a high impact case. In the second report, even though email service was delayed; other means of communication were available, therefore it was assumed gateway saturation was a moderate inconvenience for the MCI customers and the report was designated as a medium impact case. In the third report, it seems that after five hours the email service was fixed and misdirected emails were re-dispatched to the recipients. Since people did not check their email very often and the number of recipients was only 473, we concluded that this failure modestly affected the users, and assigned it a low degree of impact.

Another key feature of a failure report is the report accuracy, where we assigned an accuracy rating on a scale of 10 based on the source type. For each of these reports, the information source was given. If the information was released from an official source and had other supporting references for validation, we assigned it 9 or 10 points. If it was from an official source but no further details were given, we assigned it 7 or 8 points. All newspaper reports were given 5 or 6 points. Reports from individuals, which were difficult to verify, were normally given less than 5 points. Higher ratings were given to reports of a particular class if the reports fulfilled most of our additional criteria. For instance,

if a newspaper report had most of the required information, such as severity, duration, financial impact, description of fault origin, etc., then it was given 6 points; otherwise, it was given 5 points. In the future, we would like to use this accuracy rating to conduct a reliability analysis of the collected cases. However, the report accuracy rating in this chapter is different from the Bedell's Index values discussed in Chapter 3. Table 2.3 lists the extracted key features of a failure report and their meanings.

Table 2.3: Extracted Features and their Meaning

Fault Name	Meaning Description
Title	Title of the report. Most often this is the same as the original report.
Date	Date of the failure report.
Country	Country where fault incident originated. For global fault, it is World.
Impact Scale	Size of area affected. It could be an Organization, a City, a Region (a big part of the country), a Country, a Continent, or the whole World.
Degree of Impact	Failure impact. Could be High, Medium or Low.
Simulation	Indicates if the fault conditions can be simulated within a lab environment using I2Sim [20] critical infrastructure simulator.
Fault Intent	Fault could be intentional, due to deliberate and malicious attempts by any individual or groups, or unintentional due to human error or system flaw.
Duration	Time from the start of the fault to its full recovery.
Financial Impact	Amount of financial loss in Million USD.
Public Safety	Any public safety concern associated with a particular fault incident, such as failure of 911 service, medical emergency service, fire rescue service, or police service.
Affected Sites	Number of sites or locations affected by a particular fault incident.
Description	Description of the failure (report text).
Report Source	Reference of the report collected from RISKS forum and referred to as RISKS (i, j), where i is the volume number and j is the issue number within the volume.
Report Accuracy	Based on the source type, an accuracy rating on a scale of 10.
Fault Class	Fault type is one of the eight types mentioned in Table 2.1.
Generic Fault	A qualitative assessment of the origin of the fault.
Source Infrastructure	One of the critical infrastructures discussed in Section 2.2.1.
Affected Infrastructures	One of the critical infrastructures discussed in Section 2.2.1.
Affected Industry Sectors	Description of the industry sectors affected by the failure.
Comment	Comments on specific interesting aspects of these faults.

Many of these features are intended to capture the extent of the failures, their impact in spatial and temporal dimensions, their effect on public safety, and the propagation of the

failures from CITI to other critical infrastructures and vise versa. Table 2.4 groups these features into different categories according to the intended use.

Table 2.4: Features that Capture Different Failure Dimensions

Analysis Dimension	Feature Names
Extent of failure	Fault class, Degree of impact, Fault intent, Fault type
Impact (spatial)	Country, Impact scale, Affected sites
Impact (temporal)	Date, Duration
Public safety	Public safety
Failure propagation	Source infrastructure, Affected infrastructures

2.2.4 Data Analysis

Many of the public domain failure reports we collected had some missing attributes. For example, the duration of a failure and the number of sites affected by failures were not clearly specified for almost half of the cases. The financial impact of failures is mentioned in fewer than 10% of the cases. As a result, we could not use concepts like "Customer Minutes" (product of average number of customers affected and average outage duration) as used by Kuhn [35] to measure the severity of failures in PSTN networks. Use of such concepts is possible for FCC reports, as each FCC report has to include date, time, failure duration and the number of affected customers [35]. Unlike FCC, however, our failure reports did not have such uniformity and universal impact dimensions. To compensate for this, we used a frequency−based approach to quantify results from the extracted features of the failure database (Section 2.3). This way, we tried to determine the most likely cause of infrastructure failure, the types of localities affected, and the implications for public safety. As mentioned before, due to the absence of clearly specified values for many key attributes, we had to use our own judgment to estimate some of the values of these key attributes. In doing so, we were limited by the description of the data, and there were no mechanisms for obtaining further detail.

2.3 Failure Database

The collected cases and their extracted features were compiled in a MS Excel database. A sample record from this database is shown below (Figure 2.1). Each record represents a

5	Ground-cable removal blows Iowa City phone system upgrade			
Date	**Country**	**Impact Scale**	**Deg of Impact**	**Simulation**
11/19/1994	USA	City	High	Unsure
Fault Intent	**Duration**	**Financial Impact**	**Public Safety**	**Affected Sites**
Unintentional	6 hours	Unknown	Yes	Unknown
On November 19, 1994, Iowa City's US West telephone system shut down at about 3:30 p.m., local time, and service was gradually restored between 7:30 and 9:30 p.m, affecting about 60,000 people. Analysis showed that a new switching system had been installed in July 1994. In removing the old system, an electrical grounding cable had been inadvertently removed.				
Report Source	Iowa City Press Citizen, November 22, 1994; see discussion by Douglas W. Jones, RISKS (16, 58)			
Report Accuracy	6			
Fault Class	Human Error			
Generic Fault	Inadequate safety measures.			
Source Infrastructure	Electrical Power System			
Affected Infrastructures	Telecommunications Infrastructure			
Affected Industry Sectors	All kinds of industries in Iowa City			
Comment	Fault in electrical system due to human error.			

Figure 2.1: A Sample Database Record

single row in the MS Excel spreadsheet. The analysis performed on these records was done in another sheet within the same MS Excel file.

Each record in this database has a report ID. A report ID is a sequential number assigned based on the incidence date. Other fields have their own set of valid values. Table 2.5 summarizes acceptable values for each of these attributes.

Table 2.5: Acceptable Values for Failure Database

Field Name	Legal Values
Report ID #	A sequential number assigned based on the report's date.
Title	Text String
Date	MM/DD/YYYY
Country	Country Name / World
Impact Scale	Organization / City / Region / Country / Continent / World
Degree of Impact	High / Medium / Low
Simulation	Yes / No / Unsure
Fault Intent	Intentional / Unintentional / Unknown
Fault Class	One of the eight fault classes as (Table 2.1)
Generic Fault	Origin of the fault that belongs to one fault class (Table 2.2)
Duration	Hour
Financial Impact	Million USD
Public Safety	Yes / No / Unknown
Affected Sites	Number of sites or Unknown
Description	Text String
Report Source	Text String
Report Accuracy	Accuracy rating number
Fault Origin	Text String
Source Infrastructure	One of the infrastructures from Section 2.2.1
Affected Infrastructures	One of the infrastructures from Section 2.2.1
Affected Industry Sectors	Text String
Comment	Text String

2.4 Results

The failure data we collected from the RISKS forum came from two different sources. The first type of data are those events that received much attention and were conveyed to the readers through global news distribution networks, such as the Associated Press, Reuters, etc. The second type of data are those events that did not receive similar attention but were made public through regional newspapers, radio or television stations, or different organizations' websites. We found that in 20% of cases, the reports we collected were broadcast through large news networks. The other 80% came from national or local news sources. These sources included major national newspapers like Washington Post, New York Times, USA Today, Guardian, Toronto Star, Vancouver Sun, New Zealand Herald,

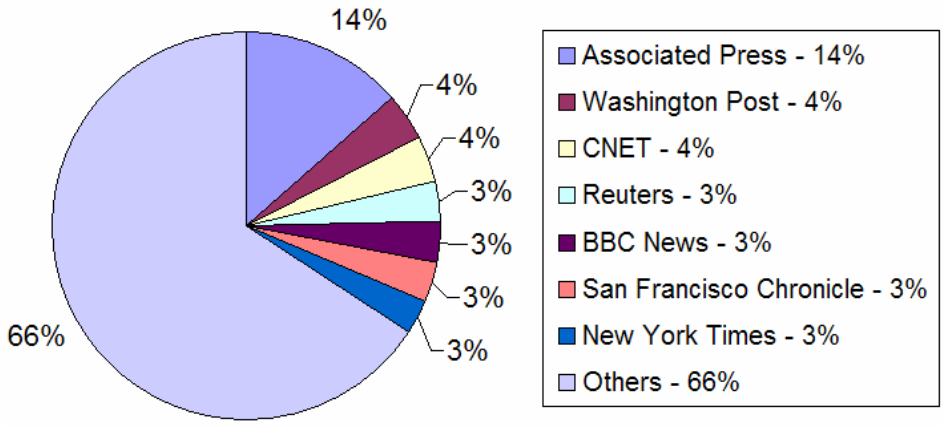


Figure 2.2: Distribution of Report Sources

etc. These reports were forwarded to the RISKS forum by forum users. Figure 2.2 shows different report sources based on their contributing ratio. In this figure, the category 'Others' (66%) includes all sources that individually contributed 2% or less. It appears that most of the reports in our study (about 60%) fall into this category and are from small, local sources. Because our study draws upon a wide range of sources, it can be considered to be broadly representative of actual failure scenarios.

We analyzed 347 cases that occurred in a 12 year period (1994 to 2005). Figure 2.3 shows that the frequency of reporting infrastructure failure to the RISKS forum changed during this period. The trend is nearly linear, except for the year 2003. The linear increase of CITI and other critical infrastructure failure reports may imply that these infrastructures are becoming increasingly dependent on CITI services. However, the sudden rise of failure cases during 2003 was due to the significant escalation of malicious attacks against IT infrastructure (also see Figure 2.10). These attacks included different kinds of worm attacks (Slammer, MSBlaster, Nachi) and DoS attacks. We also observed a significant number of Authorization Violation cases (22 reports) in 2003. These cases included major identity thefts (Report ID # 172, 183), unauthorized digital content sharing (Report ID# 189, 208), change of stock market index using malicious techniques (Report ID# 199), the presence of a fake US government organization on the Internet (Report ID# 167), online auction frauds (Report ID# 207), and similar other cases. The trend was worldwide, and was even visible in remote parts of the world (Report ID# 211, 217). One explanation for the large number of failures during this period is that corporate cyber security mechanisms were not mature enough to compete with the power and availability of automated hacking tools.

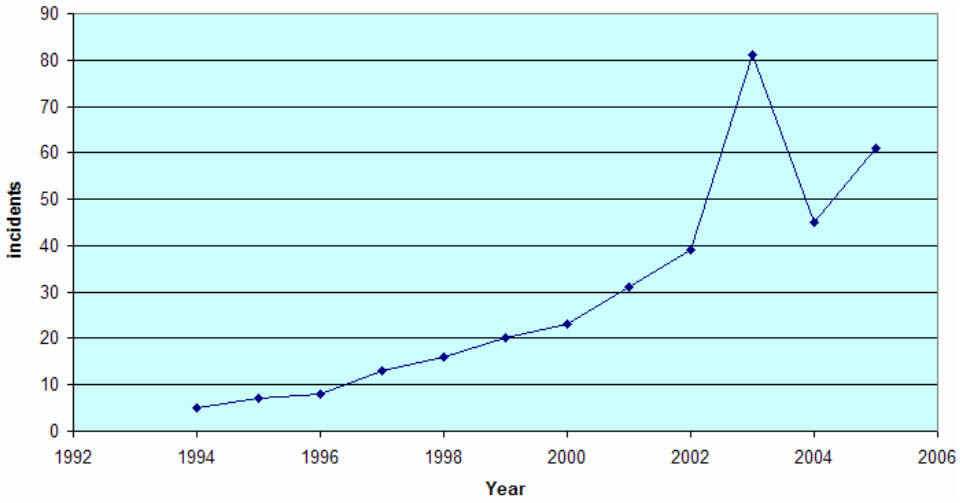


Figure 2.3: Reported Failures over Time

The increase of malicious attacks during 2003 can be cross-checked by performing a trend analysis on other bounded repositories similar to the RISKS forum, unlike the Internet, which is unbounded.

2.4.1 Failure by Category

Figure 2.4 shows percentages of failures by fault class. It is interesting to note that software-related failures constituted more than 65% of all reported failures, if we include malicious logic and authorization violation within this group.

Each fault class can be further categorized by their generic type. Figure 2.5 shows software and hardware faults related to all infrastructures classified according to their generic type. The most common cause of software failures is software glitch (45%), followed by software design or implementation flaw (29%). Software glitch is a generic term we use to indicate software failure due to unknown reasons; most often they are related to design or implementation flaws. The high percentage of failures due to software design and implementation defects suggests that better software engineering practices are essential to increase the CITI infrastructure safety and reliability. Similar results were obtained for hardware failures, where the most common cause is device failure due to unknown origin (45%).

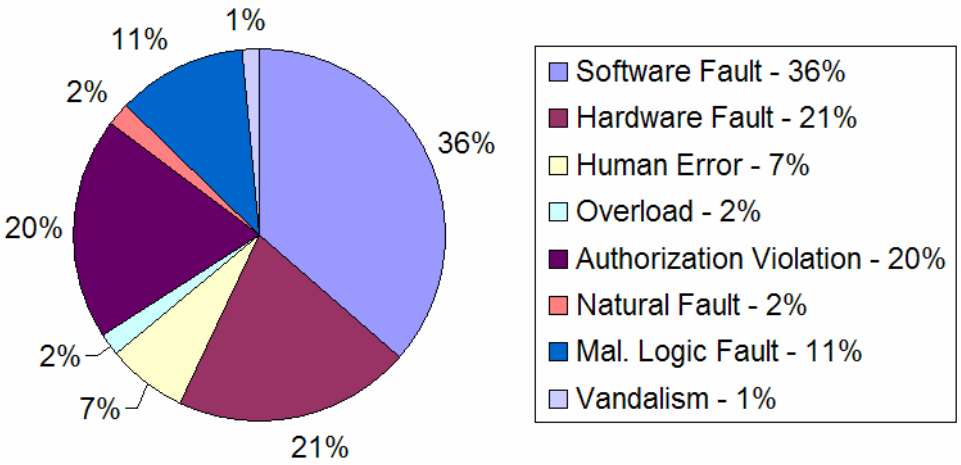


Figure 2.4: Faults that Lead to all Infrastructure Failure

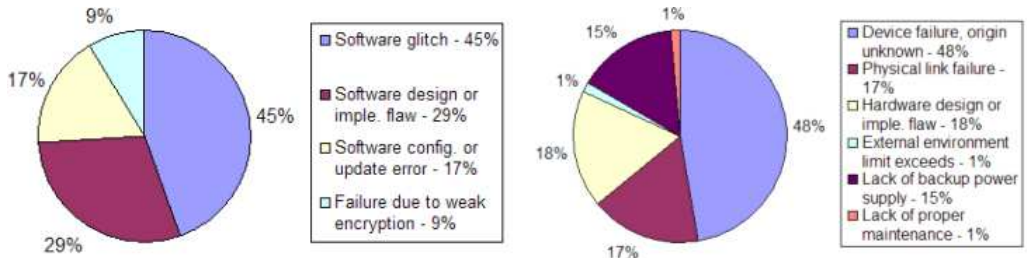


Figure 2.5: Software and Hardware Faults are further Categorized by Generic Type.

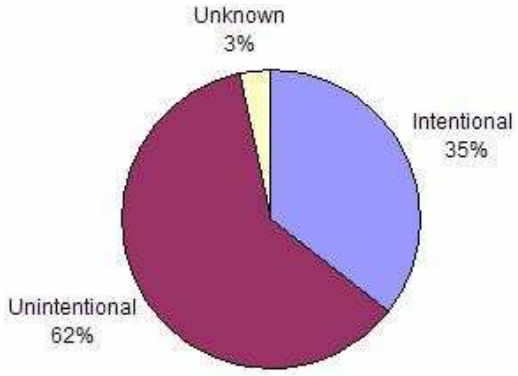


Figure 2.6: Failure Type Distribution for all Infrastructures

2.4.2 Impact of Failure

The failure reports revealed that the root causes of most of the CITI and related infrastructure failures were unintentional or accidental (Figure 2.6). Critical infrastructures are the lifelines of modern societies. As such, even though their root causes may be unintentional, the impact of failures is significant (Figure 2.7). Accidental causes include hardware or software faults, configuration problems, human error, etc. In contrast, malicious logic faults, authorization violation attempts, and vandalism account for fewer than 33% of the cases. This situation illustrates the subtle fact that system reliability deserves more attention than it is getting now in relation to system security. An example of this discrepancy is the air transportation industry. Of the 27 air transportation failure cases we reported, 26 were due to various non-malicious hardware and software faults in the air traffic control system. One such case (Report ID # 83) says:

> *On 17 Jun 2000, thousands of would-be passengers were stranded when the main air-traffic control computer collapsed. The National Air Traffic Services computer was fixed later in the day, but the resulting congestion caused many people to spend the night at airports around the UK, and many flights were cancelled the next day as well. Heathrow and Gatwick were hardest hit, although other UK airports experienced severe delays. This was the second time in a week that the computer system had failed.*

2.4.3 Public Safety Concerns

Public safety is a major concern for critical infrastructure failure. Although our study shows that a majority of failure cases did not have public safety implications (Figure 2.8),

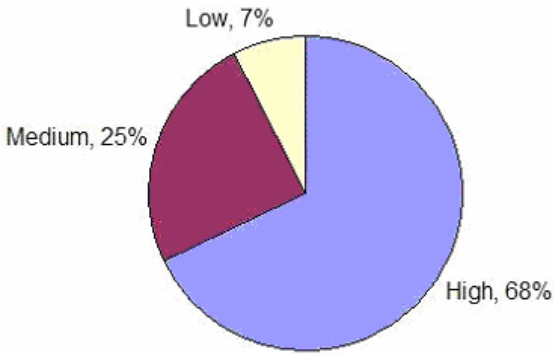


Figure 2.7: Infrastructure Failure Impact Distribution

we observed that in nearly 20% of the cases, the failure affected public safety to some degree. Many times, these failures were due to improper design or set up of public safety-related devices, lack of backup power supply, etc. Since infrastructure failures are on the rise, there are increasing concerns for public safety. An example can be given from the following 911 systems report (Report ID # 230):

> *Houston has deployed a new 911 emergency response system which has had a number of failures since it went "live" a week ago. Pictures of the new facility look somewhat like Mission Control - large consoles with multiple displays in front of each operator. It sure looks nice, but the system does not appear to work reliably. The latest incident occurred during the day when technicians were working on the link between the computers and units within the cars. To quote: When the system started slowing, technicians reverted to the backup, which crashed within minutes. From 9:50 a.m. to 10:30 a.m., dispatchers resorted to dispatching by radio instead of by computer. Without the computer's locator system, they frequently had to ask emergency workers to volunteer for individual assignments rather than assigning them to calls. Another notable quote is But city officials say the only way to test the system was by going "live."*

2.4.4 Change in Degree of Impact over Time

Figure 2.9 shows that the frequency of high impact infrastructure failures is on the rise. Figure 2.10 shows that many of these failures are due to malicious intent. Examples of origin of failure include DoS attack against the Internet infrastructure (Report ID # 156), worm or virus attack (Report # ID 163, 166) and identity theft (Report # ID 172). From 2001 on, failures due to intentional causes is on the rise. This change of trend is

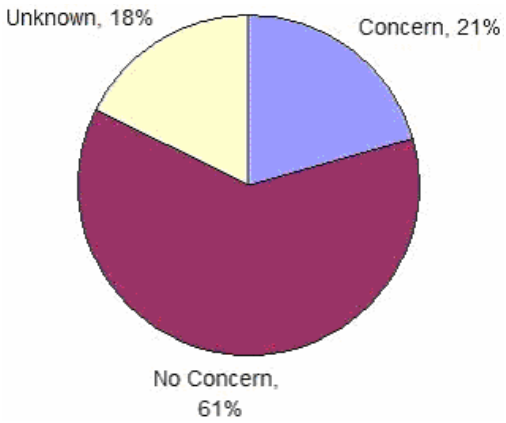


Figure 2.8: Public Safety Impact Distribution.

more apparent in recent years. For instance, during 2005, we had 30 intentional failure reports, most of which were related to identity theft, system hacking, phishing, and spamming. These kinds of cyber attacks have become the major form of threat against critical infrastructures. Other contributing factors include the emergence of automated and high-speed worms (e.g., Code Red), increasing deployment of off-the-self software systems for critical infrastructure management (e.g., MS SQL Server), and inadequate expert manpower to manage more complex interconnected infrastructure systems. This increasing dependency on IT infrastructure is making other critical infrastructures ever more vulnerable. There is no sign that this trend will change in the near future. The following example shows how a healthcare system can be affected due to its dependency on computerized prescription systems that depend on electrical power systems (Report ID # 186):

> *Thousands of patients could have received the wrong prescription drugs after a power outage at Kaiser Permanente's computer center in Southern California knocked the pharmacy's labeling system out of sync – printing the wrong labels on filled prescriptions. There were no reports yet of patients suffering from adverse reactions. About 4,700 patients from Fresno to the Oregon border were affected, including those ordering prescriptions by telephone. After the error was discovered on 14 Mar 2003, hospital officials attempted to contact the affected patients, although by 17 Mar, 152 remained uncontacted – including those for whom they had only PO-box addresses.*

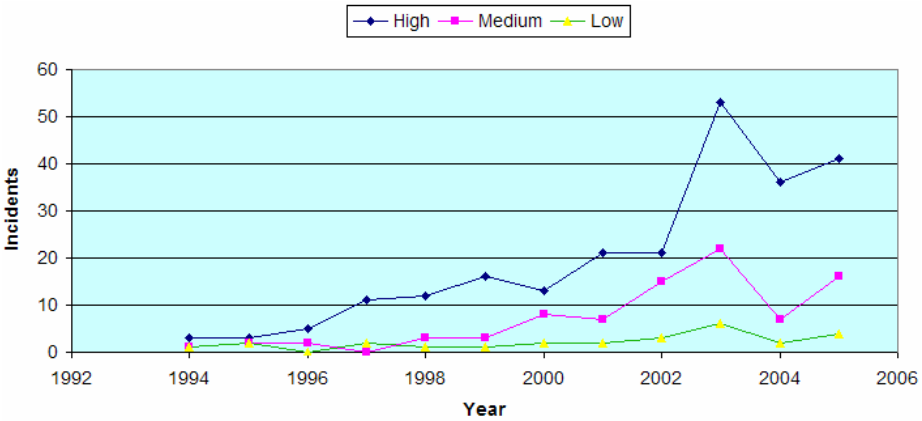


Figure 2.9: Change in Degree of Impact over Time

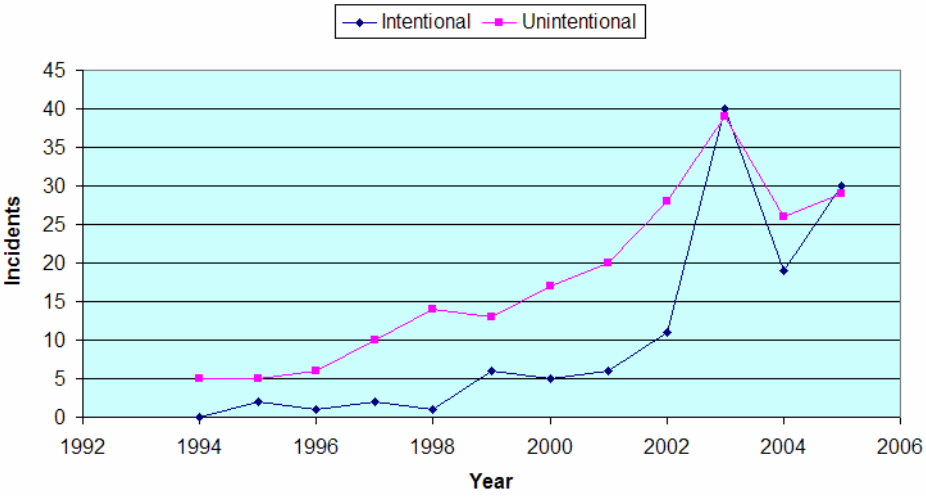


Figure 2.10: Change in Intentional and Unintentional Failure over Time.

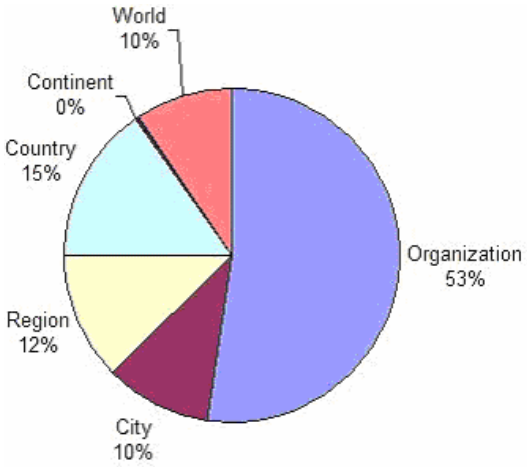
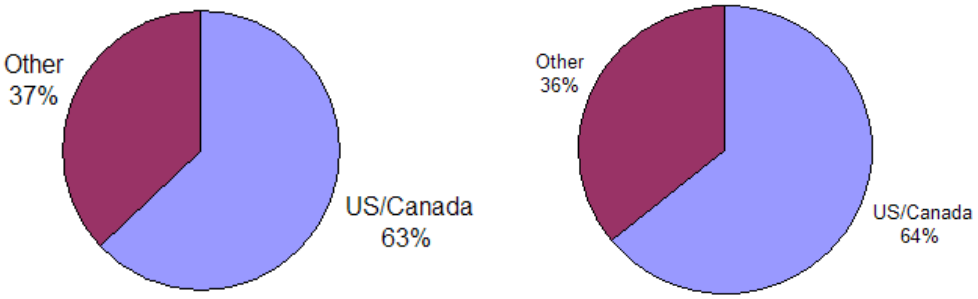


Figure 2.11: Localities Affected by Infrastructure Failures.

2.4.5 Localities Affected by CITI Failures

Figure 2.11 shows that almost half of the CITI and connected infrastructure failures studied propagated beyond organizational boundaries (47%). Crossing the national boundary was relatively rare, unless an attack was targeted internationally. Figure 2.12 shows that North America (US/Canada) was the most vulnerable region for CITI infrastructure failure (63%) in our study. One possible explanation is that this region has a much higher proportion of computer use than any other part of the world. Figure 2.12-a includes worldwide failure cases (e.g., worm attack), whereas the Figure 2.12-b excludes those cases. In both figures, most of the reported failures (above 60%) took place in North America (US/Canada). Inclusion or exclusion of worldwide failure does not significantly change these patterns.



(a) Includes Worldwide Failures (b) Excludes Worldwide Failures

Figure 2.12: Failure Location US and Canada.

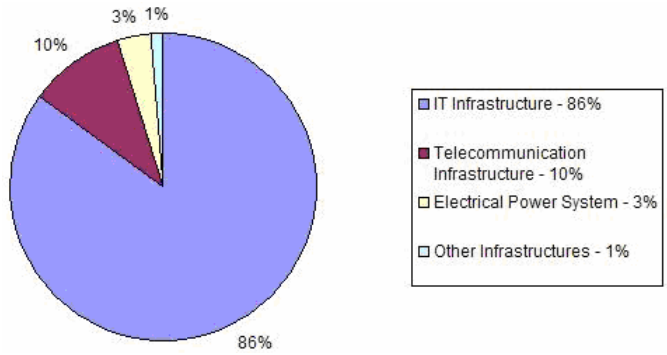


Figure 2.13: Source of Failures Affecting CITI.

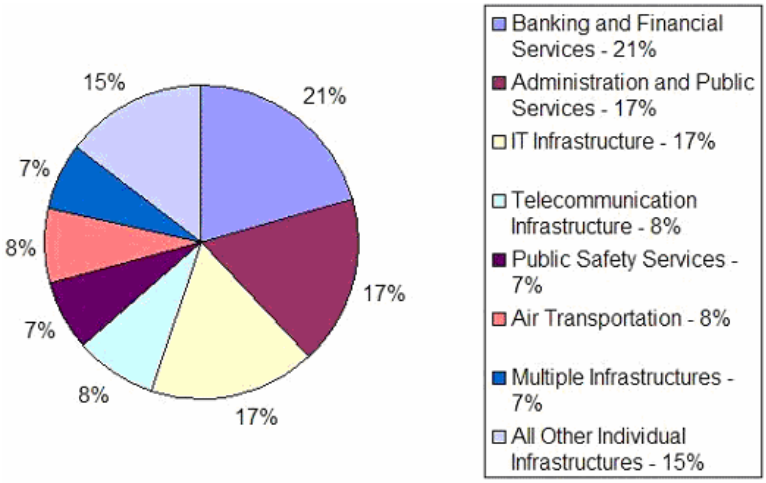


Figure 2.14: Infrastructures Affected due to CITI Failures.

2.4.6 Interdependencies among CITI and other Infrastructures

Figure 2.13 shows that in most of the cases studied, CITI failures originated from within the CITI infrastructure. The role of other infrastructures was relatively minor. Figure 2.14 shows that most of the CITI failures affected banking and financial services, administration and public services, and the CITI infrastructure itself (IT and telecommunications).

Results of a more detailed analysis of the failures of the first four infrastructures shown in Figure 2.14 are presented in Figures 2.15 to 2.18. Figure 2.15 shows that software systems were the most vulnerable points for banking and financial services and that a large percentage of these failures had malicious origins (45%). Better software engineering practices and incorporation of adequate security measures can improve the reliability of banking and financial services.

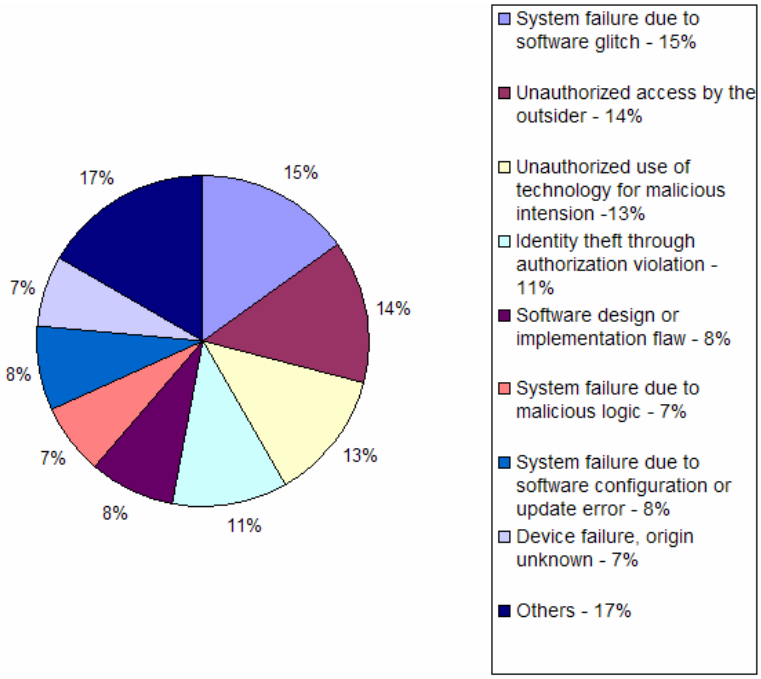


Figure 2.15: Generic Faults that led to Banking and Financial Services Failures.

Administration and public services infrastructure includes large government organizations, universities, and other educational institutions. Figure 2.16 shows that these organizations were also susceptible to software-related failures in the reports we studied, and that large percentage of them had malicious origins (37%).

Figures 2.17 and 2.18 show most of the CITI failures originated within CITI. Detailed analysis reveals that IT infrastructure in our study was largely vulnerable to software-related failures, whereas telecommunications infrastructure was vulnerable to different hardware-related failures. Therefore, improved software and hardware related techniques in CITI infrastructure design, implementation, and management will ensure a greater stability in its operation, which will eventually improve the reliability of other connected infrastructures.

2.5 Chapter Summary

Our society relies upon continued services from interdependent critical infrastructures to function. CITI failures are particularly pervasive in their penetration of all infrastructures, and can have a very large impact on the workings of society. Understanding and classifying

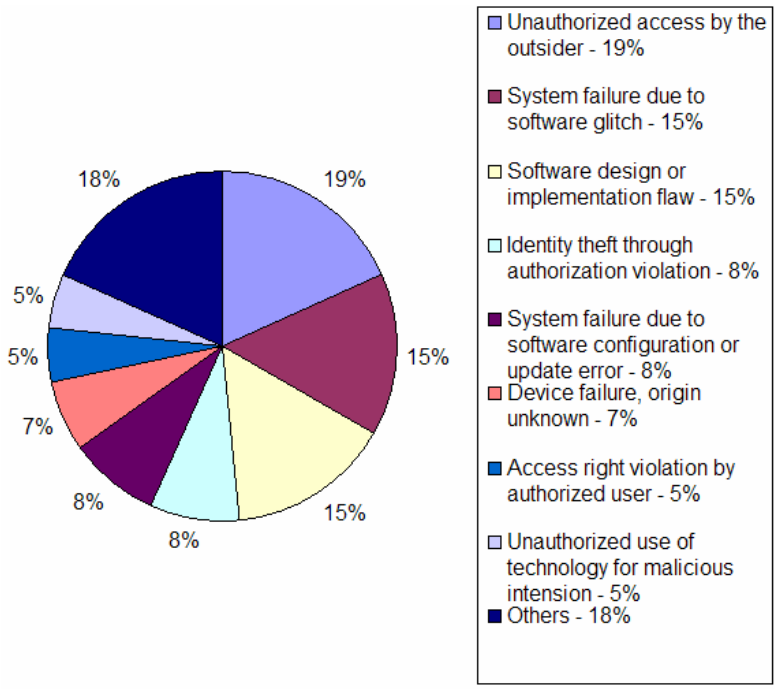


Figure 2.16: Generic Faults that led to Administration and Public Services Failures.

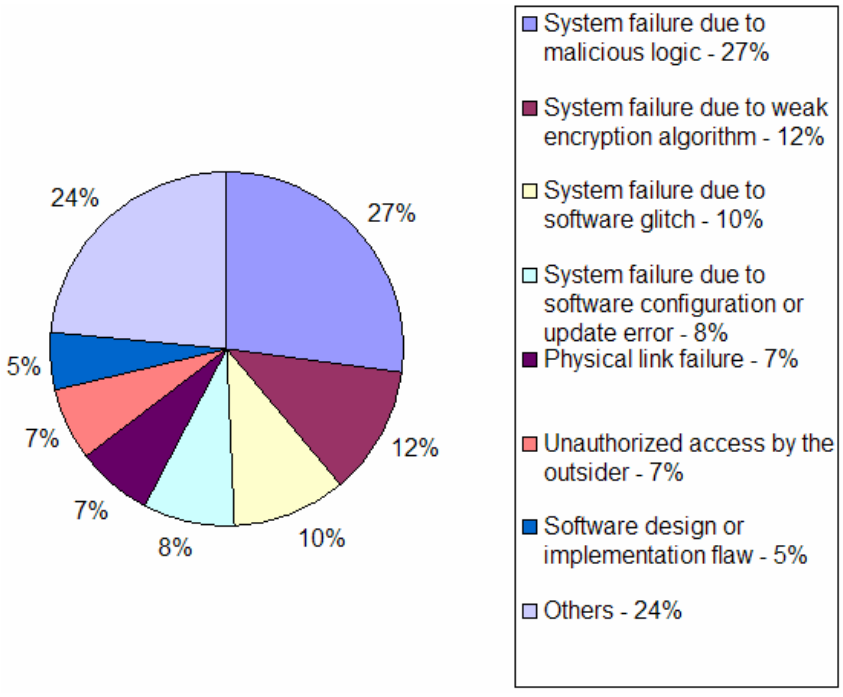


Figure 2.17: Generic Faults that led to IT Infrastructure Failure.

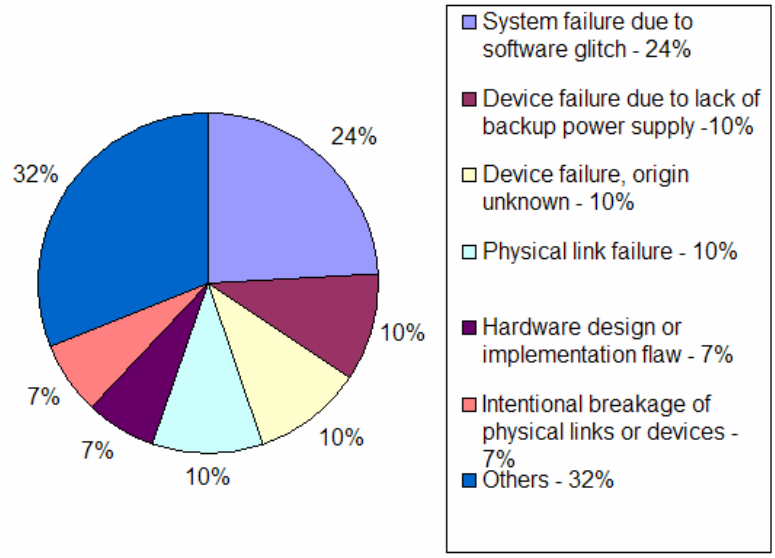


Figure 2.18: Generic Faults that led to Telecommunications Infrastructure Failure.

patterns of CITI failures is an important step towards quantifying interdependency analysis in CITI-dependent infrastructure systems and identifying preparedness and mitigation strategies to ameliorate the impact of system-wide failures. In this study, we have used public domain data over 12 years to understand CITI interdependencies. To our knowledge, this is the first attempt of this type of analysis in this area (using either public or private data sources). We identified infrastructure failure patterns, propagation, impacts on public life, and historical trends. We have developed a CITI failure database for our Infrastructures Interdependencies Coordination (I2C) research group at UBC in setting up realistic test-case scenarios involving CITI failures during large-scale system failure situations [10].

Chapter 3

CITI Interdependency Estimates

In this chapter, we present a set of quantitative estimates of cyber interdependencies for different critical infrastructure networks. The results proposed here will be useful for computerized simulation of critical infrastructure networks. Pederson et al.'s survey [13] shows that paradigms of present-day critical infrastructure interdependency modelling can be classified into two major categories: a) based on a qualitative approach, and b) based on a quantitative approach. Rinaldi et al. [1] propose the idea of qualitative modelling of critical infrastructure interdependencies. These authors view infrastructures as a complex collection of interacting components, where coordination and decision-making can be most appropriately represented by complex adaptive systems (CAS). Rinaldi et al. [1] propose that interdependencies can be understood as "emerging behaviors" of the interacting components, which are not predictable by the knowledge of any single entity. Most of the agent-based infrastructure simulation frameworks adhere to this paradigm of modelling [22].

The quantitative approach has been adopted by members of the systems engineering discipline. This is a causality-based approach, where physical infrastructures are viewed as system of systems [52]. Interdependencies among different system components are captured using mathematical functions. Marti et al. [10] follow this quantitative paradigm and propose the Cell-Channel model, as a generic approach for modelling interdependencies among critical infrastructures. According to this model, conceptual entities of different infrastructures are mapped to a single equivalent semantic, on the assumption that all infrastructures can be represented by two kinds of components, cells and channels. Based on the type of infrastructure, different types of service tokens are delivered through

the channels from one cell to another. Interdependencies among different infrastructure elements are represented using nonlinear functions, which can also be described in the Human Readable Table (HRT) format [10, 20, 53].

Our present work follows a quantitative modelling paradigm, where infrastructure entities are represented using the Cell-Channel [10] model. We have presented several empirical functions to describe cyber interdependencies for different critical infrastructure entities. The approach to develop these empirical functions is based on ranking the contribution of the different CITI services to the infrastructure entity's output. Our study uses Bedell's method [28, 54] to rank the different CITI services according to their importance and effectiveness. In this study, we have used findings from contemporary research, interviews of some of the infrastructure operators and public domain failure reports that we have collected from our previous research (see Chapter 2). As such, this chapter also summarizes important CITI technologies used in different critical infrastructures.

3.1 Related Work

The impact of information systems on an organization has been studied for more than the last four decades [54–57]. This significant interest originates from the fact that CITI investment constitutes a large portion of any organization's capital expenditure. As such, one of the major objectives of these studies has been to develop methods for prioritizing CITI investment proposals. Other related objectives have been to develop techniques to control IT expenditures. Despite considerable research in this field, information system evaluation remains a difficult problem [54, 57]. We may not be able to give comprehensive overview of this widely studied subject in one section, but we have chosen to discuss some of the important surveys that were done at different points in time to give a snapshot of this research field. At the end of this section, we provide some rationals about the methodology (Bedell's method) we adopted for our analysis. As a cursory note we would like to mention that, even though information systems evaluation is a mature field, the determination of the changes in any infrastructure's (organization) output for different level of CITI services through simulation is a new approach [1, 13].

Peter Sassone (1988) has compiled a survey [55] of evaluation methodologies for the usefulness of information systems in large organizations. The methods discussed in Sassone's survey were decision analysis, cost displacement/ avoidance, structural models,

cost effectiveness analysis, break-even analysis, subjective analysis, time-savings times-salary, and the work-value model. Cost justification was the main criteria for Sassone's survey, which identified the above methods to estimate the utility of information systems. Renkema and Berghout [54] have published a survey (1996) on the methods used to evaluate information system investment proposals in an organization. Their study lists that more than 65 methods available for such analysis. Renkema and Berghout grouped these methods into four classes, based on the similarity of their features. These four classes are financial approach, multi-criteria approach, ratio approach and portfolio approach. Important methods belonging to these four categories were compared and their merits summarized. Recent studies conducted by Smithson and Hirschheim (1998) [57] and Chou et al. (2006) [56] show that newer evaluation techniques are likely to consider influences from social, political and behavioral aspects related to the information systems of an organization. This wider scope of evaluation criteria is guided by the increasing complexity of information systems. Conventional cost-benefit analysis or technology based considerations do not adequately capture the impact of present day more complex information systems. However, it is also noted [57] that these recently proposed techniques are largely limited to academic literature; organizations and practitioners have been slow to adopt them in business practice [57].

Organizations' productivity (output) is strongly reliant on information systems' contribution. Hitt and Brynjolfsson's (1996) [58] investigation of 370 large companies demonstrates this fact. In our present work, we have focused on the impact of communication and information systems on the output of different types of infrastructures (organizations). We have selected Bedell's [28, 54] method (1984), which we find to be the most suitable to develop quantitative relationships of CITI interdependencies. However, we have extended some of the assumptions and procedures of Bedell's method to address changes that have taken place in the CITI systems over last two decades. Bedell's method has the advantage that it can capture the relationship between information system services (input) and the organization's effectiveness (output). This is important in order to relate CITI systems to the Cell-Channel model interdependencies simulation model, which is based on the Loentief input-output model [10, 20]. Another advantage of Bedell's method is that it is well documented, well structured and can be applied at both the system level and the organization (infrastructure) level. There are reports that Bedell's method has been used in practice in the recent years [59, 60].

3.2 Overview of Bedell's Method

Bedell's method [28] is a set of techniques for the evaluation of information systems. These include techniques for IT effectiveness measurement, data-sharing strategy planning, development resource allocation and project cost management. Renkema and Berghout [54] have classified Bedell's method as a portfolio methods, where investment projects are plotted against several evaluation criteria to help decision-making. Among Bedell's method, effectiveness measurement is important for our work. Below we provide an overview of this technique and propose extended assumptions needed to make it useful for interdependency analysis between CITI and other critical infrastructures.

The central concept of Bedell's effectiveness estimate is the assignment of numeric ranks to the information systems based on their importance and effectiveness in an organization. An information system is important if it supports important activities in an organization. If the design and implementation of the information system is good, its effectiveness (or quality) to serve the intended activity is high. Quantification is conducted through the calculation of several effectiveness and importance indices to estimate the following characteristics:

- Importance of an activity
- Importance of the information systems in supporting the activity
- Quality of the information systems in terms of effectiveness to support the activity.

We have calculated importance and effectiveness indices at the organization level. We have considered an organization as a physical representation of an infrastructure. For instance, a hospital is an infrastructure entity within the healthcare sector. According to Bedell's method [28], the following five indices are necessary to calculate the effectiveness of the information systems (EIO) at the organization level:

- ISA - how Important a particular System is to the Activity it was built to support
- ESA - how Effective (quality) a particular System is to the Activity it was built to support
- IAO - how Important the Activity in question is to the Organization
- ISO - how Important a particular System is to the Organization as a whole
- EIO - how Effective (quality) Information systems are to support the entire Organization

Among the five indices, the first two are system level and the last three are organization level. These are ranked between 10 to 0, from the highest to the lowest. Bedell's method

assigns specific index values to the first three based on their importance or effectiveness, as shown in Table 3.1:

Table 3.1: Bedell's Index Values

ESA Index	ISA Index	IAO Index
10 - Highly Effective	10 - Strategic Factor	10 - Critically Strategic Activity
5 - Moderately Effective	5 - Major Support Factor	8 - Strategic Activity
1 - Ineffective	1 - Minor Support Factor	6 - Contributory Activity
0 - No Support	0 - Not Useful	4 - Support Activity
		2 - Overhead Activity
		0 - Detrimental Activity

The last two indices are calculated from the first three indices. The ISO index is defined as $ISA_i \times IAO_i$, where $i = 1, .., n$. Here n denote the number of information system within the organization. This relation implies that a system becomes more important to the organization as it becomes more important to the activity it is supporting. Similarly, the EIO index is defined as $\sum(ESA_i \times ISO_i)/(\sum ISO_i)$. The idea here is to weight the effectiveness of each of the system activities by the system's importance to the organization. The merit of Bedell's method is that it provides a way to quantify the effectiveness of information systems of an organization in the context of the organization's own strategic objectives. The usefulness of these indices has been tested in both large and small organizations and was found to be high [28].

There are important issues in our use of Bedell's method. First, Bedell's method itself is related to information systems. The information system's relationship to the communications infrastructure was not explicitly addressed in the original model. However, both communication and information technology have changed significantly in the last two decades [57]. Current CITI systems are deeply integrated with each other and are diffused widely throughout an organization. We consider them as a single infrastructure service (CITI) within the organization's overall operation, rather than as isolated information systems supporting individual business services. Second, we assume information system's effectiveness (EIO Index) is directly related to the particular infrastructure's output. This assumption is justified because for many critical infrastructures the CITI has become the primary mode of operation, rather than just a support mechanism. For others, it is an important contributing factor. Internet banking and air traffic control systems are typical examples. Our study on critical infrastructure failure reports [12] demonstrates this fact

as well. As a result, in this study, even though we maintain the same definition and numeric values for Bedell's indices (Table 3.1), the scope of the information system is more inclusive. It includes communication systems and direct dependency between the CITI and the infrastructure cell's output.

One related question may arise at this point that, while for most physical infrastructures we can determine input-output relationships directly through measurement or interviews, why do we need some special methods (e.g. Bedell's method) to estimate CITI related interdependencies? The answer is simple; unlike physical infrastructure inputs (e.g., electricity, water supply), the contribution of the CITI on infrastructure' output is not directly measurable, as the CITI inputs are distributed in different layers, and their overall contribution to an organization's output is not always identifiable from the measurement of any physical quantity (e.g. data network packet flow, application software up-time). Even if we could measure these quantities for a large number of cases and could build some interdependency estimates based on those measurements, such anticipated estimates might possibly be captured more easily from emerging behavior in agent-based systems. Since we do not have enough data to estimate such interdependency relationships for the systems engineering approach and as it is not practically feasible to capture most of these interdependency data in a lab environment, Bedell's method gives us a way to move forward. We can justify the importance and effectiveness of a CITI system in an organization from our experience. These experiences are shared in the interviews and in the research reports. Infrastructure failure data also show the usefulness of these systems. Bedell's method gives us a way to translate this accumulated experience (contribution of CITI systems) to a number, which enables us to build an interdependency relationship.

3.3 The CITI Interdependency Function

To quantify the interdependency of an infrastructure cell on CITI services, we introduced the concept of the CITI interdependency function. The CITI interdependency function represents a linear model of dependency of a cell's output to the CITI services input. The effectiveness of CITI's contribution to the cell's output is determined by the Bedell's organization (at cell-level) effectiveness index (EIO). In the CITI interdependency estimate, we normalized EIO index by a factor of 10. The interdependency function is formally defined as follows:

Definition: When output of an infrastructure cell depends on information flow from the CITI services, the cell's output can be expressed in the following relation:

$$\begin{aligned} & f(x) = (1 - \tfrac{EIO}{10}) + \tfrac{EIO}{10}x, \quad 0 \leq x \leq 1 \\ & \text{where,} \\ & x \text{ is CITI service input} \\ & f(x) \text{ is infrastructure cell's output} \end{aligned} \tag{3.1}$$

The first part of the RHS of Equation 3.1 represents the output of the infrastructure cell which is independent of the CITI services input. The second part represents cell output that is dependent on the amount of information flow 'x' from the CITI services. This is to note that CITI input may be one of the many inputs to the cell. The other inputs could be electricity, water supply, etc. Figure 3.1 shows a graphical representation of the CITI interdependency function where the EIO index is 9. The availability of CITI services is represented on the X axis and infrastructure's output is represented on the Y axis. The EIO index 9 implies, 90% of infrastructure output is directly related to the CITI service input. The EIO index also represents the slope of the CITI interdependency function, where a higher slope implies a deeper coupling between CITI and infrastructure cell's output.

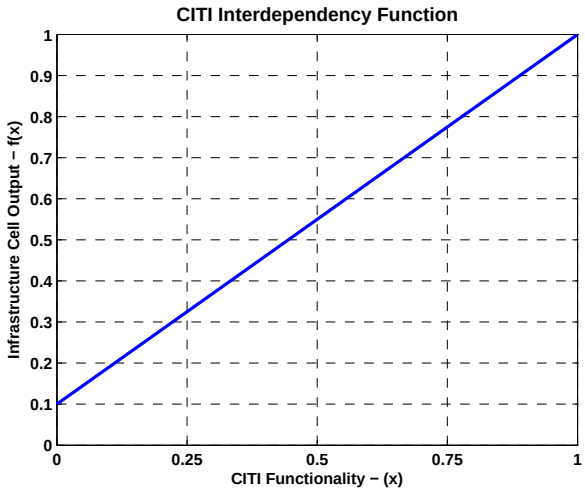


Figure 3.1: CITI Interdependency Function - Service Input vs Infrastructure Output

The CITI interdependency function is defined as a linear function. This assumption of linearity is justified due to several reasons. First, due to the statistical nature of CITI interdependency relationship and rarity of historical data to support such relationship

makes it difficult to construct any conclusive function of nonlinear interdependency. Second, we have observed during our interviews with the infrastructure operators that they tended to assume a linear performance decease of the operating infrastructure due to the corresponding decrease of the CITI services. The operators were not inclined to any nonlinear interdependency relationship. Considering these factors, the assumption of linearity implies a reasonable approximation. There is at-least one study by Anderson et al. [61] that shows the assumption of linearity was also observed among the information system managers on their perception of software usefulness in an organization.

3.4 Approach to Quantify CITI Interdependencies

In this section, we describe a simple interdependent infrastructure example and explain the assumptions, methods and scope of our work for defining the CITI interdependency functions. This background will help to understand our approach and results.

3.4.1 The CITI Interdependency Context

CITI services are comprised of different types of information systems and telecommunication services that contribute to the output of the infrastructure. These services are typically control and decision-making services that run within the infrastructure facility or may be provided by another CITI service provider through communication links. Interdependency results from the infrastructure's reliance on the CITI services for its consistent output. To quantify the CITI interdependencies we need some explanation of the concepts related to the Cell-Channel model [10, 20], such as 'cell', 'channel' and 'token'. The Cell-Channel model includes two other concepts 'clusters' and 'controls' (aggregator and distributor) that follow identical properties as in other infrastructures, and are hence omitted from further explanation.

Cell: A cell is an entity that performs a function. For example, a hospital is a cell that uses input tokens, such as, electricity, water, medicines, etc. and produces output tokens, such as, patients served. Normally, functional blocks within a cell are physically co-located, i.e. different departments of a hospital is located within a building or a set of buildings in the same campus. However, for CITI services, these physical entities may reside at distant locations. For instance, an Internet-based retailing site may have a single

web front-end (e.g. Amazon.com), but may have physical facilities located in different cities. In this chapter, we have used the word 'organization' to mean 'cell' in some places.

Channel: A channel is a means through which tokens flow from one cell to another cell. Depending upon the type of communication medium, the CITI channels can have different characteristics. For instance, a data communication channel may have queuing behavior (stochastic throughput and delay), while a voice communication channel may be of a circuit switching type (deterministic bandwidth and delay), etc.

Token: Tokens are goods and services provided by some entities to another entity that uses them. According to the Cell-Channel model, these tokens can be water, electricity, medical supplies, etc. However, for CITI services, the CITI's token is the most difficult concept to formalize. In any cell that uses a CITI service, a token is actually a contribution from different types of CITI services running in that infrastructure and contributing to the output of that particular infrastructure. Since different CITI services run in different departments within an infrastructure facility and have different levels of contribution at the organization level (cell) output, we need some mechanism to consolidate their contribution to a single CITI token type (contribution unit) for that infrastructure facility. Bedell's method, that we selected for our analysis has an organization level effectiveness measure (EIO index), which enables us to calculate this consolidation.

According to our extended assumptions of Bedell's methods, the CITI contribution to the infrastructure is proportional to the output, which we assume is represented by the effectiveness index (EIO Index). Therefore, flow of CITI service tokens directly contributes to the particular infrastructure's output. CITI tokens are measured as a number (ratio) in relation to the data packet throughput available to a particular cell at any moment in time, to the data packet throughput required for the cell when all CITI services are running smoothly. Thus, CITI service tokens are unit-less (ranging between 0 to 1) and can be compared with the per-unit representation of electrical quantities. According to this assumption, the relationship between the CITI service token and the infrastructure's output is represented in the form of a CITI interdependency function, where X axis represents CITI token input and Y axis represents cell output, both of which are ranged between 0 to 1 (Figure 3.1).

The interdependencies relationship between three critical infrastructure units according to the Cell-Channel model is shown in Figure 3.2. In this figure, there is a CITI control center, a power house and a water station. The CITI control center provides data

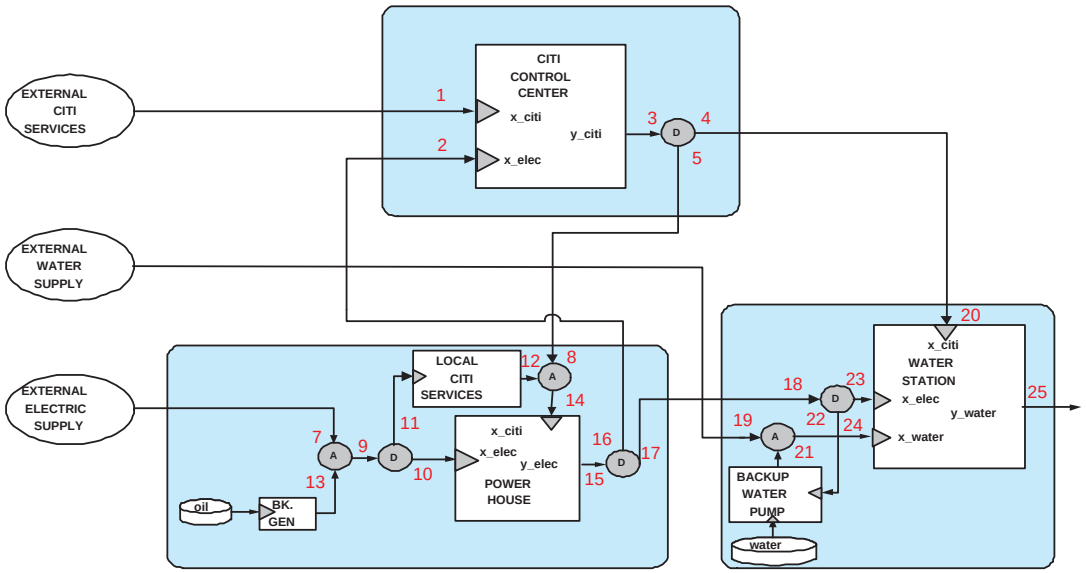


Figure 3.2: CITI Service Interdependency

communication and other information system services to the other two infrastructures. However, it depends on the power house infrastructure to run the CITI service facilities. The CITI control center gets an external connectivity and also gets some contribution of information system services from external sources through telecommunication links. The power house has its own internal control unit that provides some of the CITI-based monitoring and control services. But, it also gets some CITI services from the external CITI control center. Examples are Internet access, different web based information dissemination, sharing control information with other regional control centers, etc. We assume 60% of the power house's CITI service tokens come from internal CITI services and 40% from the CITI control center. Both of these CITI services are aggregated using a service aggregator and are fed into the power house functional block. This power house function block has two inputs (electricity and CITI) and one output (electricity). The output electricity from the power house function block is divided through a distributor and goes to the CITI control center and to the water station. Similarly, for the water station we have three inputs (water, electricity and CITI) and one output (water). However, for the water station, its control and decision-making is completely controlled from the CITI control center.

For the given example, the functional relationship between the power house input and output is represented by a three-dimensional function [10, 20], where one dimension is for

the CITI, another for the electricity input and the third one for the electricity output. In this chapter, we have shown two dimensions of these functional relationships, the CITI input and the electricity output that are represented by the CITI interdependency function. The CITI interdependency function (Figure 3.1) can also be represented as a Human Readable Table (HRT) [5] and is used in the results given in Section 3.5. Expressing interdependencies in terms of a table is a valid approach from a systems engineering perspective [52].

3.4.2 Quantification Method

In our research, we have followed three steps to formulate cyber interdependency functions (Figure 3.3). In the first step, we identify the important CITI services for the corresponding infrastructure cell. The service identification phase was based on contemporary research reports and from interviews. These research reports came from various types of domain specific studies. They included surveys related to infrastructure reliability, control center design, investigative reports on large scale failures and policy related papers. Our interviews with the infrastructure operators also helped to identify some of the important CITI services. Once the services were identified, in the next step, we ranked their effectiveness in relation to the infrastructures' output using Bedell's methods. The initial ranking was based on our understanding of the importance and effectiveness of the services from the research reports and interviews. Using Bedell's effectiveness index, we constructed CITI interdependency functions and HRTs which showed the interdependencies relationship between the CITI and the corresponding infrastructures. In the third step, we validated our interdependencies relationships by comparing various similar cases from our database. The database that we developed (see Chapter 2) is related to publicly-known CITI failure cases and gives us the unique opportunity to understand the impact of any particular CITI service component on the critical infrastructures from their absence (failure) as depicted in the reports. Some validations were also done through additional interviews with infrastructure managers and operators. If a validation showed a much different level of relationship from the anticipated one, we re-examined our ranking and recalculated the functions. This validation approach was subjective to some extent. However, since we followed a systematic approach to develop these functions, our approach in the case of incomplete information was an acceptable proposition [51].

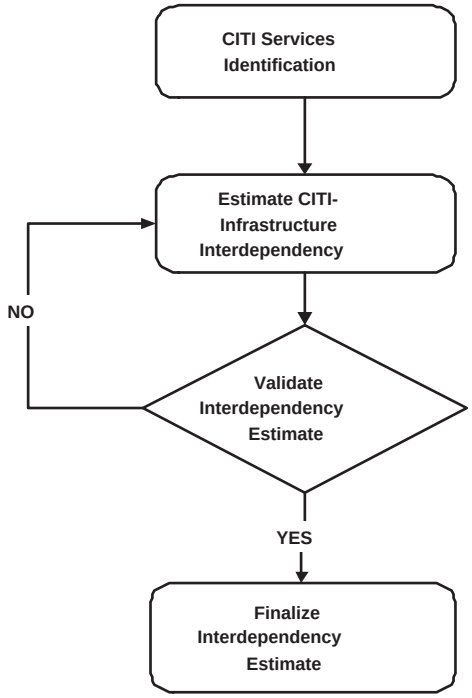


Figure 3.3: Quantification of CITI Interdependencies

We have presented some of the failure reports in this chapter to show the rationale behind our rankings. Due to space limitations, only important parts of the reports are cited. More detailed reports are available from the RISKS forum website [38] using the report title and reference-ID given at the end of each cited report. The reference id is written as RISKS (i, j), where i is the volume number and j is the issue number within the volume. In the interviews, we asked questions of the infrastructure managers and operators regarding the importance of CITI applications, their contribution in infrastructure output, impact of the unavailability of communication links to these applications, possible type of simulation scenarios and validity of our empirical interdependency functions in relation to their own experience. Many of the discussions were open ended and some were documented in our internal reports. The people we interviewed were mostly from UBC, BCTC (power infrastructure), TELUS (telecoms infrastructure) and HSBC (bank).

3.4.3 Scope of the Work

The following ten infrastructure sectors are defined as critical by the Canadian Government [6]. We have followed this classification in our work.

- Energy and utilities (e.g. electrical power, natural gas, oil production and transmission systems)
- Communications and information technology (e.g. telecommunications, broadcasting systems, software, hardware and networks including the Internet)
- Finance (e.g. banking, securities and investment)
- Healthcare (e.g. hospitals, healthcare and blood supply facilities, laboratories and pharmaceuticals)
- Food (e.g. safety, distribution, agriculture and food industry)
- Water (e.g. drinking water and waste water management)
- Transportation (e.g. air, rail, marine and surface)
- Safety (e.g. chemical, biological, radiological and nuclear safety, hazardous materials, search and rescue, emergency services and dams)
- Government (e.g. services, facilities, information networks, assets and key national sites and monuments)
- Manufacturing (e.g. defense industrial base, chemical industry)

CITI interdependencies for different organizations within each of these ten sectors were estimated. We selected an organization (cell) as a physical facility (e.g. power house, water station) that provides certain type of service (e.g. electricity, water) to the outside world. For some of the organizations we did not get enough information to estimate CITI interdependency functions. For instance, we modeled a hospital for the healthcare sector, but could not model the blood supply facility or pharmaceuticals industry due to lack of information. If enough information is available, interdependencies for these missing infrastructures can be modeled by following our methodology. In our work, we have only considered the most important CITI services needed by an organization based on the available information from research reports and interviews. Due to this insufficiency of information, our interdependency indices may be considered as the lower bound of the estimated CITI interdependencies.

3.4.4 Uses and Limitations of Our Approach

The functions we developed are key components for CITI interdependencies simulation in the I2Sim simulator (details provided in Chapter 5). In our approach, we have identified important CITI services needed for the effective operation of an infrastructure entity. Typical infrastructure entities like electrical power station, water station and hospital are represented as cells in the I2Sim simulator. The CITI interdependency functions we

estimated for each of these cells show how the outputs of the cells relate to the aggregate contribution from different CITI services running within the cell. As the empirical function represents aggregate behavior of different CITI services at the cell level, this is different from the impact of any individual CITI service within the cell. For instance, the Electronic Medical Record (EMR) database is a key component for the lab and radiology department of any modern hospital. If the EMR database is unavailable, there may not be any output from the lab and radiology departments. As such, for an EMR system, we may assign ESA and ISA indices of 10. Given this value, the CITI effectiveness index (EIA) at the department level becomes 10. This implies that the output for the lab and radiology department may vary between 0 to 1 in the CITI interdependency function representation. However, when considered as a whole for a hospital, the EMR database is identified as a "major strategic activity" (IAO = 8, ref: Table 3.1) and has lesser impact (less than 10) in the aggregated CITI function for the the hospital. Accordingly, the range of impact of the department level variation (0 to 1) is not visible in the overall CITI-Hospital interdependency function (see Table 3.9). The interdependency functions we developed in this chapter are used in the CITI interdependency simulation in Chapter 5. We assumed the same CITI services are running in the I2Sim representation of the hospital, water station, power house and other infrastructure cells. However, the assumption of identical services may not be appropriate for some simulation scenarios. In such cases, it is recommended to identify cell specific CITI services and to estimate the interdependencies functions for these particular cells. The questionnaire given in Appendix-A is helpful in this regard.

3.5 CITI Interdependency Functions for Critical Infrastructures

For each infrastructure in the following subsections, we provided an overview of the CITI services that contribute to the operation of the infrastructure. Following that, we assign a numeric rank to each of the services, according to the Bedell's index. We have used Table 3.1 to assign ranks to the services that belong to each of the infrastructures and interdependency functions were calculated based on this rank. Finally, we compare the estimated functions with real-life examples from our database to justify their validity.

3.5.1 Energy and Utilities

The energy and utilities sector includes electrical power, natural gas, oil production and transmission systems. Among these infrastructures, we only have sufficient information for the electrical infrastructure for doing interdependency estimate.

The CITI and the Electrical Infrastructure Interdependency: The electrical infrastructure has three major components for its operation. These are the electricity generation, transmission and distribution phases. The CITI services play an important role in all three phases. Supervisory Control and Data Acquisition (SCADA) systems are necessary for reliable operation of electrical generation equipments [62]. Energy Management Systems (EMS) are important for transmission and distribution network management [63]. Both EMS and SCADA systems are critical components in power system control centers. Malfunction of the SCADA or the EMS system has a major impact on electrical energy generation and distribution [64]. However, these systems are normally well designed and carefully implemented. So, most of the time they provide reliable operation in normal operating conditions. In the distribution phase, Business Management Systems (BMS) are responsible for consumer billing. For a power station's day to day operation there are other CITI services. These include, the Facility Management System (FM), the Geographic Information System (GIS), the Management Information System (MIS), etc. [63] However, except for the SCADA and the EMS, the rest of the CITI systems play only supporting role. Their absence does not critically affect electricity generation and distribution. For a power house facility where all these CITI systems are running, we estimated the following effectiveness indices (Table 3.2).

Table 3.2: CITI-Electrical Infrastructure Effectiveness Index

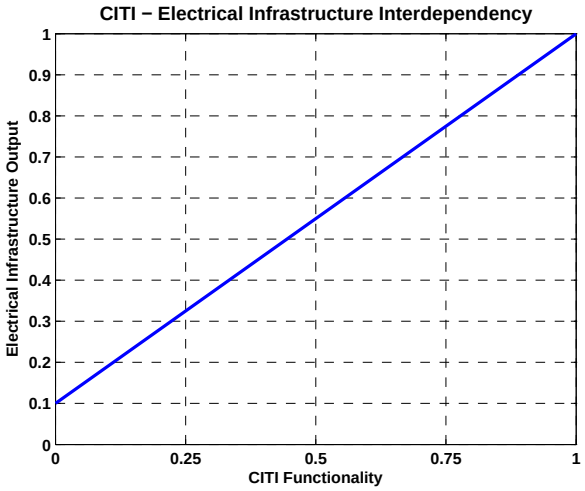
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Generation	SCADA	10	10	8	80	800
Transmission	EMS	10	10	8	80	800
Distribution	BMS	5	5	4	20	100
Facility management, Office automation	FM, MIS	5	5	4	20	100
Total					200	1800
EIO Index for Electrical Infrastructure = 1800/200 = 9						

According to the EIO index in Table 3.2, we constructed a CITI-Electrical Infrastructure interdependency function and HRT (Table 3.3). The EIO index (value 9) shows that the electrical infrastructure has a strong dependency on the CITI services. From the CITI interdependency function, it appears, in the absence of the CITI service, that the electrical infrastructure can still have some output. This implies that the electrical infrastructure may operate without the CITI support, but the output will be limited and the operating condition may be highly unstable [64].

Table 3.3: CITI-Electrical Infrastructure Functional Interdependency

CITI - Electrical Inf. HRT

CITI Contrib	Electricity Output
0	0.10
0.25	0.325
0.5	0.55
0.75	0.775
1	1



We present the following three reports from our database to validate our assessment of the CITI-Electrical infrastructure interdependency. The first two reports show the critical nature of the SCADA and the EMS systems. However, in contrast to the first two reports, the third report shows that a failure in the BMS has trivial impact on the electrical infrastructure's operation:

> SCADA system: *Amec Engineering, which designed the Beverly uranium processing plant in Western Australia, has blamed buggy software for a radioactive spill that occurred at the site last December, confirming early suspicions that computers played a role in the accident. According to Amec's report, the glitch cut power to the plant's fluid-distribution control system during a routine service exercise., "Software bug blamed in radioactive spill", CNET News.com, RISKS (21, 90)*
>
> EMS system: *A previously-unknown software flaw in a widely-deployed General Electric energy management system contributed to the devastating scope of the 14 Aug 2003 northeastern U.S. blackout. The bug in GE Energy's XA/21 system*

was discovered in an intensive code audit conducted by GE and a contractor in the weeks following the blackout, according to FirstEnergy Corp., the Ohio utility where investigators say the blackout began., "Software bug contributed to blackout", SecurityFocus, RISKS (23, 18)

BMS system: *On April Fools' Day, Randy Carrol in North Platte, Nebraska, received an electric bill of $12,344.16 for 33 days' service. But it was not a joke – except that the amount was generated by new billing software, showing use of 310,421 kilowatts (instead of the usual 300). The correct amount due later turned out to be $26.26., "Man Gets $12,000 Electric Bill", AP item, 4 Apr 2003, RISKS (22, 68)*

The CITI and the Oil and Gas Infrastructure Interdependency: CITI systems have wide application in the oil and gas infrastructures. Most oil and gas installations use the SCADA systems and many other types of control and monitoring technologies [65, 66]. However, we could not interview any oil and gas service providers to make any reliable estimate of cyber interdependencies. Furthermore, we did not have enough information in our report database to validate any CITI-Oil and Gas network interdependency relationship. As a result, we did not attempt to estimate the interdependency relationship between the CITI and the oil and gas infrastructures.

3.5.2 Communications and Information Technology

Consistent operation of the CITI based systems depends on effective service to many of their core components. These include software services (e.g. web server, email gateway, virtualization system, database system); low level operating system tools, protocol and device drivers; communication devices and links; and different supporting utilities (e.g. firewall, virus scanner, spam filter) [11]. Any typical organization that has a modern information system infrastructure has all these services in use every day. Among these CITI core components, software services are critical. An organization's operational output strongly depends on trouble-free operation of the above listed software services. However, these systems (web server, email gateway, etc.) can often fail due to a malicious attack or may suffer performance bottlenecks [11]. On the other hand, system software and communication links are important and should be reliable. However, most of the protection and utility software performs less than expected in most cases. We may assume that a CITI service center within an organization has all these services running. For such a service center, we estimated the set of indices in Table 3.4.

Table 3.4: CITI Infrastructure Effectiveness Index

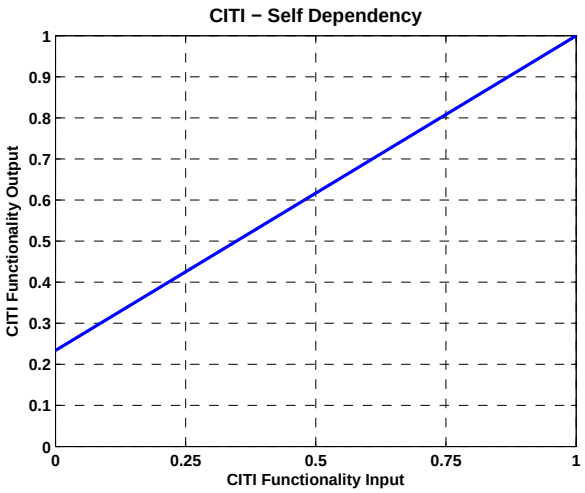
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Software services	Services	5	10	10	100	500
System software tools	Protocols, drivers	10	10	8	80	800
Communication channel	Com devices, links	10	10	8	80	800
Protection, Performance	Utilities	5	5	8	40	200
Total					300	2300
EIO Index for Electrical Infrastructure = 2300/300 = 7.66						

Using the EIO index in Table 3.4, we constructed the CITI interdependency function and HRT in Table 3.5. The CITI interdependency function shows that, even though the CITI core components are not working, some CITI outputs are still available. One explanation is that, the core components are related to network-based services. In the absence of network support, CITI systems work as individual systems. In such case, they can be useful to run non-network based applications. In the personal computer-centric environment, there are many such applications and their contribution is counted.

Table 3.5: CITI Self-dependency

CITI Functionality HRT

CITI Input	CITI Output
0	0.234
0.25	0.427
0.5	0.617
0.75	0.808
1	1



We present the following five reports to support of our ranking of the CITI's self dependency. The first report shows major consequences due to a malicious attack to the website of some of the important online business organizations. The second, third and

fourth reports present major impacts due to communication related software failure. In contrast, the fifth report shows that failure in an email distribution list has a relatively minor impact compared to the previous cases.

> Service failure: *The previous week saw three days of distributed denial-of-service (DDoS) attacks, disabling Yahoo, Amazon, eBay, CNN.com, Buy.com, ZDNet, E*Trade, and Excite.com for a few hours each. The flooding attacks were triggered from a variety of unknowing intermediate zombie systems that had been penetrated, although the launched DDoS attacks required no penetrations of their target systems., "Distributed denial-of-service attacks", Peter G. Neumann, RISKS (20, 79)*
>
> Protocol failure: *Netcom, Inc. one of the largest retail ISP's [450,000 subscribers, 230 POPs] went down for 14+ hours during the week of June 17, 1996, because of an extra "\&" in the border gateway protocol code in the MAE-East router in the Washington, D.C., area. Recovery required that all of the more than 100 routers be brought down, "The Great Netcom Crash", RISKS (18, 23)*
>
> Communication device failure: *MCI's inbound Internet gateways were saturated during July 1994, resulting in days of delay in delivering e-mail to MCI customers. A fix was considered to be months in the offing. "MCI Internet gateways choked", The Washington Post, August 1, 1994, RISKS (16, 30).*
>
> Communication link failure: *Australia's largest international Internet cable was severed on 20 Nov 2000 , partially disrupting Internet traffic in Singapore, Indonesia and Australia. The cable, carries about 60 percent of Australian ISP Telstra's international Web traffic. While Telstra has since managed to redirect most of its Internet traffic to another undersea cable, bringing its Internet services back to around 75 percent of capacity, its not yet been able to determine how long it will take for Internet traffic across the cable to return to normal. "Australian Internet cable severed", Dave Farber, RISKS (21, 13)*
>
> Utility (email distribution list) failure: *It appears that a gaping security hole at the University of Maryland led to an unexpected "canceling" of classes for Friday, April 11th. One or more students sent an e-mail to an address on campus which sent out to 3500 students and had no protection on it. From speaking to students at the school, it appears that they were signed up for an e-mail list without their knowledge, a list which accepted submissions from anywhere. "E-mail hoax at University of Maryland", Paul Kafasis, RISKS (22, 72)*

3.5.3 Finance

The financial sector includes different types of organizations, such as commercial banks, credit unions, stock exchanges, insurance companies, brokerages houses, etc. All of these

organizations rely heavily on CITI services for their daily operations. Among these entities, we have many reports related to the operations of a number of banks in our database. Banks offer a wide range of services, which can be grossly classified into online operations and offline operations [67]. Online services include, Internet banking, ATM /Debit card services and Credit card service. Offline operations are traditional banking services, which include mortgage and loan disbursement, money depositing and check cashing, mutual fund trading, etc. Both online and offline operations heavily use CITI services for day-to-day operations [67]. For a typical banking facility, Table 3.6 shows the CITI services. Among online banking operations, ATM, debit card and credit card services are the most important ones. However, the credit card service is often subject to fraud and identity theft and has not yet achieved trouble-free user confidence. Internet banking is mostly limited to a subset of functionalities smaller than the user can get from the offline banking and remains a supplementary service. For offline banking, mutual funds and the stock market have less to do with most banks' mainstream financial services, hence they are considered as a contributory service. Given this scenario, we estimated the following set of indices for a banking facility (Table 3.6).

Table 3.6: CITI-Bank Effectiveness Index

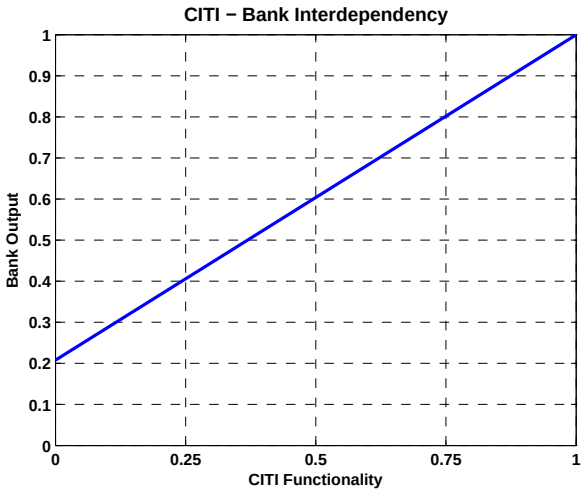
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Online banking	Internet banking	5	10	6	60	300
	ATM/Debit card	10	10	8	80	800
	Credit card	5	10	8	80	400
Offline banking	Mortgage/Financing	10	10	8	80	800
	Deposit/Checking	10	10	8	80	800
	Mutual fund/Stock	5	5	6	30	150
Total					410	3250
EIO Index for Bank = 3250/410 = 7.92						

According to the EIO index of Table 3.6, we constructed the CITI-Bank interdependency function and HRT (Table 3.7). The EIO index of 7.92 indicates that some of the bank's activities are independent of CITI services. As such, a bank can perform those functions even when there is no CITI service available. Such activities correspond to the non-electronic side of banking. This may include paper based operations, such as preparation of mortgage financing agreements and processing lending documents, etc.

Table 3.7: CITI-Bank Functional Interdependency

CITI - Bank HRT

CITI Contrib	Bank Output
0	0.208
0.25	0.406
0.5	0.604
0.75	0.802
1	1



The following four reports illustrate some of the real-life impact of different CITI services on banks' operation. The first two reports show that CITI services have a major impact on online banking. The last two reports show that offline banking is also affected considerably due to CITI services failure.

ATM and Online Payment: *Citibank's network of 2000 automated teller machines went down on the evening of 4 Sep 2001, due to software problems. It was still down the next day. Citibank's online Internet system also crashed at the same time. Basic service was restored about two hours later, but various problems persisted., "Citibank ATM network outage", Reuters, 5 Sep 2001; RISKS (21, 65)*

Debit and Credit Card multiple billing: *The approximately 100,000 merchants using CyberCash's IC Verify transaction software were given the opportunity to install a free upgrade for Y2K, but some of the merchants apparently did not get the fixes. Those who did not do so are apparently rebilling each customer transaction in the new year once each day until the fix is installed. Visa and Mastercard have installed software that attempts to catch the multiple transactions., "Y2K multiple billings", Peter G. Neumann, RISKS (20, 74)*

Offline banking services: *Canada's largest bank, the Royal Bank of Canada, has been unable to process deposits or report balances for the last five days. The bank is blaming "a processing disruption during a routine programming update to one of our computer systems". Direct deposit to a bank account is a common way to pay salaries here, especially for large employers, and among those affected are employees of the Government of Ontario, Canada's largest province, which apparently uses the Royal Bank for its payroll. The Royal Bank has 10 million customers, about a third of*

> *Canada's population. The bank confirms that "the processing disruption was national in scope so we expect a significant number of clients have been affected"., "Canada's largest bank has processing disruption", Yves Bellefeuille, RISKS (23, 43)*
>
> Both online and offline banking services: *The Swedish bank Nordbanken has suffered repeated computer outages during late December and early January. The outages, each with a duration of several hours, shut down ATMs, Internet bank services, debit card purchases and office teller services for Nordbanken's 3.5 million customers., "Repeated computer outages for Swedish bank", Ulf Lindqvist, Swedish CNN Web site, RISKS (21, 18)*

CITI failures have had major impacts on the operation of the stock market. We have several reports in our database related to stock markets disruption. However, we could not get any research report that provides a service-wise breakdown of the CITI systems in the stock market. Accordingly, we have not developed an interdependency model for the stock market. Similarly, we do not have any research report or other data related to insurance companies or brokerages houses. The following report shows the impact of CITI failures on the stock market.

> *A software upgrade glitch resulted in the New York Stock Exchange being unable to trade roughly half of its stocks in the morning of 8 Jun 2001. Consequently, the exchange was shut down entirely (on grounds of fairness) until 11:35 a.m. EDT. The RISKS archives note a 41-minute shutdown on 24 Feb 1971 (when both primary and backup systems failed), a 24-minute outage on 22 Oct 1991 (due to a power dip), a one-hour outage on 18 Dec 1995 (also due to a botched software update), and a one-hour crash on 26 Oct 1998., "Another NY Stock Exchange outage", Peter G. Neumann, RISKS (21, 46)*

3.5.4 Healthcare

The healthcare sector is a major application area for CITI technologies. Important organizations within the healthcare network are hospitals, blood supply facilities, laboratories, pharmaceuticals, etc. Since we do not have enough information about facilities other than hospitals, we focused only on hospitals. The hospital information system is one of the most important CITI technologies within the healthcare sector [68]. A hospital information system encompasses all aspects of patient treatment, medical report archiving, hospital management, and human resource activities [69]. There are many subsystems within the hospital information system, that focus on different areas of hospital automation. These

include, Computerized Physician Order Entry (CPOE) system [70], Patient Database, Lab Database, Hospital Management System, etc. [71, 72]. For a hospital where all these CITI systems are running, we have assigned different effectiveness and importance indices to these systems (Table 3.8), based on their utility, as understood from the research literature [70–72].

Table 3.8: CITI-Hospital Effectiveness Index

Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Prescription & Medicine	CPOE	5	10	8	80	400
Patient Database	GE Centricity	10	5	6	30	300
Lab & Radiology Database	EMR	10	10	8	80	800
Hospital Management	Doctors Roster	10	5	6	30	300
	Food supply	10	5	6	30	300
	Bed scheduling	5	5	6	30	150
	HR (pay role, etc.)	5	5	4	20	100
CITI Utilities	Firewall/Antivirus	5	10	6	60	300
Total					360	2650
EIO Index for Hospital = 2650/360 = 7.36						

Using the EIO index in Table 3.8, we constructed the CITI-Hospital interdependency function and HRT in Table 3.9. An EIO index of 7.36 implies that hospitals can operate without the CITI services to a certain extent. This is because doctors can prescribe drugs based on their experience and can get some medical history simply by asking the patient. Some of hospital management decisions can also be made by humans (bed scheduling, food supply, etc.).

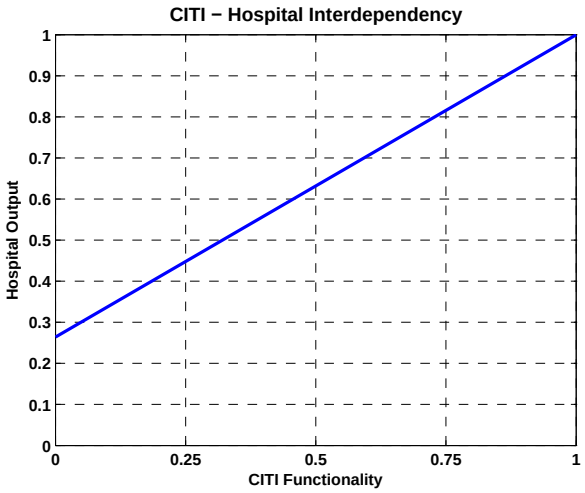
The following five reports validate our ranking by showing how hospital operations are dependent on different CITI services. The first report shows how a poorly designed or malfunctioning computerized prescription system can bring significant risk to the patients. The same is true for the lab and radiology database (report two). The remaining three reports show that absence of other CITI systems can slow a hospital's regular servicing rate to a significant extent.

> Prescription and Medicine: *Hospitals in USA those use specific medication software Hospital computer systems widely touted as the best way to eliminate dangerous*

Table 3.9: CITI-Hospital Functional Interdependency

CITI - Hospital HRT

CITI Contrib	Hospital Output
0	0.264
0.25	0.448
0.5	0.632
0.75	0.8
1	1



medication mix-ups can actually introduce many errors, according to the most comprehensive study of hazards of the new technology. The researchers, who shadowed doctors and nurses in the University of Pennsylvania hospital for four months, found that some patients were put at risk of getting double doses of their medicine while others get none at all. 22 types of mistakes were identified, such as failing to stop old medications when adding new ones or forgetting that the computer automatically suspended medications after surgery.,"Drug-error risk at hospitals tied to computers", The Boston Globe, 9 Mar 2005, RISKS (23, 78).,

Lab and Radiology Database: *According to the article "The Calgary Health Region" announced Sunday that an Internet database - which physicians use to view lab work such as blood and urine tests - mixed up results between patients and posted records under the wrong names. Officials are now contacting the offices of nearly 400 doctors and other health providers who saw the incorrect records, to ensure patients are receiving proper treatment." Doctors are concerned that the mix-up means some patients are now receiving incorrect treatments which can complicate their conditions, or that patients are receiving treatments they don't need., "2,000 patients hit by lab test mix-up in Calgary, Alberta", Robert Tremonti, RISKS (23, 94)*

Patient Database: *Lisbon newspaper "O Público" reports today that the main information system for the Lisbon Hospital Center, which supports three large Lisbon hospitals, has not worked since July 8. It appears that the master patient index has become inaccessible, and may be lost. If a patient shows up without a hospital-issued card, which includes a patient id number, the patient's records cannot be accessed. Out- patient consultations and admissions are being processed manually,*

causing "great confusion." Emergency room admissions are much slower than usual. "Information system for Lisbon hospitals stopped for ten days", RISKS (23, 94)

Hospital Management: *Beth Israel Deaconess Medical Center was paralyzed for four days by a computer crash in November 2002. Dr. Peter Kilbridge, an independent consultant who reviewed the incident at Beth Israel at the request of the *New England Journal of Medicine* editor, Dr. Jeffrey Drazen, said even if hospitals have policies in place to encourage the appropriate use of computers, those policies are often are ignored.,"Computer crashes threaten hospital operations", Associated Press, 7 Mar 2003, RISKS (22, 62)*

CITI Utilities: *The hospitals in "Västra Götaland" Sweden (west coast, population 1M) were isolated from Internet during 23 Sep 2001. Some of internal networks had to be partitioned to prevent nimda spreading further. Reservations and computer-based medical records were unavailable. "All public hospitals in Gothenburg Sweden Crippled by nimda", RISKS (21, 67)*

3.5.5 Food

The application of CITI technologies in agriculture and food distribution networks was studied by Hunt et al. [73] and Vorst et al. [74]. From these studies, it appears that agriculture and food distribution networks have been using communication and information technology services for nearly two decades. CITI technologies are employed in all four stages of the food distribution network, which include farming, food processing, distribution chain and retailing network management [74]. CITI systems are also used in some high level management, such as production planning and monitoring and quality control of food processing facilities. However, we did not have enough information in our database to validate any relationship between CITI and food sectors. Besides, we could not reach people from food industry for interviews. Due to insufficient information we have not included food related CITI interdependency modelling in our study.

3.5.6 Water Supply

There are two important components in most water-supply network: the drinking water network and the waste water management network. Although the history of virtually all drinking water distribution network is quite old, the centralized monitoring and control of such network is a relatively recent phenomena [75]. In recent years, new systems have been built to monitor and manage water reservoirs and distribution networks using different

kinds of sensors and SCADA technologies [75–77]. There are three aspects to online drinking water network monitoring, monitoring of water sources, monitoring of water treatment facilities and monitoring of the distribution network [75]. These monitoring concepts are similar to those in the electrical infrastructure. All these CITI based systems are deployed in a modern water station control room facility. Drinking water networks, water reservoir and water treatment facilities are all strictly monitored [75]. However, for water distribution networks the control mechanisms have yet to be improved. It also appears that due the recent introduction of CITI based monitoring and control of most drinking water networks, the absence of CITI systems may have moderate impact on that infrastructure's successful operation. Nevertheless, CITI services are considered to be a major support mechanism in most cases. According to this observation we estimated the following set of indices for a drinking water supply control center where all these CITI systems are running (Table-3.10).

Table 3.10: CITI-Drinking Water Infrastructure Effectiveness Index

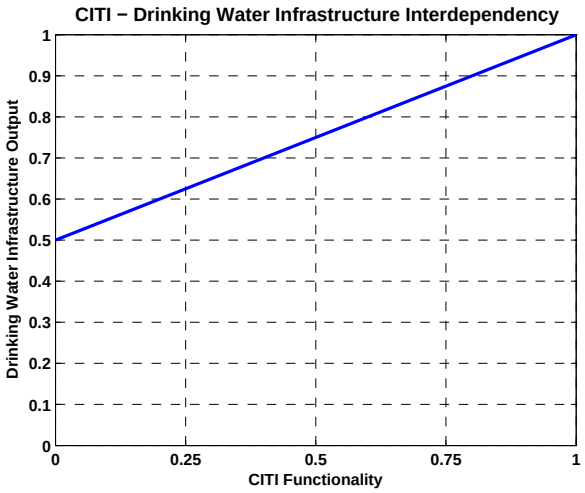
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Reservoir control	Source monitor	5	5	8	40	200
Water treatment control	Quality monitor	5	10	8	80	400
Distribution control	Distribution monitor	5	5	4	20	100
Plant Operation, Office Automation	Utilities	5	5	4	20	100
Total					160	800
EIO Index for Drinking Water Infrastructure = 800/160 = 5						

According to the EIO index in Table 3.10, we constructed the CITI-Drinking Water Infrastructure interdependency function and HRT (Table 3.11). In Table 3.10, an EIO index of 5 represents drinking water infrastructure's moderate dependency on the CITI services. From the CITI interdependency function, it appears that the drinking water infrastructure may operate without CITI support to a considerable extent.

The following three reports from our database validate our assessment of the CITI-Drinking water infrastructure interdependency. The first report is on water reservoir monitoring systems and shows a volatile condition in reservoir's operation due to the SCADA system failure. The second report presents a potentially unsafe water quality due to

Table 3.11: CITI-Drinking Water Infrastructure Functional Interdependency

CITI - Drinking Water Inf. HRT	
CITI Contrib	Water Output
0	0.5
0.25	0.625
0.5	0.75
0.75	0.875
1	1



a treatment facilities control system problem. The third report shows how an unmonitored water distribution network may sometime create problems for other infrastructures.

Water reservoir monitoring: *Four anglers were rescued by helicopter Wednesday from a small island in the Capilano River after a control malfunction at the Greater Vancouver regional district's Cleveland Dam released an unexpected torrent of water. The malfunction of the drum-gate water-control mechanism, which occurred during a computer upgrade, is expected to prompt the installation of more signs along the river warning fishermen of the potential for rapid increases in water levels. The Cleveland Dam had been releasing snow-melt water through the drum-gate at the rate of about 1,000 cubic feet per second when the malfunction occurred at 7 a.m. Wednesday. By the time dam employees brought the problem under control about an hour later, the flow had increased four-fold to about 4,000 cubic feet per second., "Fishermen rescued after dam malfunction", The Vancouver Sun, 27 Jun 2002, RISKS (22, 14)*

Water treatment monitoring: *A computer glitch in Lewiston, Maine, shut down the chlorination system and caused the chlorine content of the city water to drop below the safety threshold, affecting 40,000 residents. This occurred during the night, and was not discovered until a routine check 14 hours later. Notices were then sent out to 9,000 homes advising people "to boil the water before drinking. It took 30 hours to solve the problem., "Computer flaw makes water undrinkable", USA Today, 17 Aug 1998, RISKS (19, 92)*

Water distribution network failure: *In Durham, NC (USA), a water pipe break on early Saturday (12-Jan-2002) morning forced the closure of the city police department building and 911 center. The water flooded a subbasement and took out the electrical equipment and backup power generators. Callers to 911 got busy signals*

> *or disconnects (I suppose that's better than hold muzak) until the temporary location (at Duke University) was online about 12 hours later, with dispatchers taking call information on paper (no computers)., "Water line break closes 911 center & police department", The News & Observer, 24 Jan 2002, RISKS (21, 89)*

The real-time monitoring and control for waste-water and urban drainage systems is not very common [76]. Lack of real-time monitoring makes CITI services irrelevant. Besides, we do not have sufficient information in our database to validate any assessment. As such, CITI-Waste-water infrastructure is outside the scope of present work.

3.5.7 Transportation

In the transportation sector, the CITI services play a very important role. This sector includes four subsectors: road, rail, air and marine. A wide rage of CITI based technologies are in use in all these four subsectors [78]. A brief description of CITI interdependencies for each of these transportation sectors is given in this section.

3.5.7.1 Road Transportation

CITI technologies have been used for a long time in road transport systems [78]. Concepts such as intelligent highway or real-time traffic monitoring and control technologies have been studied for more than a decade. CITI applications used in the road transportation network include network control and traffic management, traffic information dissemination, service fee collection, and parking and facilities management systems [78]. Each of these applications may have many smaller management functions; for instance, the network control application may have traffic light control systems, congestion detection systems, route guidance systems. Among these applications, a network control system is the most critical in order to run the road networks. Network control has been around for quite long time and has proved to be a reliable system to control road networks. Toll collection system is important for government revenue. However, there are reports of inconsistent performance. The traffic guide and facilities management systems are utility services and are not as critical as the other two. For a road-network control-room [79], we may assume that all these CITI systems are running and we estimate these indices in Table 3.12.

According to the EIO index in Table 3.12, we have constructed the CITI-Road Networks interdependency function and HRT in Table 3.13. A low EIO index of 6.82

Table 3.12: CITI-Road Networks Effectiveness Index

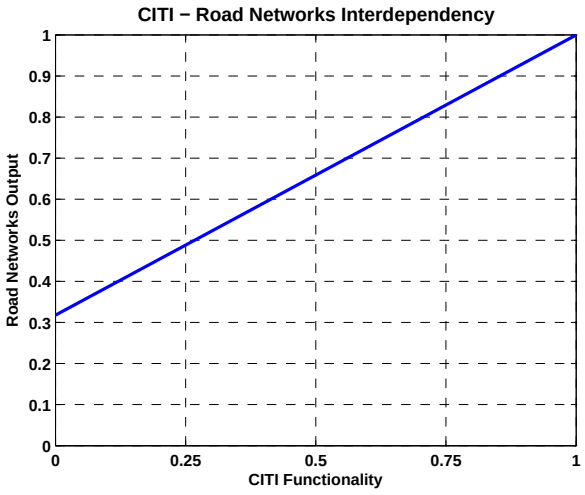
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Network control and traffic management	Signaling	10	10	8	80	800
Service fee collection	Toll collection	5	10	8	80	400
Traffic information dissemination	Traffic guide	5	5	6	30	150
Parking and facilities management	Facilities management	5	5	6	30	150
Total					220	1500
EIO Index for Road Networks = 1500/220 = 6.82						

implies an inadequate contribution of some of these systems in road network management and they do not constitute critical elements for the road transportation network.

Table 3.13: CITI-Road Networks Functional Interdependency

CITI - Road Inf. HRT

CITI Contrib	Road Output
0	0.318
0.25	0.488
0.5	0.659
0.75	0.829
1	1



The following four reports are presented from our database to validate our assessment of CITI-Road networks interdependency. The first report shows the huge consequences of the signaling systems' failure. The other three reports present the non-critical nature of the utility systems' failure.

Signaling: *Users of London city road transport system Central London was brought to a standstill in the rush hour today when 800 sets of traffic lights failed at the same time – in effect locking signals on red. "The worst gridlock the capital has seen*

for years was caused by a computer which crashed as engineers installed software designed to give pedestrians longer to cross the roads.", "Gridlock as 800 London traffic lights seize", Adrian Lightly, RISKS (22, 18)

Toll collection: *Fast Lane double-billed 8,498 accounts this week, an error Massachusetts Turnpike Authority officials attributed yesterday to the electronic toll company running the system. The computer glitch drew money Tuesday out of credit card and checking accounts belonging to Fast Lane customers, then mistakenly docked the same customers Wednesday. The total wrongly withdrawn could amount to tens of thousands of dollars, said the Turnpike spokeswoman, Mariellen Burns., "Some Fast Lane accounts double-billed", Boston.com, RISKS (24, 9)*

Traffic guide: *Drivers on southbound Interstate 75 in Michigan saw a construction message board that previously had been alerting drivers in Genesee County near Clio that construction was soon to start. One morning it said "speed limit 100 mph go go go." (The speed limit in that area is 70 mph. The sign is controlled remotely by a subcontractor's computer.), "Michigan message board says speed limit 100 mph", AP item from The Boston Globe, 8 Apr 2005; RISKS (23, 84)*

Parking Facility: *The system controlling York's newly installed intelligent traffic variable-message signs (VMS) were hit by a computer virus on 4 Oct 2003, freezing 21 VMS displays at car parks that were intended to show the number of available parking space. Motorists thus went into full car parks expecting to find space. One VMS at St George's Field showed 349 spaces when there were *none*, causing an enormous traffic tie-up. Yorkshire Evening Press, 6 Oct 2003;, "Parking chaos in York", RISKS (22, 92)*

3.5.7.2 Railway

In contrast to road network, a railway network is completely planned environment. All trains are pre-scheduled, their route and transit times are predefined, signaling and control mechanisms are also set ahead of time [78]. Any change in train schedules, disruption in power supply, or signaling can cause major disturbances in a railway networks' operation. In this environment, communication and coordination play very important roles. In the early days, these information systems were coordinated through phone or signal boxes. Now, they are replaced with high-speed computer networks and information systems [80]. Each train has unique reporting number, its travel duration, its progress and locations are logged in a scheduling system [78]. Typical CITI applications that run in a railway control room [78, 80] are the Railway Traffic Management System (RTMS) for scheduling and conflict resolution, the Communication-based Train Control (CBTC) for remote monitoring and maintenance, the Ticketing and Reservation Systems (TRS), and the Station

Information Management System (SIMS) for public display of train schedule and location update. Among these CITI systems, the first three are critical and have major impacts on the railway network. Accordingly, we have come up with the following indices for a railway network control center (Table 3.14).

Table 3.14: CITI-Railway Networks Effectiveness Index

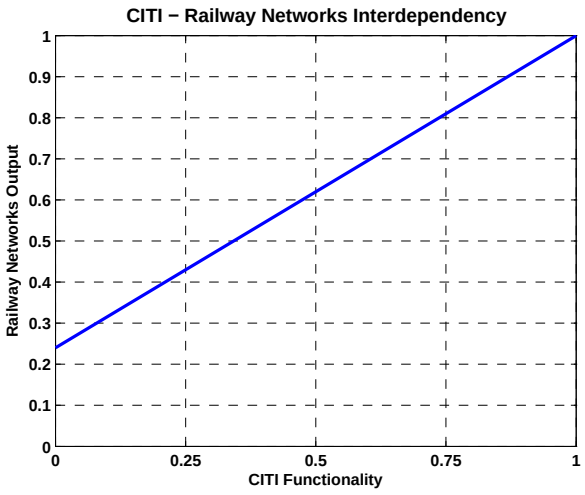
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Train scheduling	RTMS	10	10	8	80	800
Monitoring and control	CBTC	5	10	8	80	400
Ticketing	TRS	10	5	8	40	400
Information display	SIMS	5	5	6	30	150
Total					230	1750
EIO Index for Railway Networks = 1750/230 = 7.6						

According to the EIO index in Table 3.14, we have constructed the CITI-Railway network interdependency function and HRT (Table 3.15). The EIO index of 7.6 implies that manual operation is possible only in some aspects of the railway network, such as ticketing and information display.

Table 3.15: CITI-Railway Networks Functional Interdependency

CITI - Railway Inf. HRT

CITI Contrib	Railway Output
0	0.24
0.25	0.43
0.5	0.62
0.75	0.81
1	1



We present the following four reports from our database to validate our assessment of CITI-Railway network's interdependency. The first three reports show the major

interdependency of CITI systems on railway networks. The last report is related to information display and represents a relatively minor problem for the network's operation.

Train scheduling: *On Monday, 7th February the central computer at the rail control center for Zuerich main station in Switzerland failed. The outage was noticed at 08:40, and had deleterious consequences for further control centers which were dependent on the Zuerich center. The Associated Press reported that trains between Zuerich and Pfaeffikon, a commuter line on the left bank of Lake Zuerich, were all canceled for nearly four hours. Buses were used to ameliorate the situation, for example for trains in the direction of Chur., "Zuerich Main Railway Station Outage", Peter B. Ladkin, RISKS (23, 70)*

Monitoring and control: *City dwellers using railway transportation CSX Transportation's (CSXT) information technology systems experienced significant slowdowns early today after a computer virus infected the network. The cause was believed to be a worm virus similar to those that have infected the systems of other major companies and agencies in recent days. The infection resulted in a slowdown of major applications, including dispatching and signal systems. As a result, passenger and freight train traffic was halted immediately, including the morning commuter train service in the metropolitan Washington, D.C., area., "Lots of railroad traffic affected by so-big" , Danny Burstein, RISKS (22, 87)*

Ticketing: *Londoners were faced with travel problems this morning after an IT error meant hundreds of commuters could not renew journeys on their Oyster card. The error, which affected the whole of the London Underground (LU) and Docklands Light Railway (DLR), was caused when an overnight electronic updating process went wrong. Transport for London (TfL) and TranSys - the consortium that operates the Oyster card scheme - automatically updates the system each night to add new records and block stolen and canceled cards. But a glitch in the system early this morning means commuters are unable to use machines at Underground or DLR station this morning to add new journeys onto the smart cards., "Oyster card fault causes problems on London Underground", Computing, 10 Mar 2005, RISKS (23, 79)*

Information display: *Metro's $20 million central computer system crashed at 5:15 p.m. during the evening rush hour on 24 Apr 2001. The central system provides real-time graphics to the downtown control center. Similar malfunctions occurred in 1998 and 1999 (e.g., RISKS-20.60). In the 15 months following its installation, this BDM system crashed 50 times, according to the Metro. Coincidentally, a six-car train that had broken down 8 minutes earlier was stuck in the tunnel between Friendship Heights and Bethesda, and had to be towed out. The outage caused system-wide delays, with some passengers facing platform delays up to 45 minutes., "Computer system crash stalls D.C. Metro", Washington Post, RISKS (21, 36)*

3.5.7.3 Air Transportation

CITI services play a major role in air transportation networks [78, 81]. A typical airport operation as described by Fields et al. [82], has many critical CITI service related components. The key functionality of aviation is controlled by three or four controllers. These include, the tower controller that controls aircraft by issuing instructions and approving aircraft movements on the runway and in the air in close vicinity of the airport. The ground controller is responsible for approving the start of aircraft engines and issues clearances to moving aircraft. The planner maintains contact with the pilot and provides information regarding airspace around the airport to coordinate flight movement. There are other CITI service based operations in the airport related to passenger services. These include, reservation and baggage handling, security and custom clearance, etc. Among these services, controllers are the most important for an airport's operation, while the other services remain major contributing factors. Therefore, for an airport facility we came up with the following effectiveness indices for the CITI services (Table 3.16)

Table 3.16: CITI-Airport Effectiveness Index

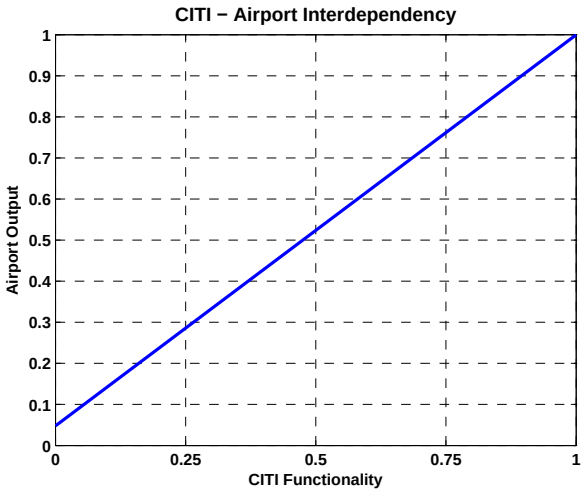
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Tower controller	TC	10	10	8	80	800
Ground controller	GC	10	10	8	80	800
Facilities planner	FP	10	10	8	80	800
Reservation & baggage	SABRE	10	5	8	40	400
Airport operation (security, customs, etc.)	Utilities	5	5	6	30	150
Total					310	2950
EIO Index for Airport = 2950/310 = 9.52						

According to the EIO index in Table 3.16, we constructed CITI-Airport interdependency function and HRT in Table 3.17. The high EIO index (value 9.52) demonstrates that an airport has a very strong dependency on CITI services. From the CITI interdependency function, it appears that in the absence of any CITI service there is virtually no output from the airport facility.

Table 3.17: CITI-Airport Functional Interdependency

CITI - Airport HRT

CITI Contrib	Airport Output
0	0.048
0.25	0.286
0.5	0.524
0.75	0.762
1	1



We present the following six reports from our database that validate our assessment of the CITI-Airport interdependency. The first three reports show the critical nature of the controllers and the remaining three reports show that the support services (baggage, security and customs) are also very important for the airport facility.

Tower Controller: *Yet again the new multi-million-pound air-traffic computer system at Swanwick near Heathrow crashed last Friday (May 17, 2002) shortly after 6.30 am. This is a time of maximum inbound flights from the Middle and Far East – with full 747s arriving at one a minute. Also too it is just when the morning rush hour for domestic and European departures and arrivals begins to build up. The crash was the result of a 'routine upgrade' which made half the air traffic controllers' computer screens inoperable. This meant that only half the normal flights could be handled. This meant that airlines had to cancel most of their flights into and out of Heathrow - a situation which lasted for most of the day., "Computer failure grounds over 300 flights in minutes", Chris Brady, RISKS (22, 9)*

Ground Controller: *A computer glitch kept Atlanta-bound Delta Air Lines flights on the ground for about two hours Saturday, but the company was gradually restoring service to its main hub. Delta told the Federal Aviation Administration it had a problem with dispatch computers, which calculate weight and balance and handle information related to preparation for flight, plus gate information, FAA spokeswoman Kathleen Bergen said.,"Computer glitch grounds Atlanta flights", Fredric Rice, RISKS (23, 35)*

Facilities planner: *A virus apparently attacked an AC Jazz flight-planning computer that provides essential information on fueling, weather, and other variables. Without*

the computer's flight information releases, aircraft cannot take off. The problem affected only Air Canada's regional operations. About 200 flights were affected, some canceled, some delayed.,"Air Canada "Jazz" airline grounded by computer glitch", National Post, 6 Feb 2003, RISKS (22, 54)

Baggage handling System (SABRE): *Hundreds of American Airlines flights were delayed Tuesday evening after its Sabre Group computer stopped functioning at 5:18 p.m. CT and was out of service for at least four hours. An American Airlines spokesman said the delays ranged from 8 minutes to a few hours on domestic and international flights, due to a software problem. This was the second Sabre central-system outage in a week. The outages affected about 50 airlines that use Sabre for reservations, seat assignments, baggage info, etc., "Computer glitch snags airline travel", USA Today, 1 Jul 1998, RISKS (19, 84)*

Airport operation, report #1: *Errors by screeners were responsible for false alarms over weapons that sparked the recent evacuation of Midway Airport and two other U.S. airports, according to the Transportation Security Administration. The confusion that led to the terminal evacuation on 15 Nov was prompted by a hand grenade appearing on an X-ray scanner. The image of the grenade, part of an exercise used to test screeners. A screener operating the X-ray scanner thought the grenade, artificially projected inside a carry-on bag, was real. But the passenger was able to leave the security checkpoint with the suspect bag before screeners could search its contents, leading to the evacuation order., "Midway scare is blamed on glitch", Chicago Tribune, 23 Nov 2004, RISKS (23, 61)*

Report #2: *A U.S. Customs database system in Virginia shut down for about 5.5 hours beginning around 6pm on 18 August. The system is used to process incoming international air passengers, but its absence caused havoc at Miami International Airport, where up to 2000 people were waiting to clear immigration., "Customs Computers Fail", Associated Press/AP Online, 22 Aug 2005, RISKS (24, 2)*

3.5.7.4 Marine Transport

CITI systems have important applications in marine transport [78]. However, we did not have much information in our database regarding waterborne transportation. Due to insufficient information, we have not included marine transportation-related interdependency analysis in our study.

3.5.8 Safety

The Canadian Government's description for public safety are related to chemical, biological, radiological and nuclear safety, hazardous materials, search and rescue, emergency

services and dams [6]. However, in this section we particularly discuss emergency services. Emergency services include 911 and emergency medical services (EMS), fire department and law enforcement services. All these three organizations work for public safety. A typical emergency service request starts from the event report through a 911 call. The service call is answered through the control center named Public Safety Answering Point (PSAP) [83]. PSAP routes the call to an appropriate emergency service provider. Based on the event type, either single or multiple emergency service providers take care of the incident individually or jointly [84]. For large-scale disasters, help from volunteers is sought [85]. There are several key components in this process where CITI services play major functions. First, emergency service providers are connected through Public Safety Networks (PSN) [86] for information sharing and interoperability. The PSN has a call-recoding database that keeps track of the detail of the service request [83]. The caller's location is provided by the telephone company or by the 911 database service named Master Street Address Guide (MSAG) [83]. Sometimes, emergency service providers heavily depend on the coordinated support from crisis response teams [85]. Among these CITI related services, public safety networks (PSN) work with moderate effectiveness [86]. The call database, location based service and public alert systems are not yet very reliable [83,85]. For a disaster scenario where all these CITI services are involved we came up with the following effectiveness indices (Table 3.18).

Table 3.18: CITI-Emergency Services Effectiveness Index

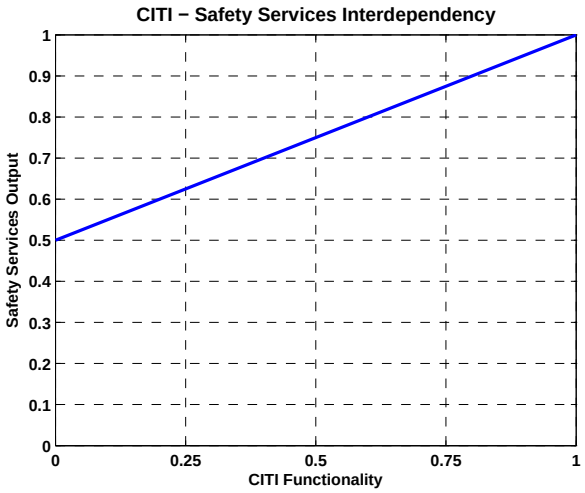
Activity	System	ESA	ISA	IAO	ISO	ESA x ISO
Inter-agency connectivity	PSN	5	10	10	100	500
911 call database	PSAP	5	10	8	80	400
Location based service	MSAG	5	10	8	80	400
Volunteer mobilization	Public Alert	5	10	8	80	400
Total					340	1700
EIO Index for Emergency Services = 1700/340 = 5						

According to the EIO index in Table 3.18, we have constructed the CITI-Emergency Services interdependency function and HRT in Table 3.19. The EIO index 5 implies that emergency services are not much dependent on CITI technologies. There are many

potential areas for improvement. However, such a low coupling may be good in the sense that emergency operations are somewhat immune to CITI service related problems.

Table 3.19: CITI-Emergency Services Functional Interdependency

CITI - Emergency Services HRT	
CITI Contrib	Emr. Serv. Output
0	0.5
0.25	0.625
0.5	0.75
0.75	0.875
1	1



We present the following four reports from our database to validate our assessment of CITI-Emergency Services interdependencies. All four reports show there are many places that need attention:

PSN call routing: *The Washington Post reports on a number of cases where calling 911 from a cell phone was routed to the wrong jurisdiction, so "response to a life-threatening – and ultimately fatal – emergency was delayed because a cell phone call to 911 didn't work the way it was supposed to". The examples given were a caller in Chillum MD routed to 911 in Washington DC (an immediately adjacent jurisdiction). They note that in the Chillum case, the problem occurred because "a wireless signal can get picked up by the wrong cell phone tower". "Cell phones & 911 service", Jeremy Epstein, RISKS (22, 67)*

911 call database: *A 911 dispatcher in Buncome County, North Carolina, clicked on a box to transfer the house address of a caller into the Computer Aided Dispatch system. But that system, installed in March 2003, did not yet have information on all Buncombe County roads, and suggested an incorrect alternative (Briarcliff Drive, instead of Lane, in West Asheville), which the dispatcher accepted. As a result, the paramedics were significantly delayed and the self-inflicted victim died. Attempts are now being made to complete the database. "Botched 911 call led to man's death", Citizen-Times, 15 Aug 2003, RISKS (22, 87)*

Location based service: *Two teenagers died when their rowboat sank in Long Island Sound on 24 Jan 2003. The 911 operators who took their last-minute phone call have*

been charged for not handling the call and delaying the search for a day. The story suggests that the operator attempted to enter "Long Island Sound" as the location but the software prevented that and, after consulting an equally ill-informed superviser, the operator simply gave up and dropped the call. "Deadly input validation?", Chris Adams, RISKS (22, 58)

Volunteer mobilization: *Connecticut state emergency management officials said a worker entered the wrong code during the weekly test of the emergency alert system, leading television viewers and radio listeners to believe that the state was being evacuated: "Civil authorities have issued an immediate evacuation order for all of Connecticut, beginning at 2:10 p.m. and ending at 3:10 p.m." The code that was mistakenly entered appeared on a monitor one line above the intended code for the test. As soon as the error was detected, faxes went out to every police department in the state. "Off-by-one error: Evacuate the entire state!", The Hartford Courant, 1 Feb 2005, RISKS (23, 70)*

3.5.9 Government

Many government services are heavily dependent on CITI services. These include, electronic voting systems, various administrative databases, online tax-filing systems, employee payrolls, welfare services, etc. However, government departments are huge entities and many of them do not have any single point of control (e.g. control center for coordinating actions), which could be an entity to formulate an interdependency function. As such, government organizations have to be modeled on a case by case basis following our methodology. In this section we give several examples from our database to show CITI interdependency for government organizations. However, we refrain from developing any interdependency function for a particular government administrative unit.

Election or eVoting: *Venezuela's high court on Thursday suspended this weekend's general elections, saying fair balloting is impossible until the problems are resolved. Conditions for "credibility and transparency" in Sunday's presidential, congressional and regional elections do not exist, said Ivan Rincon of the Supreme Tribunal of Justice. President Hugo Chavez had earlier blamed an Omaha (Neb.)-based company for the technical problems, saying it was part of an overall plan to "destabilize" the country's electoral process, "Venezuela cites computer glitch, postpones elections", Chicago Tribune, 26 May 2000 RISKS (20, 89)*

Administrative Databases: *Government's bureau of motor vehicle The Department of Motor Vehicles in Colorado was disabled all of last week by a computer virus. New and renewed licenses and ID cards were unabled to be issued during the time. Every computer in the system had to get fresh software installs and nearly 4.5 million*

documents had to be reloaded. No cost estimates have been given for the outage and no details released about the nature or origin of the virus., "Virus disables Colorado DMV for nearly a week", Brad Hill, RISKS (23, 56)

Online Tax filing and other Online Services: *KYW News Radio in Philadelphia reported on 17 Apr 2001 that there had been a problem when tax procrastinators attempted to file their Pennsylvania State returns just before the midnight Monday deadline. Apparently in the last few hours, users received an error message from the filing Web site, and they were unable to complete their transaction. Because of this, the state decided to give ALL late filers an extension through 18 Apr. Officials were quoted as saying that "a glitch on the Web server" was the cause of the problem (whatever that means)., "Denial of Tax Service", Rebecca Mercuri, RISKS (21, 35)*

Payroll and Government Payments: *On 1 Jun 2001, the majority of people on the government payroll were paid with a one-day delay. The same goes for refunds for VAT and taxes. The reason: Belgian postal services are tasked with doing the money transfers towards the different banks. Seems that they had a special situation: on 31 May, not only people had to be paid, but the next weekend (02-04 Jun) being a long one, an 'exceptionally large number' of transactions were fed to the system., "Payday delayed by one day in Belgium", Kris Carlier, RISKS (21, 45)*

3.5.10 Manufacturing

Manufacturing facilities use different types of enterprise management applications [87]. These systems include, Enterprise Resource Planning (ERP), Supply Chain Management (SCM), Customer Relation Management (CRM), etc. These CITI systems have a significant impact on production and marketing facilities. For instance, the following report shows production loss due to a virus attack on an electronic production facility.

Dell Computer's plant in Cork, Ireland suffered five days of downtime after the company discovered that 500 of its computers had been infected with the FunLove virus. Staff had to track down the source of the infection and eradicate the virus from all its systems. Paul Taylor (Reuters) wrote, "the attack is regarded as one of the most damaging seen in Europe." In addition to the lost production time, the incident damaged customer relations, with some customers complaining about the delay in delivery of their systems., "Dell loses five days' production time to FunLove Virus", Mich Kabay, RISKS (20, 66)

However, we do not have enough information both from the research literature and form our database to propose any functional relationship for CITI interdependency for a typical manufacturing facility. Thus manufacturing facility interdependency is not included in our present study.

3.6 Effect on CITI Services from other Critical Infrastructures

Continuity of CITI services is directly related to the availability of uninterrupted power supply. The following report shows the impact of electricity failure on CITI networks. This Electricity-CITI interdependency relationship (function) is a ON/OFF type (a step function).

> *More than 200 New England businesses experienced a four-hour Internet blackout on 7 Aug 1997 after an explosion knocked out electrical power in the Boston area. One person was killed in the blast, which overloaded a panel switch at MIT, causing a fire and cutting off Internet access to BBN Planet customers. Access resumed around 10:00. The speed with which the incident happened made it impossible to reroute traffic, said a BBN spokesman. "Explosion causes Internet blackout in New England", Edupage, 10 Aug 1997, RISKS (19, 29-30)*

The CITI network may be affected by a water system failure. Some reports about World Trade Center attack on September 11, 2001 [88] show, water flooded cable vaults and that apparently disrupted telecommunication networks. However, the interdependency of CITI on the water system network is largely situation-dependent and has no generalized rule. The following report shows one such scenario. Interdependency of CITI on other critical infrastructures are still unidentified at this point.

> *In Durham, NC (USA), a water pipe break on early Saturday (12-Jan-2002) morning forced the closure of the city police department building and 911 center. The water flooded a subbasement and took out the electrical equipment and backup power generators. Callers to 911 got busy signals or disconnects., "Water line break closes 911 center & police department", The News & Observer, RISKS (21, 89)*

3.7 Chapter Summary

Cyber interdependencies have a large impact on critical infrastructures due to the growing use of CITI services in the infrastructures' operation. Up to now, the exact nature of cyber interdependencies was an unknown area. The work presented in this chapter is an initial attempt to systematically understand cyber interdependencies and to formulate some useful rules. The findings presented here are supported by real-life infrastructure-related data. These functions show the level of coupling of an infrastructure with the

CITI services. There may be other CITI services which were not included in our study and therefore, the estimated functions can be considered as the lower bound of the interdependency relationship. A more detail study of these infrastructures may improve these estimates. The functions we estimated are implemented in the I2Sim simulator for CITI interdependency simulation (details in Chapter 5). A large number of simulation results can be used to fine-tune these functions. Our approach can also be extended to identify cyber interdependency functions for those infrastructure cells that are not covered in the present study. The interdependency relationships we have quantified will be useful to predict critical infrastructures dynamic behavior, stability and resilience for any large-scale disturbance.

Chapter 4

Design and Implementation of a Prototype I2Sim Simulator

Well-designed simulation frameworks can be used to provide significant insight into the interdependencies of critical infrastructure networks. The simulation results can predict the critical infrastructures dynamic behavior, as well as their stability and resilience for large-scale disturbances. The results can be useful for vulnerability identification, reinforcement planning, and adding or discarding redundancies in the infrastructure networks. Such tools can also be very helpful to manage disaster scenarios in order to coordinate and prioritize resource allocation and recovery activities. Therefore, critical infrastructure simulation tools can be very useful to infrastructure service planners and policy makers for secure and reliable infrastructure design and operation.

The modelling and simulation of individual infrastructures is a well-developed and mature field. Different formal models and simulation tools exist for individual infrastructures. These tools can precisely determine the operational characteristics of each of the individual infrastructures. Using those tools, one can predict an infrastructure's operational states for different failure conditions. In contrast to this, the modelling of multiple interdependent infrastructures is a relatively undeveloped area. This is largely due to the fact multiple infrastructures' control dynamics are governed by a very large number of variables that include both physical and human elements. As a result, their joint operational characteristics can be very nonlinear and complex, which makes the development of a unified model particularly challenging [89]. As a result, there had been very few multi-infrastructure simulation models or tools available in the recent past. However, in the last few years there has been considerable interest on the part of the governments of North

America and Europe [7, 33, 90] in critical infrastructures related research. Many new models are being proposed and simulation frameworks are being built. Pederson et al. [13] have compiled a survey on contemporary research on critical infrastructure modelling and simulation. This is a comprehensive-study that shows the wide variety of ideas proposed in recent years. Modelling approaches include techniques based on game theory, graph theory, risk-based models, Petri-net based models, etc. However, many of these interdependency models are in the conceptual phase. Only a handful simulation frameworks have been built that are available in the published literature. It is also observed that the vast majority of these recently implemented frameworks are based on agent-based technology [13, 91].

In the agent based solution framework, a population of autonomous interacting agents coordinates decisions to reach a higher-level global objective. An agent is an entity that has knowledge of its physical space (location), can modify its environment (capability) and can learn from experience (memory) [22]. Use of agent-based solutions in multiple infrastructures simulation is driven by the idea that infrastructures are distributed entities and their coordination and decision-making can be most appropriately represented by a community of autonomous, intelligent agents that act like Complex Adaptive Systems (CAS) [1, 22]. One additional benefit of this approach is the observance of "emerging behaviors" that are not predictable by the knowledge of any single agent. Despite this conceptual similarity and apparent benefits, agent based frameworks have some important limitations. The most significant of these is scalability. It has been observed that the performance of multi-agent systems starts to degrade exponentially when the number of agents in the simulation exceeds a certain threshold of a few hundreds agents [92]. This is an important limitation because thousands of agents may be required for a realistic representation of the critical infrastructures in a certain geographical location [15]. Another potential disadvantage of agent-based models is that the complexity of the implemented system tends to obscure the underlying assumptions and can distort the results significantly [93].

To address the difficulty of a multiple infrastructure simulation, Marti et al. [5, 10] have decoupled the problem into two distinct functional domains; physical relationships between the infrastructures that are considered to exist in the physical layer, and human decision factors that are considered to exist in the human layer. Marti et al. proposed the Cell-Channel model [10] as a generic approach for modelling physical layer inter-dependencies among the critical infrastructures. According to this model, conceptual

entities of different infrastructures are mapped to a single equivalent semantic on the assumption that all infrastructures can be represented by two kinds of components, cells and channels. Based on the type of infrastructure, different types of service tokens are delivered through channels from one cell to another. This unified modelling paradigm provides a single intuitive framework to understand complex interdependent behavior of multiple infrastructures. Lu Liu (Lucy) [53] built an initial version of the cell-channel model Infrastructure Interdependencies Simulator (I2Sim). Liu used MATLAB/Simulink [94] functional blocks to represent the cells' functionality. Her solution-algorithm also used capabilities of the Simulink framework.

In the present implementation phase of I2Sim developed in this thesis, we have used the matrix partition-based technique named 'Multi-Area Thevenin Equivalent' (MATE) [25]. The decision to use the MATE model was guided by Dr. Marti's research group's experience with the real-time simulator 'Object Virtual Network Integrator' (OVNI) that simulates large-scale power system networks using the MATE solving algorithm [25, 31, 95]. Our MATE-based implementation of I2Sim critical infrastructure framework is an important extension of the MATE concept. Because the MATE model is based on electrical quantities, Kirchhoff's voltage or current laws are applicable in the original definition. However, present solution technique generalizes Kirchhoff's current law for all kinds of flow based, multiple input, multiple output networks where a Thevenin-equivalent solution can be applied. Additionally, a MATE-based formulation of infrastructure matrix gives interdependency relationship between different infrastructure elements as off-diagonal matrix entries. This is a significant advantage to understand and analyze critical infrastructure network's interdependent behavior. The MATE algorithm can easily be parallelized and can solve a large-scale networks very efficiently [96]. As such, MATE is considered an efficient alternative to the existing agent-based simulation frameworks. This approach can also be used to improve the performance of other simulators. While using the MATE model for multiple infrastructure simulation, we made several important decisions regarding its generalization. This chapter gives a formal overview of those concepts. We also discuss some of the simulation results obtained with the implemented simulator.

4.1 Related Work

Rinaldi et al. categorize [1] interdependencies among critical infrastructures into four principal classes: physical, geographical, cyber and logical interdependencies. There are several critical infrastructure simulators [13] available to simulate these four types of interdependencies in varying degrees. In terms of Rinaldi's classification, I2Sim's cell-channel model encapsulates the physical layer interdependencies, where output of one infrastructure depends on the material flow from other critical infrastructures. To have a contextual background for the Cell-Channel model, in this section we give a brief description of the approaches used in the four physical infrastructure simulators that are available from the published literature.

Min et al. at Sandia National Laboratories developed the Critical Infrastructures simulator IDEF0 [97]. The core of IDEF0 is based on a System Dynamics model, and includes a decision support front-end tool that is based on non-linear optimization techniques. The system dynamics approach was introduced by Forrester at MIT in the early 1960s, and provides a methodology to study and manage complex feedback systems. There are two entities in this model: the Stocks are the accumulation of resources in the system and the Flows are the rates of change that alter the Stocks. Stocks and Flows are connected through feedback loops. IDEF0 is based on this methodology and can show cause and effect relationships among different infrastructures. However, the descriptions of these infrastructures in IDEF0 exist as high-level, abstract entities, like power supply, water supply, economy [97]. Although IDEF0 shows material flow relationships among these infrastructures, they are not connected to any particular physical entity. In addition, the system-dynamic model uses differential equations to express these relationships, which makes it computationally costly and requires smooth functions to define the infrastructure states.

Panzieri et al. [22] implemented an agent-based model called Critical Infrastructure Simulation by Interdependent Agents (CISIA). In CISIA, the agents' dynamic behavior is described by Fuzzy Logic. This agent-based approach helps, because Panzieri receive qualitative data regarding interdependencies from the service providers. The agents in the CISIA simulator exchange three types of information with each other. These are: the operational level (OL) that gives the capability of the system to perform the required job, the requirement level (R) that shows the target to reach, and the fault level (F) that shows

if the system is in a fault condition. Each type of information exchange is represented by a $n \times n$ incidence matrix using fuzzy numbers, where n is the number of agents in the simulation. Physical, geographical and cyber interdependencies defined by Rinaldi et al. [1] can be simulated in this framework.

The Critical Infrastructure Modelling System (CIMS) simulator was developed at Idaho National Laboratory [98]. In CIMS, each physical infrastructure is represented by an agent. In the simulation framework, interdependencies between infrastructures and their relationships are represented as connected graphs. The simulator works like a war game software and can express interdependencies as highly interactive visual interaction. Satellite images, GIS data and live sensor inputs can be incorporated into the CIMS framework.

The Electric Power and Communication Synchronizing Simulator (EPOCHS) [99] was developed by integrating three domain-specific simulators (electrical power transmission PSCAD, electrical distribution PSLF and data communication NS2) using an agent-based framework. The idea is to build a federated simulation-framework where each infrastructure has detailed knowledge of their specific domain. The coordination between different simulators is maintained by a module named Runtime Infrastructure (RTI). The RTI routes messages between components and maintains time synchronization throughout the system. Simulation results related to the interdependencies of the electrical network and the data communication network were presented in [99].

4.2 The Cell-Channel Model

The Cell-Channel model proposed for the I2Sim framework by Marti et al. [10] captures physical interdependencies among different critical infrastructures using a precise mathematical description. The conceptual entities of different infrastructures and their interactions are mapped to a single equivalent semantic. Components defined in the physical layer can interact with the decision-making layer through event forwarding mechanisms. This model has the following five components :

- **Cell:** A cell is an entity that performs a function. For example, a hospital is a cell that uses input tokens, such as, electricity, water, medicines, etc. and produces output tokens, such as, patients served.

- **Channel:** A channel is a means through which tokens flow from a generator cluster to a load cluster.
- **Token:** Tokens are goods and services that are provided by some entity to another entity that uses them. These tokens can represent water, electricity, medical supplies, etc.
- **Cluster:** A cluster is a group of one or more cells (it is called node in [10]). Clusters reduce the modelling granularity and give a mapping to the MATE model [25, 31]. Two clusters are separated in time or space and are connected by channels. Each cluster generates and/or consumes tokens.
- **Control:** These are Distributor and Aggregator units. They change their state based on the events received from the decision-making layer.

The generation and flow of tokens between different entities is determined by the physical capability of each of the cells (e.g., power generation capacity, or water supply capacity), their environmental constraint (damage to cells or delay in a transportation channel) or human decision factor (e.g., redirection of water supply to a hospital rather than to a swimming pool). The operational characteristics of each of the cells can be represented by state variables that are algebraically related to the dynamic behavior of the infrastructure network (e.g. token flow, cell capacity change, etc.). The linear part of the cell functions can be characterized by a Leontief input-output model [10, 52]. However, since cells may also have non-linear input-output relationships, the standard Leontief model is extended to handle non-linear cases. In the extended model, part of the cell's functionality is characterized by linear network equations. Non-linear functionality is encapsulated within a block that characterizes the relationship between different output levels to various input conditions. The nonlinear block gives a set of coefficients to the linear equations. These coefficients are calculated at every time-step, so that the overall cell model becomes a set of linear equations for a particular time instant (piece-wise linear around the operating point), for any combination of linear and nonlinear elements. The beauty of this approach is that it makes the design of the entire cell much simpler and the simulation much faster. These equations are grouped into a matrix called an interdependencies matrix in I2Sim (or G-matrix, in its power systems counterpart). This relationship is similar to the flow of current in an electrical network. Since, the whole network is a combination of many

cells and channels, their unification can form a very large interdependencies matrix at the highest level. However, I2Sim's mapping to the MATE framework enables us to partition the interdependencies matrix into smaller blocks and then further reduce these blocks into simpler equivalent forms (Thevenin-Equivalent forms). Each of these blocks can be solved on separate computers in a very efficient manner [25, 95]. The MATE-based description of an infrastructure matrix has the following form (using power system concepts):

$$\begin{bmatrix} A & 0 & 0 & p \\ 0 & B & 0 & q \\ 0 & 0 & C & r \\ p^t & q^t & r^t & -z \end{bmatrix} \begin{bmatrix} v_A \\ v_B \\ v_C \\ i_\alpha \end{bmatrix} = \begin{bmatrix} h_A \\ h_B \\ h_C \\ -V_\alpha \end{bmatrix} \tag{4.1}$$

where:

$[A]$, $[B]$, and $[C]$ are cluster coefficient matrices.

$[p]$, $[q]$ and $[r]$ are clusters' current (tokens) injection vectors

$[z]$ is a matrix of links' Thevenin impedances

$[h_A]$, $[h_B]$ and $[h_C]$ are vectors of clusters' accumulated currents (tokens)

$[V_\alpha]$ is a vector of link Thevenin voltages

$[v_A]$, $[v_B]$ and $[v_C]$ are subsystems' nodal voltages

$[i_\alpha]$ is a vector of link currents (tokens)

Modelling of cells and channels and their solution algorithm is at the core of the I2Sim implementation developed in this work. A formal description of these concepts is discussed in the following section.

4.3 Infrastructure Simulation in the I2Sim Framework

The present implementation of the I2Sim simulator solves the Cell-Channel model for infrastructure networks using the MATE solution algorithm. The MATE algorithm was originally implemented in the OVNI power system simulator [25]. I2Sim extends OVNI's solution procedure originally developed for the power system infrastructure to multiple infrastructures. The I2Sim framework can be described as a time-driven discrete event simulation architecture [100], where simulation states change due to a sequence of chronological events. Each event occurs at a particular time step and changes the state of the simulated infrastructures. In this section, we describe the components of the I2Sim

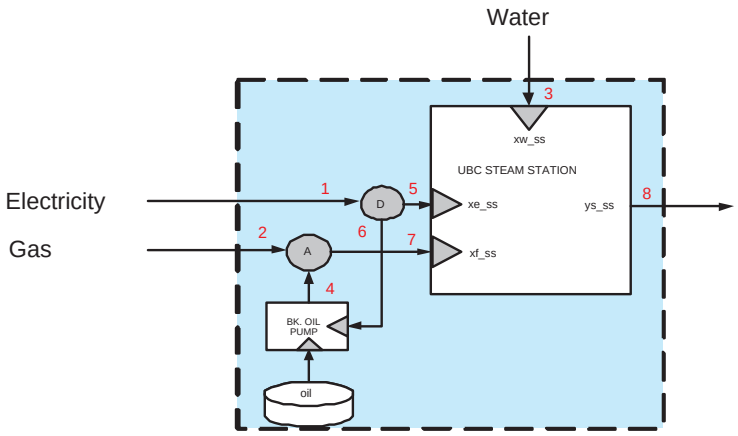


Figure 4.1: Steam Station Model

framework implemented in this work and explain how they extend OVNI's implementation of the MATE model [25].

4.3.1 Models of Cells, Channels and Infrastructure Networks

Cells and channels are the basic infrastructure elements that form the core of Cell-Channel model (Section 4.2). An example of cells are different physical infrastructure entities that include powerhouses, substations, steam-stations, hospitals etc. Figure 4.1 shows a steam-station cell in the UBC campus as represented in I2Sim. The steam-station cell has a non-linear functional block that converts water to steam using the electricity and gas energy. There are three external inputs connected to the cell. The first input (port #1) is electricity. It is connected to a human decision element as distributor, which divides the electricity between the non-linear block and the backup oil pump. The second external input is connected to the fuel supply (gas), which is also connected to a human decision element labeled as aggregator. In the absence of external fuel supply, backup oil supply is available by turning on the aggregator. The third external input is water that is directly connected to the non-linear block inside the cell. The relationship of how different input tokens produce an output token is contained within the functionality of the non-linear block, which is described by the Human Readable Table (HRT) [53]. The HRT for the steam-station describes the amount of water, electricity and gas required to produce a certain amount of steam.

According to the cell-channel model, the cell's functionality can be described by a series of linear equations [10], where the contribution from the non-linear block is adjusted by using scalar coefficients that are calculated for that particular instant of time. For the steam-station cell, the equations are:

$$\begin{aligned}
&x_{e1} = u_{e1} \\
&x_{g2} = u_{g2} \\
&x_{w3} = u_{w3} \\
&x_{o4} - M1 * x_{e6} = M1 * u_{o4} \\
&D1 * x_{e1} = x_{e6} \\
&D2 * x_{e1} = x_{e5} \\
&A1 * x_{g2} + A2 * x_{o4} = x_{g7} \\
&M2 * x_{w3} + M3 * x_{e5} + M4 * x_{g7} = y_{stm}
\end{aligned} \tag{4.2}$$

In Equation 4.2, x represents the input, y is the output and u is the external connected sources. Subscript e denotes electricity, g gas, o oil and w water. The number following each subscript represents the node # from which the tokens originate. Coefficients $A1$, $A2$ are for the aggregator, $D1$, $D2$ for the distributor and $M1$, $M2$, $M3$, $M4$ are normalizing coefficients for the non-linear block. Equation 4.2 can be written in the following matrix form:

$$\begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & -M1 & 0 & 0 \\
D1 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\
D2 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\
0 & A1 & 0 & A2 & 0 & 0 & -1 & 0 \\
0 & 0 & M2 & 0 & M3 & 0 & M4 & -1
\end{bmatrix}
\begin{bmatrix}
x_{e1} \\ x_{g2} \\ x_{w3} \\ x_{o4} \\ x_{e5} \\ x_{e6} \\ x_{g7} \\ y_{stm}
\end{bmatrix}
=
\begin{bmatrix}
u_e \\ u_e \\ u_e \\ -M1 * u_{oil} \\ 0 \\ 0 \\ 0 \\ 0
\end{bmatrix} \tag{4.3}$$

Any cell in the I2Sim model can be written in the form of Equation 4.3. In compact notation, Equation 4.3 is $[A][x] = [h]$, where $[h]$ represents the vector of the external or history sources.

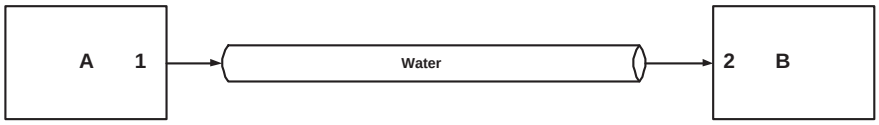


Figure 4.2: I2Sim Channel Model

Cells are connected to each other through channels. In Figure 4.2, water (token) flows from cell A to B through a channel connected between nodes 1 and 2. There are two types of channels in I2Sim. The first type, labeled 'channel', connects cells located within same cluster. A Cluster is a group of cells forming a sub-network (subsystem in OVNI) and have a Thevenin equivalent in the MATE model [25]. The second type of channel is called a 'link'. A link is a connection between two cells located in two different clusters. In other words, Thevenin-equivalents are calculated against links. Clusters are mapped to the MATE model as blocks of coefficient matrices as shown in the Equation 4.1. Channels and links can be loss-less, lossy, and with or without time-delay properties. If we send $x(t)$ tokens from A and receive $y(t)$ tokens at B, then for the lossy channel, we have:

$$y(t) = \alpha x(t) \quad where\ (0 \leq \alpha \leq 1) \tag{4.4}$$

where α is the loss factor. If we further generalize Equation 4.4 to inlcude a time delay, we get the following general form of the channel model, where τ is the time delay [10]:

$$y(t) = \alpha x(t - \tau) \tag{4.5}$$

There are two subtle differences for the flow of tokens in I2Sim compared to the flow of electricity in MATE. In MATE, electrical transmission is driven by voltage and the flow of current is bidirectional. However, for the cell-channel infrastructure model, the flow of tokens is governed by human decisions and these flows are unidirectional. Because of these properties an infrastructure network is represented by a directed graph $(N,\ L)$ where a set of vertices's N denotes cells and a set of edges L denotes a collection of channels connecting two cells. For the I2Sim implementation in this thesis, channels are assumed to be ideal, and all loss and delay are ignored. However, links are treated differently. Loss and delay are calculated for them. Links have a separate location in the MATE matrix formation (Equation 4.1).

4.3.2 Representation of Interdependencies between Infrastructures

The interdependencies between different infrastructures are non-linear. In the cell model, this is encapsulated in the non-linear block, which represents input-output dependencies among one or more token types. The functionality of the non-linear block is described by the Human Readable Tables (HTR). To establish benchmark cases for I2Sim, we studied interdependencies among different infrastructures within the UBC campus for the last two years. The UBC campus was selected because it has all the attributes of a small city [8]. UBC has its own infrastructure service networks and those are managed by UBC service personnel. We have compiled different HRTs for UBC's infrastructures to determine input-output relations between these networks. Expressing interdependencies in terms of a table is a valid approach from a systems engineering viewpoint [52]. These tables were constructed through active support from the infrastructure operators by collecting information related to input and output specifications, operating assumptions, backup facilities, management procedures and practices, and other physical and environmental constraints. System states and control parameters were determined through a number of interviews and by conducting different experiments at the infrastructure premises. For instance, building earthquake sustainability was determined through vibration tests at different infrastructure facilities [101]. Lu Liu (Lucy) in her master's thesis [53] documented interdependency models of the hospital, substation, steam station, water station and powerhouse for a reduced-scale test case of the UBC campus (Section 4.5.1). Table 4.1 shows a condensed form of the HRT for the steam-station [53]. As the HRT shows, the steam-station cell has three inputs, electric power (100kW), water (16.7 m^3/hr), Gas (160 GJ/hr) and has one steam output of 115 klbs/hr. The HRT also shows that the relationship between inputs to output is linear (single range: 0-100%).

Table 4.1: HRT Table of UBC's Steam Station Cell

	Input of the Steam Station Cell						Output
	Power	fun(p)	Water	fun(w)	Gas	fun(g)	Steam
Type	float	float	float	float	float	float	float
Range	0-100%	0-1	0-100%	0-1	0-100%	0-1	0-100%
Unit	kW	-	m^3/hr	-	GJ/hr	-	klbs
Capacity	100	-	16.7	-	160	-	115

HRTs are used directly on the simulator implemented by Liu [53]. However, it has been observed that for the HRT based implementation, it is difficult to interpolate input-output relationships during run-time, especially for higher dimensional cell models. This problem was addressed by DeTao [102]. He developed a method to calculate continuous functions from HRTs. The MATE based implementation of I2Sim uses functional relationships from DeTao's model. This relationship is represented by the following function, where $x_1, x_2, x_3, ...$ are inputs and y is the output to the non-linear block at any particular time instant t :

$$y(t) = f(x_1, x_2, x_3, ..., t) \tag{4.6}$$

4.3.3 OVNI Solution Model

The OVNI simulator for electrical systems [25] introduced the MATE concepts that form the core solution algorithm of the I2Sim implementation in this thesis. The design philosophy of the OVNI simulation framework is to partition the solution of a large-scale power system network into the solution of smaller subsystems (clusters) plus the solution of the links joining the subsystems [25, 103]. In this section we give a brief description of how the MATE model (Equation 4.1) is solved in OVNI. An OVNI simulator is driven by a master clock, where time t starts from 0 and is incremented at time steps $\triangle t$, i.e. $t = n\triangle t$, where $n = 0, 1, 2, ...$ The simulation ends at time T_{end}. At time $t = 0$, the OVNI simulator reads the electrical network description from the input file, creates an internal representation of the various matrices and starts the simulation. At each point in time $n\triangle t$, the network states are evaluated, updates are made to the Thevenin-equivalent state variables, and results are calculated. These steps are described in Algorithm 1.

The OVNI solution framework does not allow for the internal state of each of the elements to be changed arbitrarily during the simulation. For instance, the user may not change the value of a resistor or a capacitor during run time. However, such a capability is vital for the infrastructure simulator. The infrastructures' internal states may change during a run-time in order to model, for example, the damage conditions produced by a disaster event. Such a parameter change can significantly affect the simulation outcome. To accommodate this dynamic capability, we extended I2Sim's components by including decision elements and added external event forwarding mechanisms.

Algorithm 1 OVNI Simulation Algorithm for Electrical Networks

```
Input:
  - cell, channel, initial-state from input file
Create:
  - A,B,C,..        {cluster coefficient matrices}
  - p,q,r,..        {current injection matrices}
  - p^t,q^t,r^t,..  {transpose of current injection matrices}
  - v_A,v_B,v_C,..  {are subsystems' nodal voltages}
  - h_A,h_B,h_C,..  {vectors of history sources}
  - z               {matrix of links' Thevenin impedances}
  - i_α             {vector of link currents}
  - V_α             {vector of link Thevenin voltages}
t = 0
while t ⩽ T_end :
  update h_A,h_B,h_C,..
  update z
  update V_α
  calculate i_α
  calculate v_A,v_B,v_C,..
  t = t + △t
end while
```

4.3.4 I2Sim Event Scheduling

The I2Sim solution model was extended from OVNI by introducing two types of decision elements: Aggregators and Distributors. We discussed in Section 4.3.1 (Equation 4.2) that aggregators and distributors are linear elements in the steam-station cell and the range of their values varies between 0 and 1. These values change due to the events received from the decision layer. Decisions are propagated as events to the specific decision element. The numerical solution of the I2Sim framework is calculated at a sequence of time points $t_1, t_2, ...t_n, t_{n+1}, ...$ Events are introduced into the network at any of these times. Events modify the internal state of the recipient cell. We can formally define an event as follows:

Definition: An event e is a 4-tuple $(cID, elemID, t, value)$ where cID is the target cell ID (e.g. steam-station), $elemID$ is the target decision element (e.g., aggregator) within the cell, t is the event triggering time and $value$ is the new value the decision element acquires because of this new event at the triggering point.

I2Sim events come from the database, and are kept in a priority queue, where the event with the lowest firing time executes first. In other words, events are sorted according to their triggering time in the priority queue. The event queue module periodically checks the database for new events and updates the priority queue if any new event is available. The scheduler module drives the simulator over the time window. While doing this, the scheduler checks the front (top) of the event queue at each point in time and adopts one of the following three strategies based on the triggering time of the first element:

$process(e)$ Event's time matches the current simulation time (has the same triggering time as of the present time point), so forward the event to the right cell for processing. Remove the event from the event queue, and check the next event (since multiple events may have same time-stamp).

$skip(e)$ Event has larger time-stamp than the present time point, so skip checking additional events in the queue at this time point.

$remove(e)$ Event's time-stamp is smaller than the present time point (the user entered the event with a time-stamp less than the current time). This means the simulation clock has passed this point. Therefore, toss the event out and check the next event from the queue.

4.3.5 I2Sim Solution Model

It was noted in the previous sections that the I2Sim solution model differs to some extent from OVNI due to the differences between infrastructure networks and electrical networks. The difference also arises due to the introduction of non-linear blocks and decision elements in the I2Sim cell model. Additional complexity in the solution framework has been added for the event processing mechanism. The I2Sim solution model is described in Algorithm 2. Since there is no concept of voltage in the Cell-Channel model, the subsystems' nodal voltages in the MATE model represent token flows. Also, the tokens flowing in the infrastructure networks are unidirectional. So, p^{tx},q^{tx},r^{tx},.. are not the exact transpose of p,q,r,.., the token injection matrices. Since events change the cell's internal states, this implies recalculation of cluster coefficient matrices at every time point. Careful study of Algorithms 1 and 2 shows that they both conform to the MATE model (Equation 4.1). Therefore, the correctness of the solution of Algorithm 2 can be derived from [25].

With the present solution technique, we have assumed uniform step size ($\triangle t$) for all subsystems. More specifically, we have used a global clock that advances by a fixed time-step up to the end of the simulation time. This may not always be the case for infrastructure networks, where different infrastructures may have different levels of time sensitivity [89]. For non-uniform time steps, the MATE based solution model can be extended using Moreira and Marti's Latency technique [27]. For infrastructure networks, we have proposed a more generalized solution technique in Chapter 5.

4.4 I2Sim Implementation

The I2Sim simulator in this work was implemented in C++ using object-oriented design techniques. The simulator is a multi-threaded application, where different components (core modules, input module and output module) run in parallel. I2Sim was compiled and tested on Windows and Linux machines using Microsoft Visual C++ and GNU C++ compilers. Description of the implementation architecture of these modules is given in this section. Additional information regarding implementation-related issues can be found in the I2Sim User Guide [104].

Algorithm 2 I2Sim Simulation Algorithm for Critical Infrastructures

```
Input:
  - cell, channel, initial-state from input file
  - events e from database
Create:
  - A,B,C,..      {cluster coefficient matrices}
  - p,q,r,..      {token injection matrices}
  - p^tx,q^tx,r^tx,..   {partial-transpose of token injection matrices}
  - v_A,v_B,v_C,..  {are subsystems' nodal token flow}
  - h_A,h_B,h_C,..  {vectors of history sources}
  - z       {matrix of links' Thevenin impedances}
  - i_α     {vector of link token flow}
  - V_α     {vector of link Thevenin voltages}
  - event_queue {priority queue to hold events}
t = 0
while t ⩽ T_end :
  get e ∈ event_queue {get event from the event queue}
  if process(e) : {forward events to the corresponding element}
    M1, M2, . ≡ f(x_1, x_2, ., t) ∈
A, B, . {coeffs for non-linear blocks}
    update A, B, C.... {recalculate cluster coefficient matrices}
  end if
  update h_A,h_B,h_C,..
  update z
  update V_α
  calculate i_α
  calculate v_A,v_B,v_C,..
  t = t + △t
end while
```

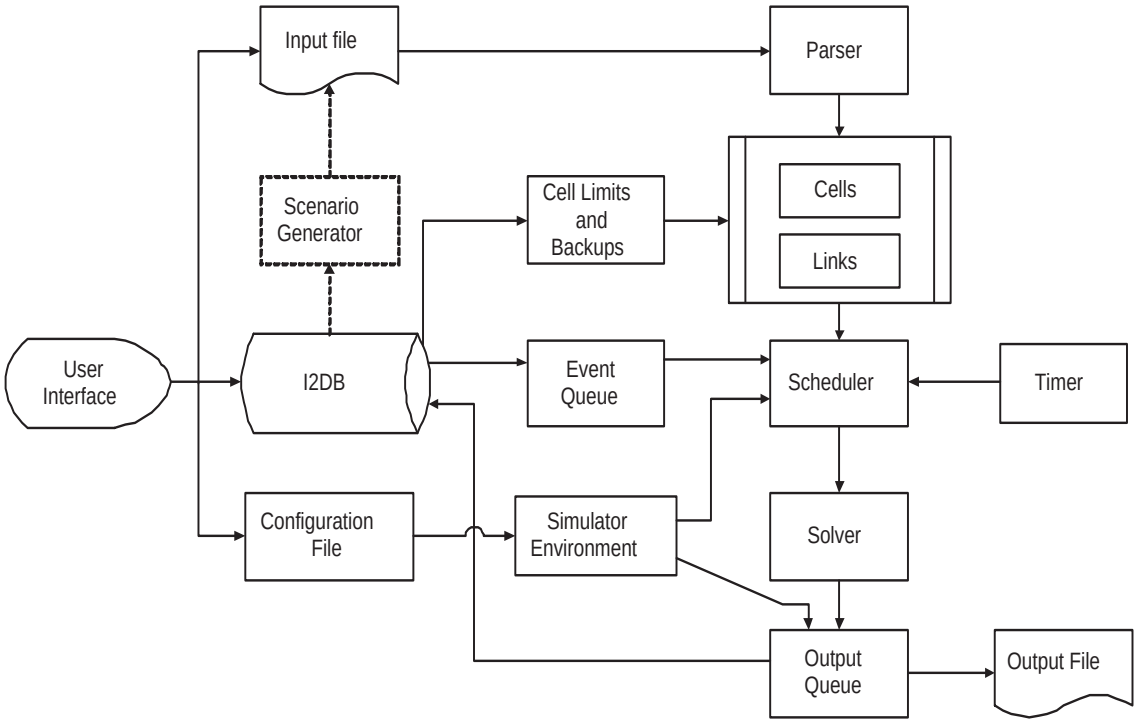


Figure 4.3: I2Sim Implementation Architecture

4.4.1 Core Modules

The solution techniques discussed in Section 4.3 form the core components of the I2Sim implementation architecture. These include modules for infrastructure elements (cells and links) and modules for solution framework (solver, event queue and scheduler). Other than core components, there are several other input and output related modules, such as an input file parser, a program options scanner, an output queue, database input/output utilities, etc. The relationship among the different functional blocks is shown in Figure 4.3.

I2Sim has an internal clock that keeps track of the simulation time and propagates simulation events with the help of a scheduler. The scheduler periodically inspects the events queue, dispatches events from the queue to the corresponding cell and calls the solver module to calculate the solution based on the modified state information. Since I2Sim is a multi-threaded application, the event queue is accessed by both the database search process and by the scheduler, which may lead to deadlock and inconsistent system states within the simulator. To avoid this problem, a thread-safe locking mechanism was implemented using the Boost [105] thread library that ensures that only a single process will access the event queue at any particular time, while another process is waiting. The Core

module is the main execution point of the simulator. It reads the input file and instantiates cells, channels and links. Following that, the core module instantiates the solver object, starts the event queue and the scheduler, and then sets up the output buffer. Once the environment is prepared, the simulation starts and continues according to the master clock. When the simulation ends, all threads are closed and all memory is released.

4.4.2 Input Module

A simulation scenario requires different combinations of cells and channels, and their aggregation (clusters). The description of the infrastructure elements and their connectivity comes from the input file. The I2SIM input file is written using simple statements, which can be formalized according to the Extended Backus Naur Form (EBNF) [106]. A parsing module is responsible for input file processing. The parser interprets a character stream received from the input file and creates an in-memory representation of the cells, channels and clusters for a valid scenario description. A EBNF specification-based small language pre-processor (lexer) works with the parser to verify the input file. The EBNF specification for I2Sim input file is shown below:

```
program :=  block | comment;
block := main_block | cluster_block ;
comment := +((comment_p("#") | comment_p("/*", "*/"));
main_block := 'main' '{' param_list '}';
cluster_block := start_block statement_list;
start_block := keyword identifier param_list ;
statement_list := statement *statement;
statement := keyword identifier;
param_list := number *(',' number);
identifier := (alpha_p | '_') *(alnum_p | '_'));
keyword :=  "cluster", "link", "channel", "power_house",
             "water_station", "hospital", "steam_station",
             "substation",  "token_source", "null" ;
```

This formal structure of the input file simplifies checking and validation. For any large-scale infrastructures simulation, writing such an input file can be difficult and error prone. As such, it is planned to generate these input files from the values stored in the database in the near future. The proposed scenario generator is shown in dotted line in Figure 4.3.

4.4.3 Output Module

In regular mode, the simulator output is written to the database. It is generally inefficient to write results to the database as soon as they are generated. A more efficient approach is to gather the results in a buffer (output queue) and periodically write them in the database. I2Sim's result queue serves this objective. It writes the result data from the queue to the database in bulk mode periodically at a user-defined time interval. This makes I2Sim's operation particularly efficient. Since the output buffer is accessed by both simulator and the writing modules, a locking mechanism is employed to give exclusive access to one of the two processes. The results queue can be written to a file instead of a database. However, in that case, the operation is aperiodic (i.e., the result is written to a file at the end of the simulation).

4.5 Simulation and Validation

We have compared our MATE based I2Sim implementation with the version developed by Lu Liu (Lucy) [53]. Multiple simulation results have validated our approach. We have used Liu's results as a benchmark, because those were validated against real disaster scenarios. In this section, we present the results for a scenario discussed in Liu's thesis. The scenario is a reduced scale representation of the important infrastructures in the UBC campus. The present scenario description has the same initial conditions and has undergone the same set of event sequences as Liu's.

4.5.1 Simulation Environment

The test case captures real-life interdependencies that exist between critical infrastructures in the UBC campus. The test scenario consists of five cells in the UBC campus that include the electrical substation, powerhouse, water station, steam station and hospital. Figure 4.4 shows these five cells and their connectivity. The functional model of each of these cells can be found in Liu's thesis [53]. However, unlike Liu's implementation, we have used the linear equation based model mentioned in Section 4.3.1. Since in both models, the functionality of the cells is characterized by the human readable tables (HRTs), we have identical input/output relationships. Among these five cells, the Hospital cell is the most complex and has many backup sources and decision elements (Figure 4.4). The other four

cells have relatively simple models. The five cells form a G-matrix (Equation 4.1) within the simulator, according to our implementation of the MATE model. The G-matrix for the five-cell test case generated in the simulator is shown in Figure 4.5. The dark blocks in the left-hand side of G-matrix represent internal states of each of the infrastructure cells. For instance, coefficients A1 and A2 in the hospital block of the G-matrix are related to the aggregator 1, where coefficient A1 is for line # 27 and A2 is for line # 33 (Figure 4.4). The output of the aggregator 1 is an input to a boiler (#43). There would be no interdependency in the boiler energy supply if all fuel were from line #27 (gas supply from the Terasen company). However, to secure the boiler energy supply, there is a backup fuel supply from the oil pump (line # 33). In addition to this, the oil pump is operated through the electricity coming from the line # 38. This backup fuel supply mechanism introduces interdependencies within the hospital system's operation. These interdependencies are represented by the off-diagonal entries in the G-matrix. This mathematical formulation of interdependency relationship in the G-matrix is a powerful construct and allows us to perform sensitivity analysis to identify the level of interdependencies and probable vulnerable points. Similar mathematical formulation is also shown for distributor 1, which is splitting input from the line # 40 into three separate lines # 38, # 49 and # 58. The entries D1, D2 and D3 in the hospital block of the G-matrix are related to the distributor 1 (Figure 4.5).

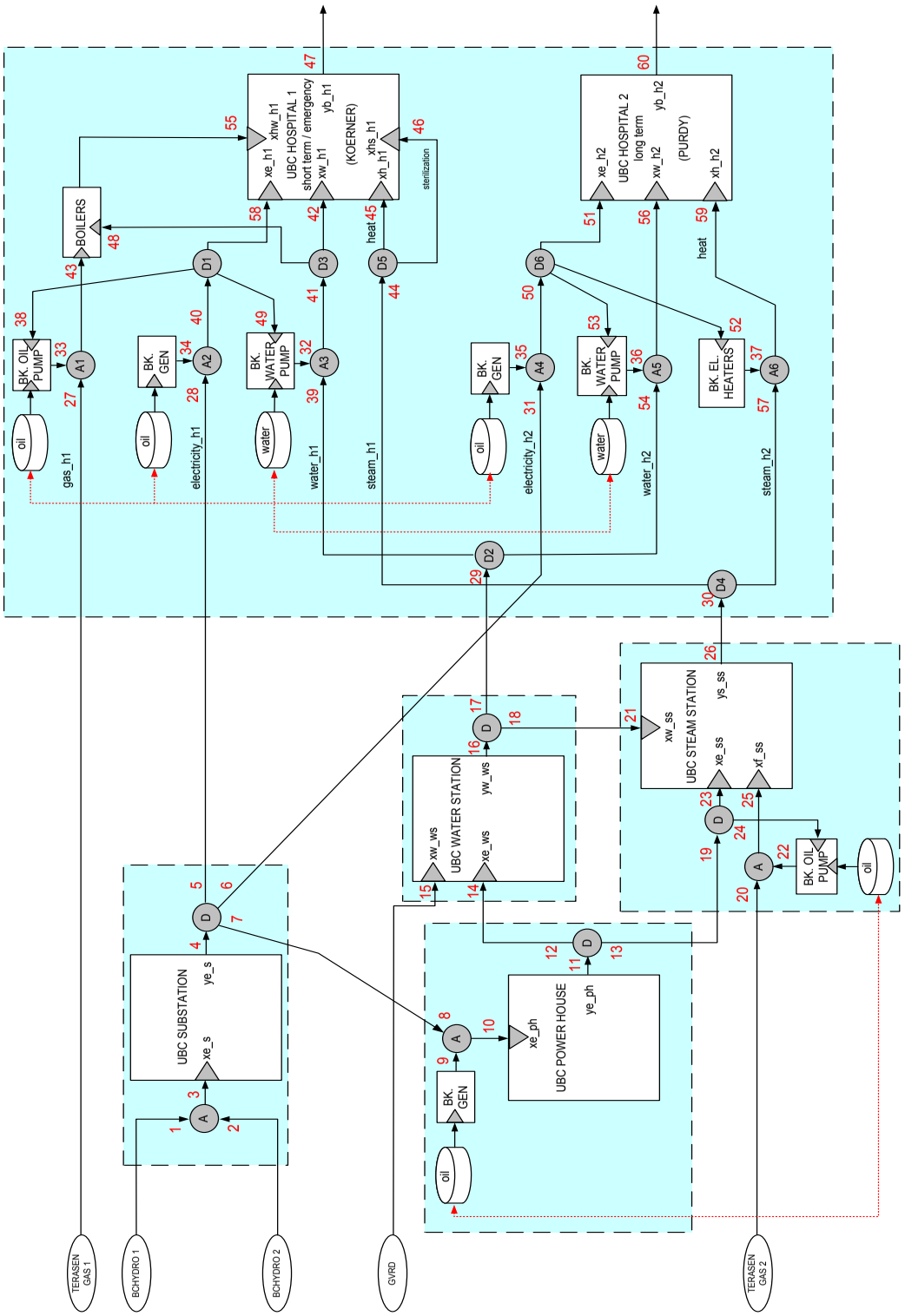


Figure 4.4: UBC's Five-Cell Test Case

The objective of our test case is to assess the impact of failure of different infrastructures on the hospital services. The outputs from the hospital are service capacities in terms of different types of beds (emergency patients, long-term patients, etc.). Other cells provide their respective tokens (electricity, water, gas, etc.) to the systems. Several external sources are also connected to these cells. These are, electricity from BC Hydro, gas supply from Terasen and water supply from GVRD. It is assumed that the capacity of the external sources is unaffected during the natural or man-made disaster. The channels between the external sources and the UBC campus may be affected by any such disaster. The operating capacity of the hospital depends on the external inputs from the different critical infrastructures (e.g., electricity, water, gas and steam) and also depends on the internal variables of the hospital (e.g., doctors, nurses and medicines). Since the performance of one cell is affected by the output of other cells, the overall picture of their interaction can be found through the simulation.

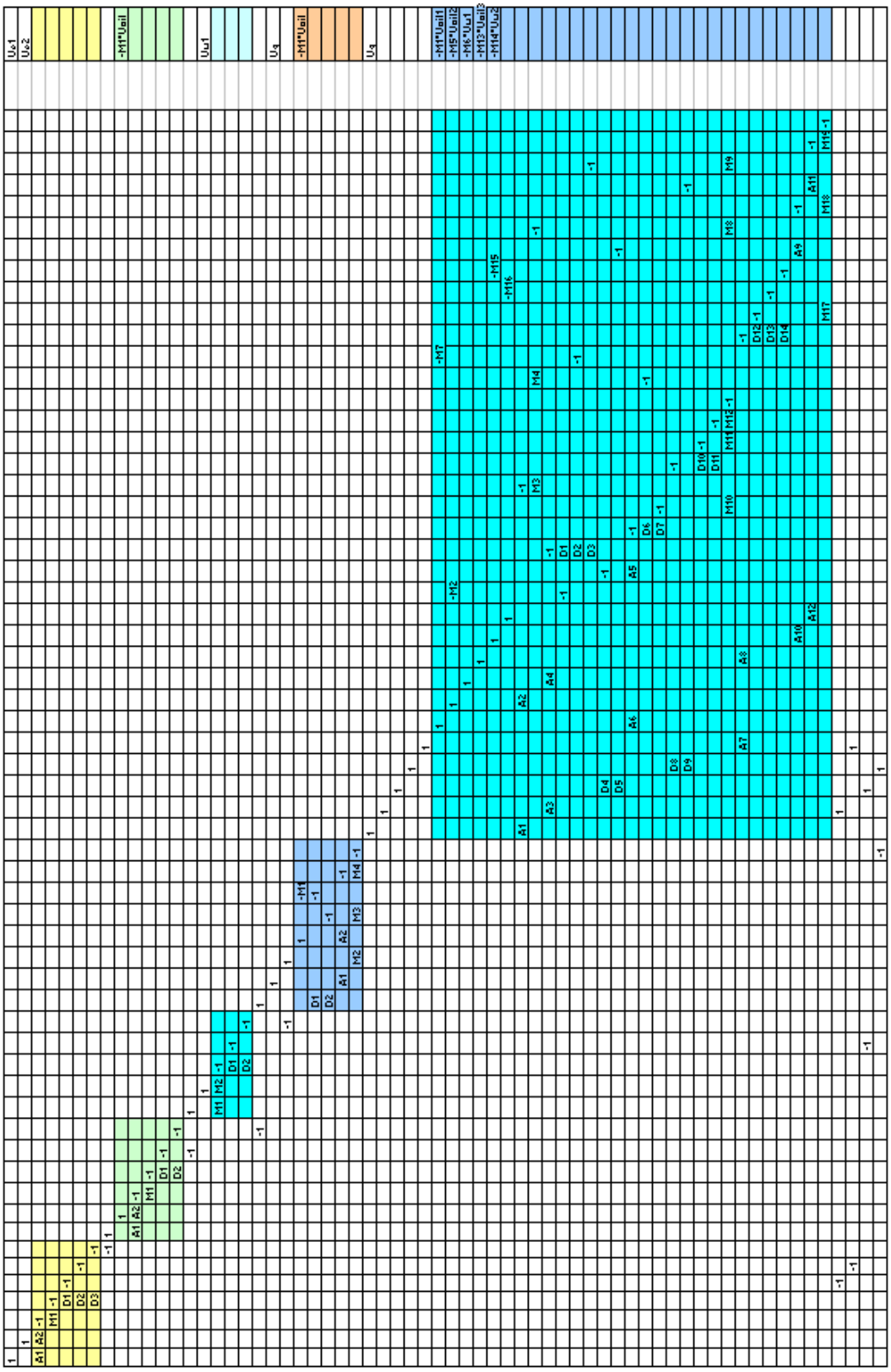


Figure 4.5: G-Matrix of UBC's Five-Cell Test Case

4.5.2 Simulation Scenario and Results

This test case scenario is similar to the first benchmark cases in Liu's thesis [53]. The scenario mimics the impact of the Pacific storm that hit UBC's campus on 19 November 2006. This is a 15 hour long scenario that starts at 1:00 AM 19 November 2006 and ends at 16:00 PM 19 November 2006. We equate the values of all external and internal sources to Liu's test case. The test scenario presented in this section was composed by describing infrastructure connectivity and initial states in an input file and setting events in the I2Sim events table (database). The events represent changes of infrastructure states due to natural or human decisions during the simulation time window. The scenario is described below:

- Initial state: t = t_0, all systems in normal state. [1:00 AM 19 November 2006]
- t=t_0 + 20 (min)
  - A power outage occurs due to fallen trees, which brings down BC Hydro transmission lines.
  - Backup generators start automatically.
- t = t_0 + 40 (min)
  - Water pipe linking water station to the UBC hospital gets broken. Water supply to the hospital is reduced to 0.
- t=t_0 + 70 (min)
  - Backup pumps start manually and supply water from the reservoir.
- t=t_0 + 780 (min)
  - Water pipe is repaired after 12 hours.
  - Backup water pumps are turned off.
- t= t_0 + 860 (min)
  - Power is restored from BC Hydro.
  - Backup generators are turned off.
- t=t_0 + 900 (min)
  - Simulation ends [15 hours - 16:00 PM 19 November 2006]

During this simulation, two major events occur, a power outage and the breakdown of a water pipe line. Figure 4.6 shows the simulation results of the substation cell, which shows power is gone for 14 hours without having much effect on the output capacity of the hospital (Figure 4.8). This is because backup generators start operating as soon as the

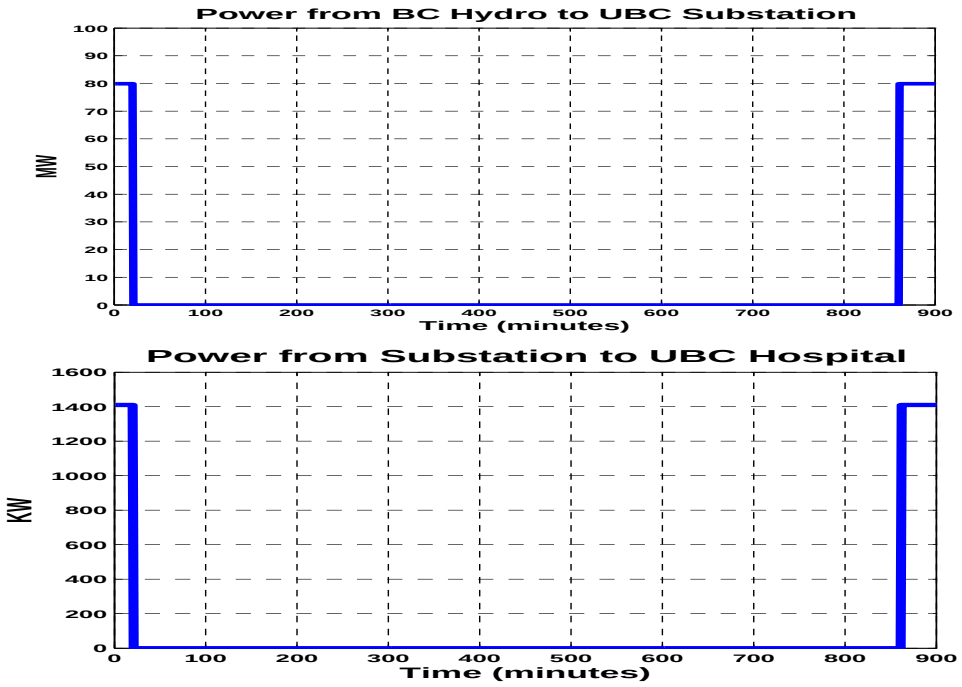


Figure 4.6: Input and Output Capacity of Substation Cell

power is out. However, the water pipe breakage affects the hospital immediately, and as a result, the hospital output capacity directly reduces to zero.

During the simulation, the hospital output (Figure 4.8) shows spikes (sharp drop and rise) at the elapsed times of 20 and 860 minutes. The switching to and from backup generators introduced these sudden changes. These were each 3 minutes delay to study the possible impact of backup sources on the hospital capacity. Figure 4.9 shows these spikes in a zoomed-in view of the first and last 100 minutes of our simulation. The results produced by the MATE I2Sim as shown here are identical to the results obtained in [53]. The results also demonstrate I2Sim's capability to identify interdependencies between infrastructures.

4.6 Chapter Summary

The development of an efficient critical infrastructure simulator with a powerful modelling construct is a challenging task. In this paper, we have introduced an implementation of the I2Sim framework that combines the efficient matrix partitioning of the MATE solution technique and I2Sim's Cell-Channel modelling. We have extended and generalized the MATE model for the case of multiple critical infrastructures and have shown its potential

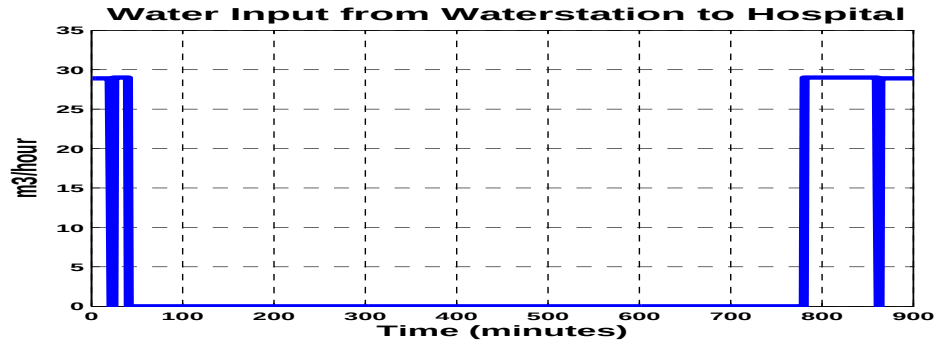


Figure 4.7: Water station Output to UBC Hospital

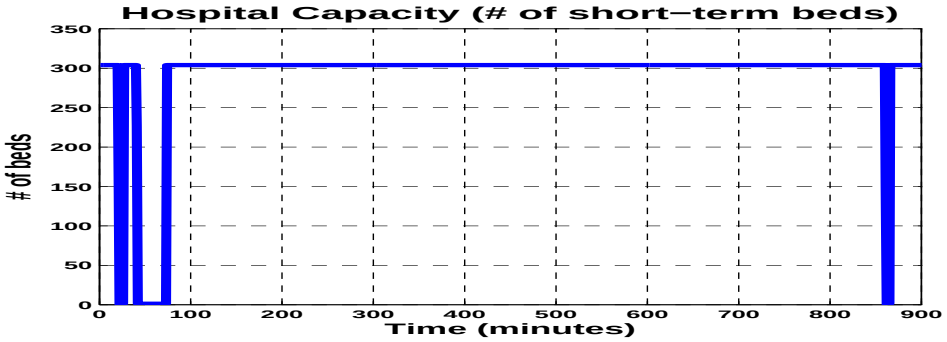


Figure 4.8: UBC Hospital Output Capacity

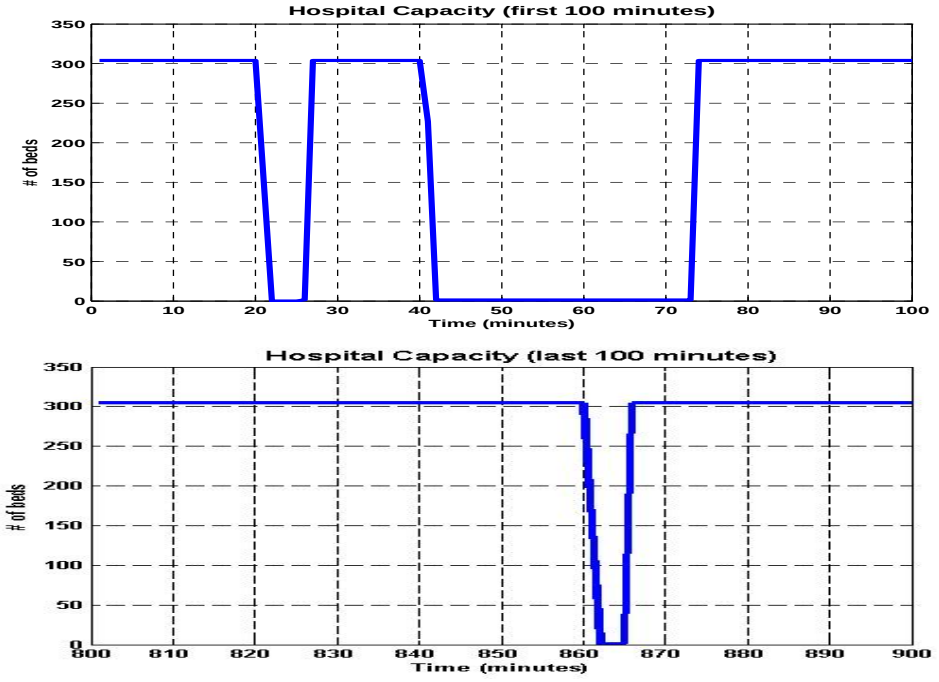


Figure 4.9: Effect of Delay of Backup Generators on Hospital Output Capacity

to solve a general kind of physical infrastructure network. We built a functional simulator and its accuracy has been confirmed by comparing with benchmark test cases obtained in previous work. The MATE algorithm is known for its efficiency to solve large-scale power system networks. This is a new approach to build an infrastructure simulator in a time when parallel hardware and cluster computing are becoming pervasive. The MATE based approach can be an efficient alternative to existing agent based simulation models.

Chapter 5

CITI Interdependency Simulation in I2Sim using a Hybrid Model

There are two important requirements for cyber interdependency simulation. The first is to know the origin of different types of CITI failures and their possible impacts on the critical infrastructures. The second is to represent infrastructure elements and events with their domain specific characteristics in the infrastructure simulator, so that we can observe rational interactions between these components. Reference [11] and Chapter 2 provide an analysis of a large number of CITI related infrastructure failure reports from the public domain. The analysis of these reports shows the origins of the CITI failures, the impacts of these failures in spatial and temporal dimensions and how these failures propagate from one infrastructure to another. In [21] and Chapter 3 we developed a set of empirical functions to formalize cyber interdependencies for different critical infrastructures. These functions are essential components for cyber interdependency analysis and establish relationships between cell outputs in the infrastructure simulator and the corresponding CITI service inputs. The second requirement is addressed by the I2Sim simulator ([5, 20] and Chapter 4) that can faithfully simulate critical infrastructures. The validity of I2Sim has been confirmed through various benchmark tests. In this chapter we focus on techniques to extend I2Sim's capability to simulate cyber interdependencies.

We followed three steps for cyber interdependency simulation in I2Sim. First, we estimated CITI interdependency functions (see Chapter 3 and [21]) and incorporated them into the I2Sim infrastructure cells as an additional dimension (see Section 4.3.2). Second, we have included an IP network simulator named Fluid Flow Model (FFM) [30] to accommodate the contribution of the data communication networks. Third, we have

implemented latency techniques [27] into a hybrid simulation approach for synchronization and results integration between the two simulators. The presence of fast changing (Δt) events (e.g. changes in data communication networks) in between slowly progressing (ΔT) other critical infrastructure events (e.g. human decisions) makes the system characteristics stiff. Simulating all system components with respect to a smaller time-step Δt can provide a correct but inefficient solution, because most parts of the simulated infrastructures are insensitive to the small time-step changes. The latency-based techniques are an efficient yet accurate solution for such multi-rate systems. We have already discussed the development of empirical functions and their integration into our I2Sim implementation in previous chapters and in [20, 21]. In this chapter, we will give an overview of the hybrid models and technical issues related to the joint operation of both simulator (I2Sim and FFM) for cyber interdependency simulation. These include system modelling, synchronization and accuracy analyses. Following these, we present the simulation results of a well-known and well-documented computer network failure case [2–4] that significantly affected the operation of Beth Israel Deaconess Medical Center of Harvard Medical School for four days in November, 2002.

5.1 Related Work

Hybrid model simulation has been studied in different branches of electrical engineering for quite some time. Schwetman [107] proposed a hybrid model (1978) to combine discrete-event simulation with analytic techniques to produce efficient yet accurate system models. In recent years a number of hybrid models have been proposed in the data network simulation community [108–111]. The objective of these techniques is to simulate large-scale data communication networks, where packet level simulation is quite inefficient. In the hybrid model proposed in this thesis, two different kinds of solution frameworks are used to solve the problem and the results are periodically combined together.

Similar concepts are used in the electrical and electronic circuit simulation community. The latency technique is used in this field to coordinate slow time-varying systems and fast time varying systems within large systems simulation. However, in these latency based solutions, both fast and slow subsystems use identical solution techniques. A good overview of latency based techniques and their relative merits can be found in [26]. For EMTP (Electromagnetic Transients Program) [112] programs, on which MATE [25] and

I2Sim [20] are based, the latency techniques have been developed by Moreira and Martí [26, 27]. The main methodology is to use numerical integration methods with different time-steps for different subsystems (fast and slow) according to the accuracy requirements of each subsystem. These latency techniques have been used successfully for MATE-based simulation programs (e.g., OVNI [25]). In the basic form of these techniques, the relationship between the large time step (ΔT for the slow subsystem) should be an integer multiple of the small time step (Δt for the fast subsystem), i.e. $\Delta T = n\Delta t$. In order to ensure stable solutions, both the slow and the fast subsystems are solved together at time instant multiples of ΔT. For the in-between solutions, only the fast system has to be solved. To consolidate the results from the fast subsystem, one history source is needed. Modelling this history source depends on the discretization property of the particular subsystem and also on the numerical integration rule that is used.

We use a hybrid model in this thesis consisting of two different solution models for fast (fine time-scale) and slow time varying (coarse time-scale) subsystems. However, we have used the concepts of Moreira and Martí [26, 27] to synchronize between the two subsystems.

5.2 The Hybrid Systems Model

In critical infrastructure networks, different system components have different timing hierarchies [89]. As such, during critical infrastructures' modelling and simulation, it is often convenient to partition the simulation components into different sets based on their timing granularities. For cyber interdependency simulation, we have partitioned the system components into two mutually exclusive sets; long time-step system components and short time-step system components. This creates a multi-rate systems model. In this hybrid model, long time-step (ΔT) simulation is used to model regular infrastructure events, which include changes in the infrastructure cells and changes in the control elements through human decisions. This part of infrastructure interdependency is simulated through I2Sim. The short time-step (Δt) part can be modeled through a technique that can capture the fast changing events that affect infrastructure cells and can be potentially overlooked in the regular long time-step part of the simulator. For the cyber interdependency simulation, such an approach is especially useful. This is because data communication networks have highly irregular traffic characteristics [113, 114], which can only be captured by domain

specific simulators. For our present research we have used a Fluid Flow Model (FFM) [30] based simulator for data communication network representation. Since many events can occur in the very short simulation time-steps (on the order of milliseconds), this part of the simulator can be expensive to operate. Because a major part of the infrastructure simulation is insensitive to small time-steps, by using our hybrid model we are able to achieve a high level of agreement with the expected simulation results at a significant reduction in computational costs. The architecture and working mechanism of I2Sim has been discussed in Chapter 4. In the following sections, we give a brief overview of the FFM simulator and then discuss integration techniques between I2Sim and FFM models.

5.2.1 Fluid Flow Model Simulator

There are three principal types [110] of data network simulation models in use today. They are based on packet models, fluid flow models and hybrid models. The packet-based approach is the most widely used simulation technique for data communication networks. NS2 [115] is a well-known simulator that falls into this category. Packet-level simulators keep track of all individual packets in the network. As such, they are not very scalable when large networks are considered [110]. In fluid flow models, aggregates of packet flows in a network are considered. Aggregate flows in a network are modeled using ordinary differential equations (ODEs) that describe how packet flow changes in the network [29]. A fluid model simulator has recently been developed by Liu et al. [30] based on the model of Misra et al. [29]. This simulator uses a discrete-time Runge-Kutta algorithm to solve TCP/IP based data network states. This discrete time-step solution approach is similar to I2Sim's and is used as a synchronized solution module for cyber interdependency simulation. Hybrid model data network simulators [108, 109] use features of both the packet model and the fluid model. Similar to packet models, hybrid models also have a problem of scalability and are not considered in the present discussion. In this thesis, we have adopted the fluid flow model simulator developed by Gu et al. [111].

The Fluid Flow Model (FFM) proposed by Misra et al. [29] assumes an aggregate (class) packet flow inside the data communication network. A network is considered as a directed graph G = (V,E), where V is the set of routers and E is the set of links. Network G constitutes a population of N classes of TCP (Transmission Control Protocol) flows. Since TCP is used in 90% of network traffic [116], this model is a fairly accurate representation

of the behavior of communication networks. In this model, stochastic differential equations based on the network description are reduced to a set of ordinary differential equations that can be solved numerically to get deterministic results [29]. The following is the set of ordinary differential equations [30] that describe the network characteristics in this model:

- TCP Window Size:

$$\frac{dW_i(t)}{dt} = \frac{1(W_i(t) < M_i)}{R_i(t)} - \frac{W_i(t)}{2}\lambda_i(t) \tag{5.1}$$

  where M_i is the maximum TCP window size, $W_i(t)$ is the expected window size at time t, $R_i(t)$ is the round-trip time and $\lambda_i(t)$ is the loss indication rate experienced by a flow class i.

- Queue Length:

$$\frac{dq_i(t)}{dt} = -1(q_l(t) > 0)C_l + \sum_{i \in N_i} n_i A_i^l(t) \tag{5.2}$$

  where n_i denote the set of TCP classes traversing queue l, and C_l is the capacity (bandwidth) of the queue.

- Traffic Propagation:
  Departure rate

$$D_i^l(t) = \begin{cases} A_i^l(t), & \text{when} \quad q_l(t) = 0 \\ \frac{A_i^l(t-d_l)}{\sum_{j \in N_i} A_i^l(t-d_l)} C_l, & \text{when} \quad q_l(t) > 0 \end{cases} \tag{5.3}$$

  where $A_i^l(t)$ is the arrival rate and d_l is the queuing delay experienced by the traffic departing from l at time t:
  Arrival rate

$$A_i^l(t) = \begin{cases} A_i(t), & l = k_{i,1} \\ D_i^{b_i(l)}(t - a_{b_i(l)}), & \text{otherwise} \end{cases} \tag{5.4}$$

Liu et al. [30] solve the above equations using a fixed step-size Runge-Kutta algorithm. Their solution works like a time-stepped network simulator for IP networks. Functional steps related to this solver are shown in Figure 5.1, adopted from Figure 2 of [30].

Test results reported in [30] claim respectable accuracy and scalability of this model. We have also found that FFM simulator is quite efficient (can calculate up to 4000 time-steps/sec for our test cases); however, it takes significant time during some of the events'

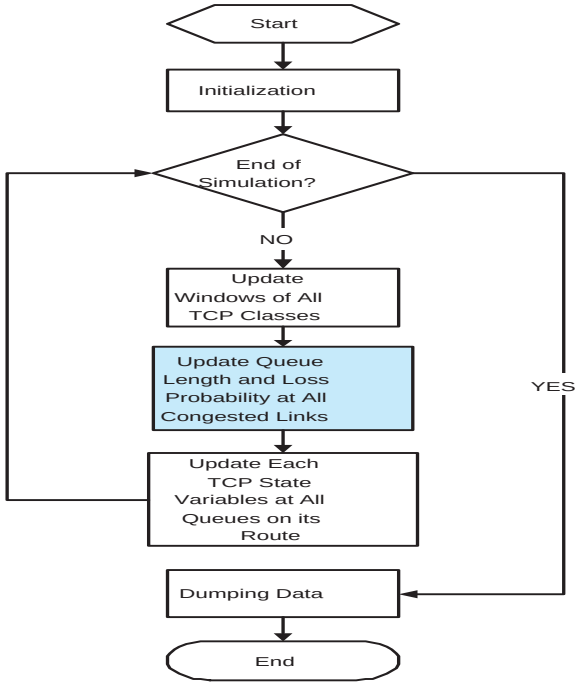


Figure 5.1: Flowchart of Fluid Model Solver

handling (e.g., changing link bandwidth to a very small value during the simulation). However, this FFM's constraint should not be considered as a limitation of the hybrid approach presented in this paper, because other simulation frameworks could be integrated similarly when required to satisfy the domain specific requirements.

5.2.2 Hybrid Model Solution Techniques

Due to its close human interaction in I2Sim, the events are slow by time-varying. As a result, I2Sim may not capture fast time-varying events that occur in some critical infrastructure networks, such as, electrical or data communication networks. If we want to capture the effect of these fast changing networks into the cell-channel simulation model, we need to extend I2Sim's capability with domain-specific layers that can capture the events specific to that domain. For this integration, a hybrid model has to be more flexible than the all-electrical multi-rate model proposed by Moreira and Martí [26, 27], in which the MATE solution applies to all subsystems. The hybrid model has to be loosely coupled in order to integrate the I2Sim solution model [20] with other solution techniques based on domain specific models. Both models have to share information analytically, as well as

through event forwarding mechanisms. Domain-specific networks have to provide dynamic updates to their proxy representations in I2Sim through analytical means. For instance, road networks within I2Sim get dynamic updates analytically from the background road network simulator. However, human decisions or high level events from the foreground I2Sim model are pushed to the background physical network model through extended event handling mechanisms. This constitutes a two-way communication between the foreground and background models. The background network status (e.g., link throughput or delay information) is conveyed to the foreground simulator through an extended I2Sim link model. Correspondingly, the extended event forwarding mechanism propagates the foreground I2Sim events to the background simulator, which then affects the background network states accordingly. In the hybrid model, extended event handling is relatively straightforward, as it is a simple extension of I2Sim's event forwarding mechanism [20], where events that are targeted to the background simulator have separate IDs. Hence, the discussion of event forwarding is omitted from this thesis. On the other hand, the analytical model integration techniques are more complex and are discussed in the following sections.

For the hybrid model, analytical integration and synchronization are achieved through the extension of I2Sim's channel model and latency based synchronization techniques. In the present I2Sim implementation [20], for a channel between A and B with loss-coefficient α and time delay τ, if $x(t)$ is the number of tokens sent from A and $y(t)$ is the number of tokens received at B, then we have the following relationship:

$$y(t) = \alpha x(t - \tau) \quad where\ (0 \leq \alpha \leq 1) \tag{5.5}$$

For a time-varying loss-coefficient $\alpha(t)$, Equation 5.5 can be written as:

$$y(t) = \alpha(t)\, x(t - \tau) \quad where\ (0 \leq \alpha \leq 1) \tag{5.6}$$

We have extended this channel model (Equation 5.6) by incorporating the calculation of the loss coefficient α and delay τ at every time-step (ΔT) with the calculations from the domain specific simulator. The intuition here is that the domain specific simulator can capture the specific network states more precisely and can give us better estimates of the channel properties for interdependency measurements.

The cyber interdependency simulation is a special case of hybrid simulation. For cyber interdependency simulation, the data network status (e.g., link throughput and delay) is

conveyed to the I2Sim simulator through a special CITI link. Since, network delays in the data communication network are typically less than a second [117, 118], assuming $\tau \ll \triangle T$, we have ignored τ for the CITI channel model. Therefore, Equation 5.6 reduces to the following form for the CITI link:

$$y(t) = \alpha(t)\, x(t) \quad where\ (0 \leq \alpha \leq 1) \tag{5.7}$$

For the CITI link, the channel loss coefficient $\alpha(t)$ is directly proportional to the packet throughput of the corresponding link in the backend FFM simulator. If $B(t)$ is the packet throughput [29] in the FFM simulator, then $\alpha(t) \propto B(t)$, that is:

$$\alpha(t) = KB(t) \quad \text{where } K \text{ is a proportionality constant} \tag{5.8}$$

For the hybrid model, $\alpha(t)$ is calculated over the entire $\triangle T$ interval. In FFM [30], the channel throughput $B(t)$ is calculated at every time-step as the product of the channel utilization and the time-step Δt. The channel utilization is assumed to be equal to the packet arrival rate $A_i(t)$ [29, 111], which is calculated at the beginning of the time epoch (Δt). Thus, the throughput calculated for a time period $\triangle T$ from FFM is given as:

$$B(\Delta T) = \sum_{i=1}^{n} A_i(t)\Delta t \quad where\ (\Delta T = n\Delta t) \tag{5.9}$$

The arrival rate is assumed to remain constant during the short interval Δt. This is shown as the stepped function (red line) in Figure 5.2. Therefore, the CITI link throughput for an interval is the area under the stepped function for that period. However, the actual arrival rate is not constant during the time-step and this introduces an error in the calculation. This referred to as *sample path error* and is discussed by Wu and Gong [119]. Their results show that the time-stepped simulation (TSS) model has good performance for heavily loaded systems. However, for low utilization networks, the TSS has acceptable performance for traffic that continues for large time scales. Liu et al. [30] mention that for the fluid model (FFM), which also uses the time-stepped simulation technique, the accuracy is bounded by the smallest round-trip time of TCP classes and the highest bandwidth of congested queues. Liu et al. also demonstrate that FFM could achieve sufficient accuracy with a small enough step-size. The actual arrival rate in the data network is shown as the blue line in Figure 5.2.

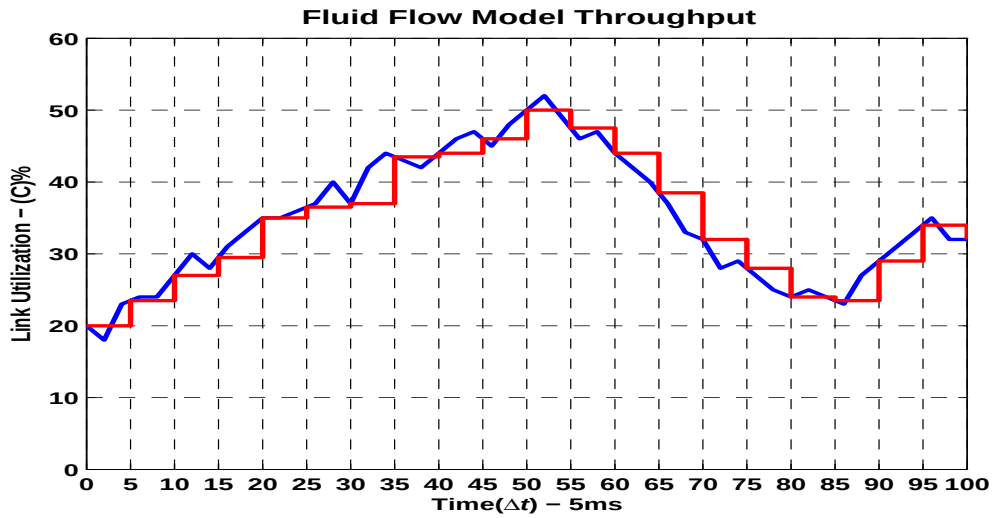


Figure 5.2: FFM Throughput as Calculated in the CITI Link

A CITI link carries CITI tokens [21] from one end to the other in the I2Sim simulator, as shown in the Figure 5.3. The channel characteristic $\alpha(t)$ is estimated from the throughput calculated from an actual data network by FFM. The other parameters of CITI links are the proportionality constant K and the channel capacity C. The channel capacity (C) represents the maximum allowable CITI token-flow possible through the channel. For the CITI link this is equal to the capacity of the physical data link. For instance, if the physical channel is 155 Mbps, then the CITI link capacity is also 155 Mbps. Since CITI tokens are defined as per-unit quantities (Chapter 3 and [21]), the channel capacity (C) is used as the base unit for per-unit representation of CITI token flow in the CITI links.

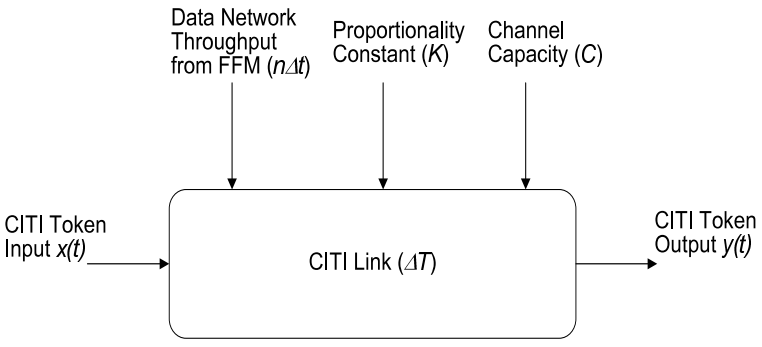


Figure 5.3: CITI Link

Determination of the proportionality constant K in Equation 5.8 is more involved. The constant K is calculated off-line (i.e., while FFM is not simulating cyber interdependency) from a stable data communication network, where traffic flow in different links are not expected to changes over time. For a stable throughput estimate B and a channel capacity C (both measured in second), K is defined as $K = C/B$. Since channel capacity varies from channel to channel, K is also channel dependent. The intention is to make the per-unit channel characteristics coefficient $\alpha(t)$ equal to unity (Equation 5.8) during stable operating conditions. This implies that there is no flow loss in a stable data network.

5.2.3 Hybrid Model Synchronization

Correct synchronization between simulator modules is important for consistent output from the hybrid systems framework. The synchronization between the I2Sim and FFM simulators is relatively simple, because the FFM simulator uses a fixed step-size algorithm (Runge-Kutta), similar to I2Sim's to calculate data network states [30]. As mentioned before, the FFM time-step is represented as $\triangle t$, while the time-step for I2Sim is represented as $\triangle T$. For synchronous solutions, we have used an approach similar to Moreira and Martí [26, 27], in which these time-steps are related as $\triangle T = n\triangle t$, where n is any integer greater than zero. Typical time-steps in our solution are $\triangle T$ = 5 minutes and $\triangle t$ = 5 ms. In our solution, no interpolated history source is needed, as we used the cumulative throughput result (Equation 5.9) that is directly available from the FFM simulator. However, the final result of the throughput calculation (Equation 5.7) was put in a history source at time $\triangle T$ in the CITI link and was propagated to the other end. The technique presented in this thesis is somewhat different form Moreira and Martí's, as we did not solve two systems jointly at time $\triangle T$ with an interpolated history source. The hybrid solution technique is rather a loosely coupled muti-rate system, where each system has their own solution model.

5.3 Hybrid Model Implementation

To integrate the FFM simulator with I2Sim, we have extended the event scheduling module of the I2Sim simulator [20], in addition to the use of CITI links-based latency techniques. The I2Sim event scheduler lets the FFM simulator run for a $\triangle T$ time-window before updating the throughput information into the CITI links of the I2Sim simulator. The I2Sim event scheduler is also the central point to decide which event will propagate to I2Sim and

which event to the FFM simulator. Figure 5.4 shows an infrastructure network topology in the hybrid simulation framework. There are two CITI links in the I2Sim part (dark arrows) that are transporting CITI tokens from one I2Sim cell to another. But, for their channel throughput estimate, they get traffic flow statistics from their corresponding nodes in the FFM simulator (shown as gray FFM nodes from which upward arrows connect the CITI link to show traffic flow statistics). The gray nodes in the fluid network (Figure 5.4) are the points, which are statically defined as FFM access point, for the corresponding CITI links in the I2Sim network. The flow characteristics at these points are received from the ODE solver. In our present implementation of the hybrid model, we assume the data traffic flow path and volume in the FFM simulator is known and can be used statically to set the CITI link characteristics in the I2Sim simulator.

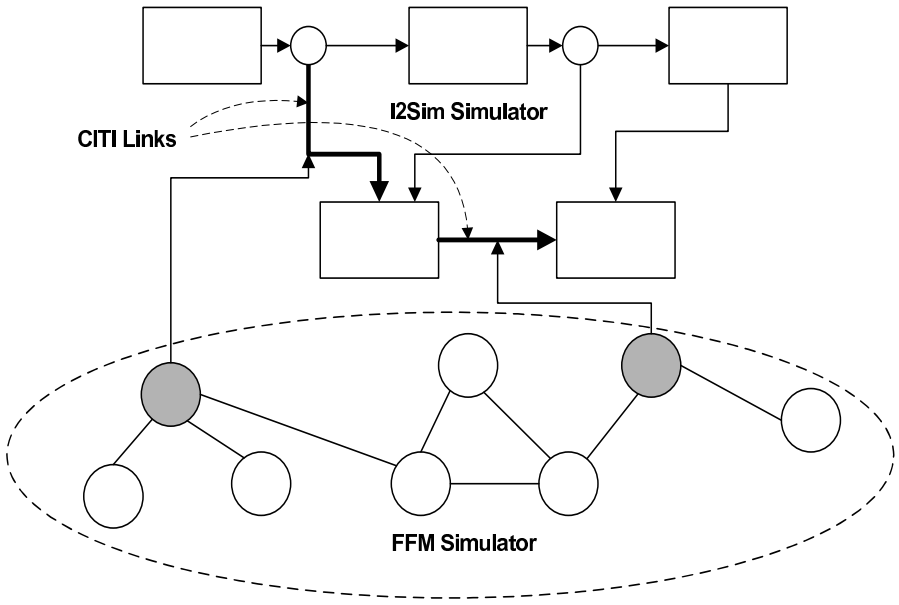


Figure 5.4: Integration of Hybrid Simulation Architecture

5.4 Case Study - Beth Israel Deaconess Medical Center

We have tested our hybrid model cyber interdependency simulation technique to simulate an actual data network failure. This breakdown occurred in November 2002 at the Beth Israel Deaconess (BIDMC), a research and teaching hospital at Harvard University, USA. This network failure is an interesting case to study because it significantly affected the

operation of BIDMC for four days. The failure case was widely reported in the news media and the consequences were well documented. One of the first reports of this incident was published in Boston Globe [3] newspaper. Following that, a narrative was written by Kilbridge [2] in the New England Journal of Medicine. A chronology of the events, with detailed accounts of the human interaction during the crash was written by Berinato [4] in CIO Magazine in the late 2003. The layout of the crashed network was available on the medical center's CIO John Halamka's blog [120]. Here, we give a brief description of the medical center's network and the failure scenario.

The BIDMC network that suffered the outage was built in 1996. It was a massive switch-based network (OSI layer 2 device [121]) that connected the three hospital campuses; namely, East Campus, West Campus and Renaissance Park, as shown in Figure 5.5. The major medical services were provided at the East Campus and West Campus locations. Multiple switched links connected these two campuses, with a mutual communication bandwidth 2Gbps. Renaissance Park was an off-campus facility, and was mainly concerned with affiliated administrative services and was connected through a SONET OC-3 link (155 Mbps). There were six other affiliated hospitals in remote locations, also connected to the BIDMC through OC-3 links.

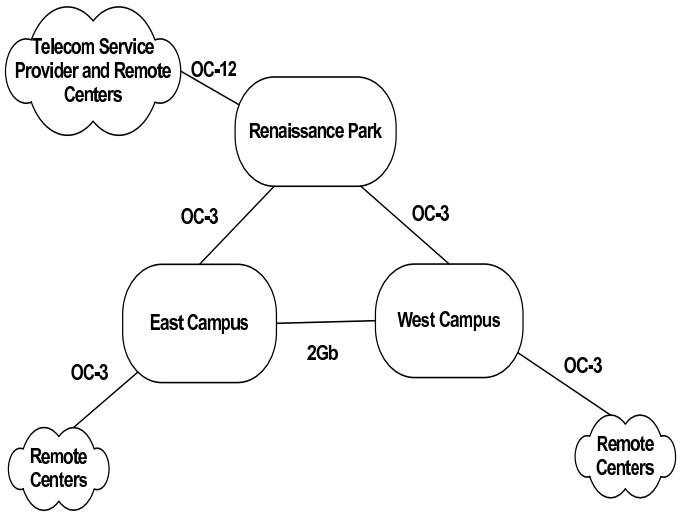


Figure 5.5: Data Network Connection Layout at BIDMC (available from [120])

Before the crash, the medical center's IT staff had implemented a wide variety of hospital automation applications for the medical center and for other affiliated hospitals. Due to the use of a switch-based architecture, a Spanning Tree Protocol (STP) was used

for the data-link layer traffic forwarding mechanism. The default values for the spanning tree protocol imposed a maximum network diameter of seven hops. This means that two distinct bridges or switches (layer 2 devices) in the network should not be more than seven hops away from each other. If they were beyond this seven hops limit, data might get stuck in a loop. However, it appears that the medical center's network operators were not aware of this limitation [3] and did not rearrange the network devices according to this constraint. Unfortunately, on 13 November 2002, one researcher inadvertently uploaded a large volume of data and flooded the network. The network was so saturated that not enough bandwidth was available to handle any user request, causing the network to slow down drastically. Doctors could not order medication or lab reports electronically, no decision support software was available, and there were many other problems.

We have compiled two sets of simulation events based on the descriptions of this failure. The first set of events is based on the event sequence described by Berinato [4]. It shows the impact of data network's failure on the medical center's output. The second set of events shows the interdependencies that emerged from other supporting infrastructures. The objective of the second simulation is to show the flexibility of I2Sim to combine different kinds of physical, human and cyber interdependencies. The details of these simulations are discussed in the following subsections.

5.4.1 I2Sim Simulation Environment

We created the simulation environment for the Beth Israel Deaconess Medical Center using the cell definitions of the hospital, electrical substation, powerhouse, water station and steam station that we used for the UBC reduced-scale simulation (Chapter 4 and [20]). This adaptation was selected due to some interesting similarities between the two hospital systems. Both hospitals have a patient capacity of about 600; and both are divided between two campuses, where one serves emergency patients and the other serves regular and long-term patients. Both hospitals focused on research and teaching. Due to this similarity, we assume that both hospitals require similar types of utility services. However, for the BIDMC simulation, we have added a data network (Section 5.4.2) as a separate layer and have also added two additional cells, the CITI service center as a control center for the utility services, and an admin center for the Renaissance Park facility. Figure 5.6 shows the internal components of each of these seven cells and their connectivity. The outputs

from the hospital are the patient discharge capacity per hour. Under normal conditions these are 27 emergency patients and 21 regular patients per hour (based on present capacity estimate [122]).

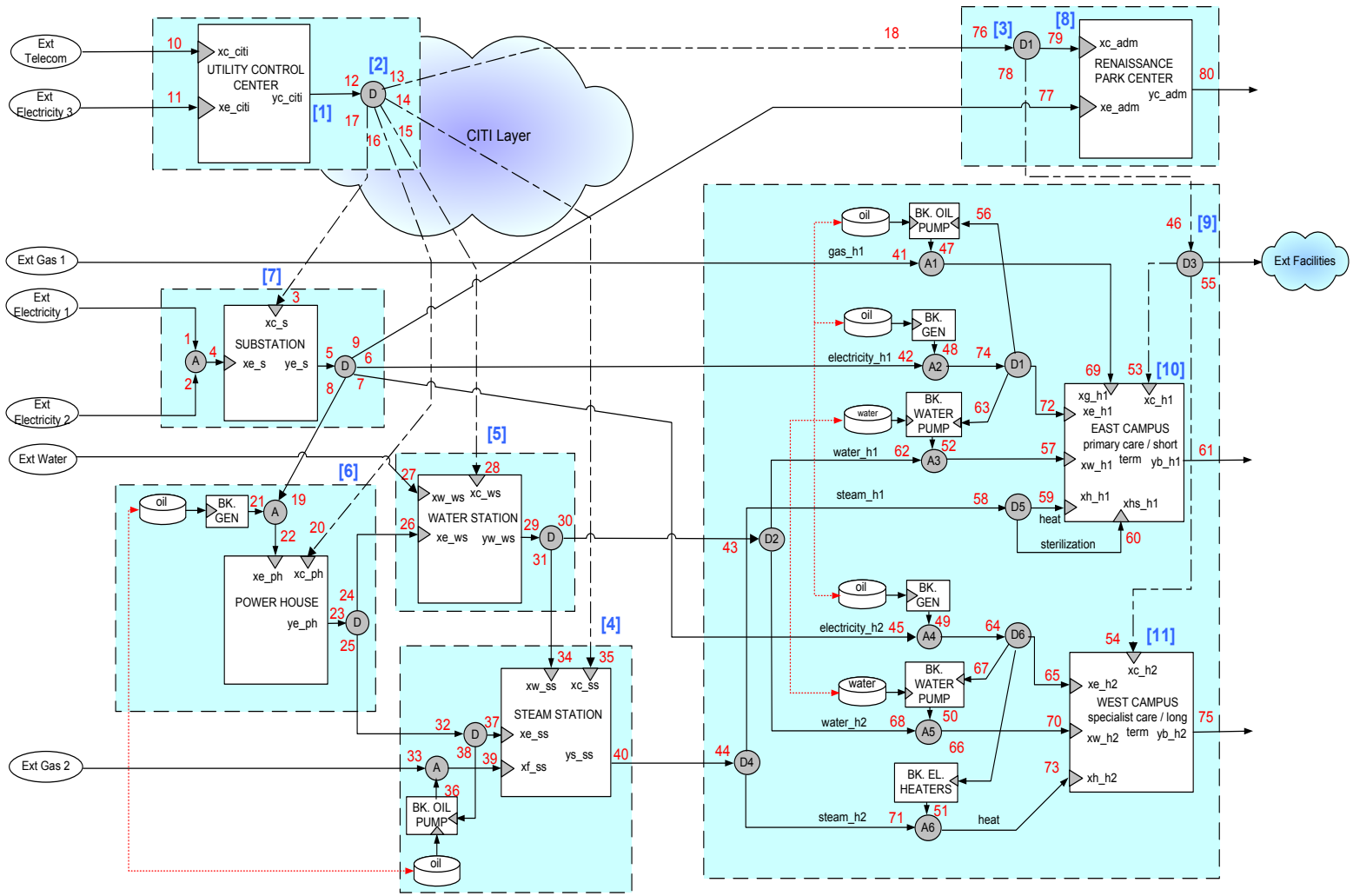


Figure 5.6: BIDMC Infrastructures in I2Sim

One important component of the BIDMC simulation is the inclusion of CITI interdependency functions from Chapter 3, which are essential components for our cyber interdependencies simulation. These functions show how the operation of a particular cell is affected by the availability of different levels of CITI services. Figure 5.7 is CITI-Hospital interdependency function from Section 3.5.4 that illustrates how the hospital's operations are affected due to the changes in CITI services. We have added CITI interdependency functions corresponding to all seven cells of BIDMC's infrastructures (from Chapter 3).

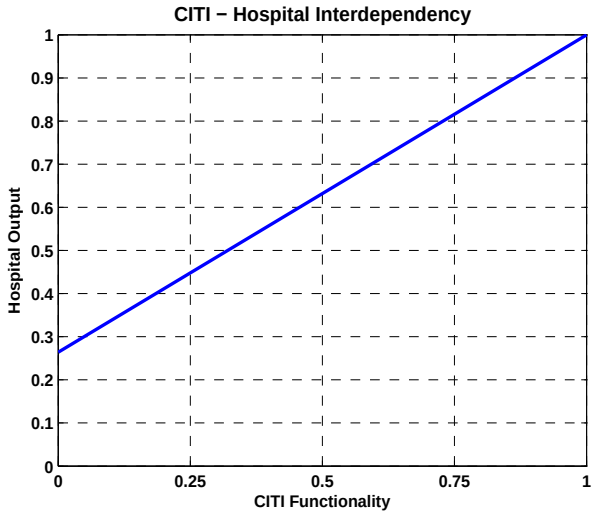


Figure 5.7: CITI and Hospital Functional Interdependency

5.4.2 FFM Simulation Environment

We have already discussed the switch-based networking topology of the BIDMC's data network shown in Figure 5.5. Due to the use of Spanning Tree Protocol (STP) in the switches, data packets propagate through the shortest paths and alternate routes are blocked in order to prevent packets from circulating in loops. Thus the BIDMC's network topology looks like a tree from the higher level protocol layers (layer 3 and above). Figure 5.8 shows the logical network connection as seen from the perspective of TCP/IP based protocols. We have used this logical topology to model BIDMC's data network layer in the FFM simulator. As shown in Figure 5.8, East Campus and West Campus are connected through 2 Gbps links. Renaissance park and all other infrastructures are connected using OC-3 (155 Mbps) links. We have also defined a number of TCP flows for our FFM simulation. These

are, 5000 flows between East and West Campuses, 3000 flows between Renaissance park and East and West Campuses, 5400 flows from Internet Service Providers to the different nodes and 400 flows for the utility control centers for monitoring and control. These numbers are assigned based on a rough estimate about the number of people that might be using different computer related services. These flows are also shown in Figure 5.8. The connections between the I2Sim and the FFM simulators are established through CITI links. In Figure 5.6, CITI links are shown in dotted lines and their corresponding FFM nodes are shown in blue numbers.

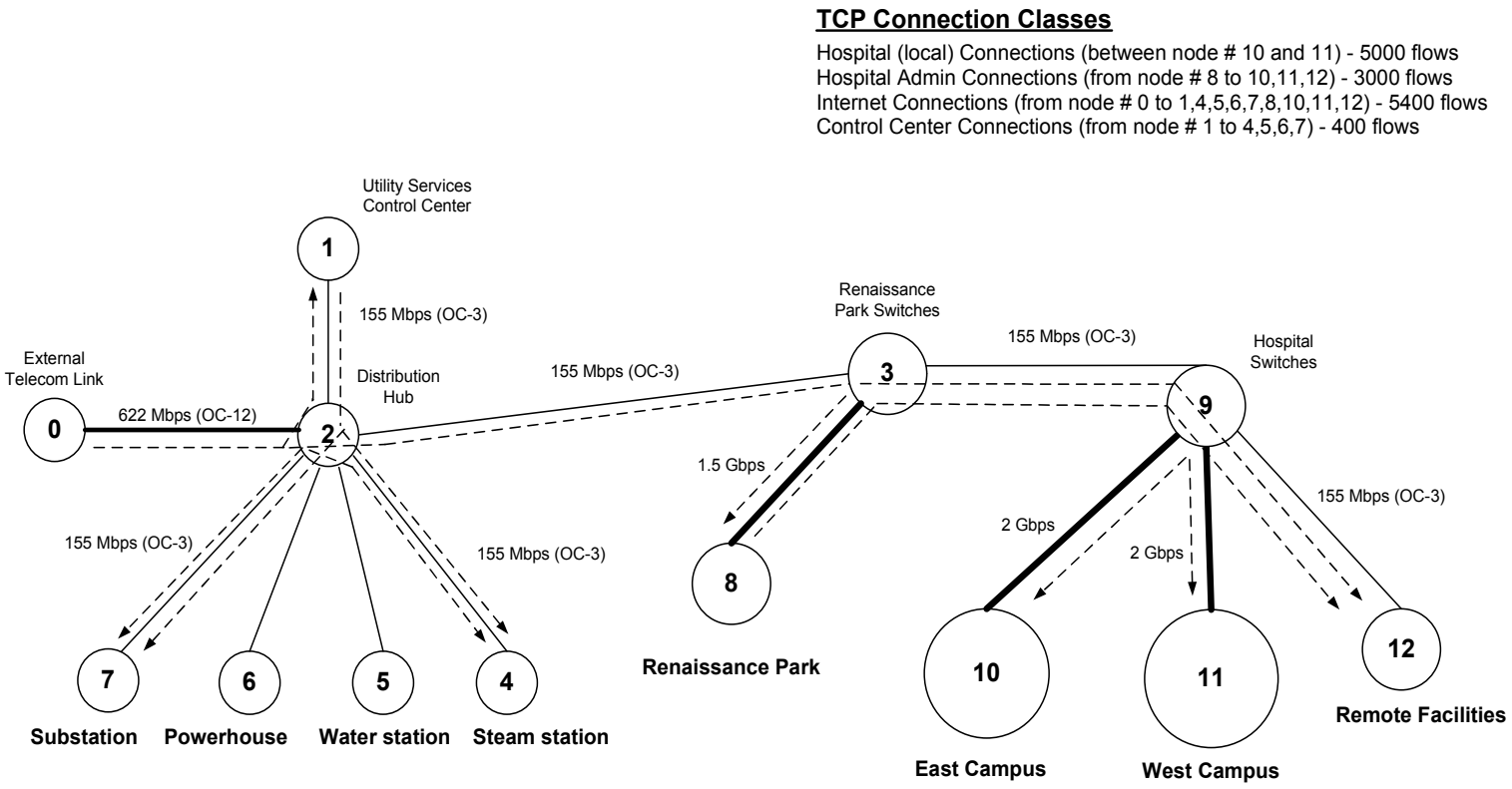


Figure 5.8: Data Network Logical Layout at BIDMC (based on Figure 5.5)

5.4.3 Scenario 1 and Results Analysis

Simulation Scenario-1 was prepared based on the description available from Berinato's article [4]. The events related to the data networks mentioned in his article are given in this section. We have simulated the first 27 hours, until Cisco System's takeover of the BIDMC's network. Our simulation of the first 27 hours suffices to show the validity of our CITI-Hospital interdependency function (Figure 5.7) and the correctness of our hybrid systems simulation model. Following Cisco's takeover, the network architecture changed significantly (from switch based to router based) and detailed information of the new architecture was not available in any documentation. The scenario starts at 1:00 PM 13 November 2002 and ends at 4:00 PM 14 November 2002. The scenario is described below.

5.4.3.1 Scenario Description

- Initial state: $t = t_0$ [1:00 PM 13 November 2002]
  - all systems run in normal state.
- $t=t_0$ + 45 (min) [1:45 PM]
  - System problem starts (network slows).
  - Decision made to shutdown local links (LANs and VLANs).
- $t = t_0$ + 60 (min) [2:00 PM]
  - Restart network switches.
  - No change in network performance observed.
- $t=t_0$ + 300 (min) [6:00 PM]
  - Daytime network load decreases.
  - Minor improvement in network performance observed.
- $t =t_0$ + 480 (min) [9:00 PM]
  - One spanning tree loop found and the loop was isolated.
  - No change in network performance.
- $t=t_0$+1080 (min) [7:00 AM 14 November 2002]
  - Network still sluggish.
- $t=t_0$ + 1140 (min) [8:00 AM]
  - Off-campus links disconnected.
  - Network performance did not improve.

- $t=t_0$ + 1620 (min) [4:00 PM]
  - Cisco CAP (Customer Assurance Program) declared.
  - Changes in network architecture starts
  - Simulation ends [27 hours - 4:00 PM 14 November 2002]

5.4.3.2 Results Analysis

Due to our selection of the FFM simulator, the description of our simulation scenarios are limited to the TCP layer. We could not simulate the actual network events related to Spanning Tree Protocol (STP), which is a data-link layer protocol. As such, our description of the network conditions were introduced by reducing the data network link bandwidth, which is the result of the STP failure visible at the TCP layer. However, the actual network event simulation is possible within a hybrid simulation model by plugging a simulation framework that is capable of simulating the OSI layer-2 protocols and events. Accordingly, we started our Scenario 1 failure conditions by reducing the data network bandwidth to half of what we would expect as 'normal' (2 Gbps) between East Campus and West Campus (between nodes 10 and 11 of Figure 5.8) at the elapsed time 45 minutes in the FFM simulator. Following this, we changed the bandwidth to values that we assumed would capture the changing nature of the failure scenario, as shown in Figure 5.9. The relationship between the CITI input relative to the hospital output is defined according to the function-description of Figure 5.7. The output capacity of the East Campus hospital is shown in Figure 5.10. Due to the CITI-Hospital linear interdependency relationship, the two graphs have identical shape. The West Campus output capacity also has a similar shape.

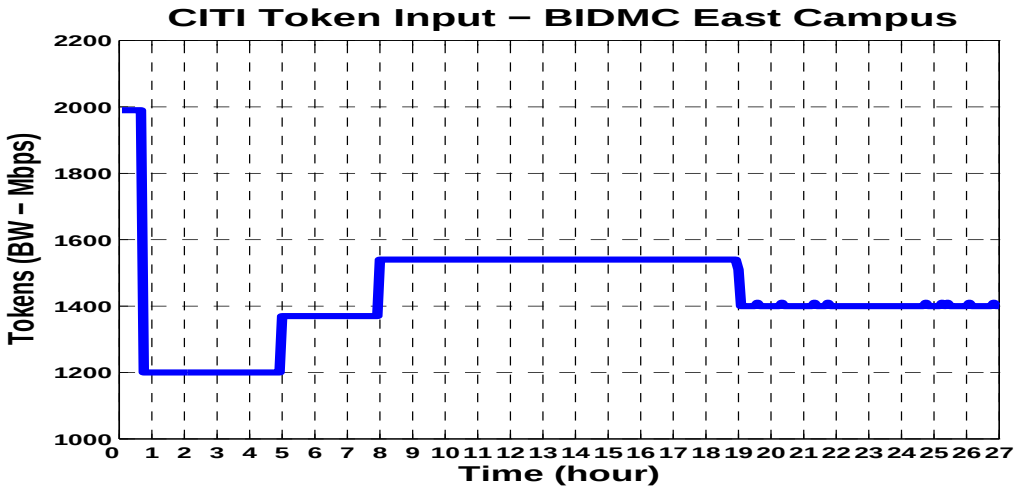


Figure 5.9: CITI Token Input to the Hospital (East Campus)

The output obtained from the simulator is consistent with the scenario described in [3, 4]. We see in our simulation that, as anticipated, the hospital's service capacity has significantly been reduced due to CITI failure. Despite one of the author's claim that the hospital was able to continue its emergency operations uninterrupted [2], other reports stated a closure of an emergency room and stated as well that lab reports that routinely take 45 minutes were delayed up to five hours [4]. It was also noted that during the outage, the hospital's management deployed hundreds of human couriers to hand-deliver patient records, lab results and other documents. From lab technicians to the CEO, everyone had worked overtime. Due to this massive enlistment of hospital staff, the hospital was able to maintain a consistent output for its emergency services [4]. Otherwise, output would likely have followed the level predicted by our model.

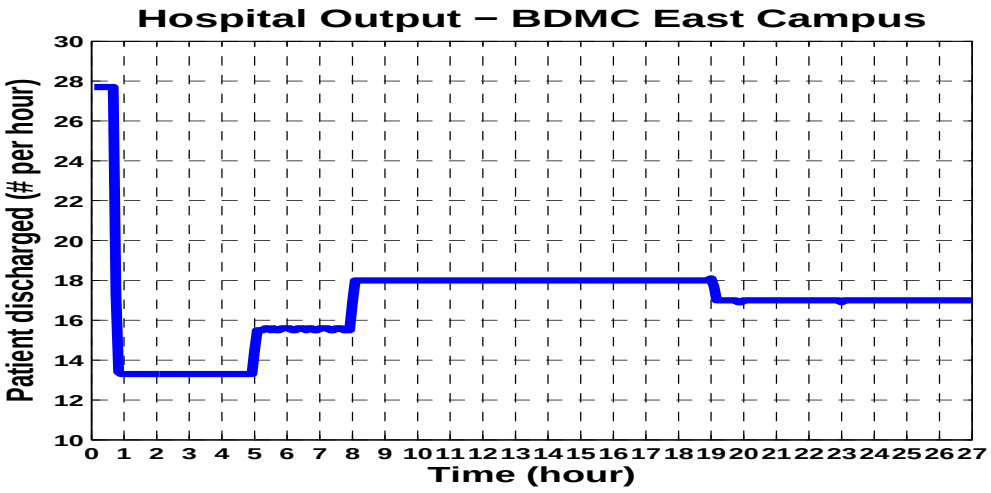


Figure 5.10: Hospital Output Capacity - Patient Discharged per hour (East Campus)

5.4.4 Scenario 2 and Results Analysis

To show the power and flexibility of I2Sim and the hybrid model, we have added one small event at the end of this Beth Israel scenario. For this we modeled into our simulation that the water station's control network at the hospital had suffered a congestion that started at the 1300 minute-mark and continued for another two and a half hours. The relationship between the water station output and the CITI input is defined through function [21] shown in Figure 5.11.

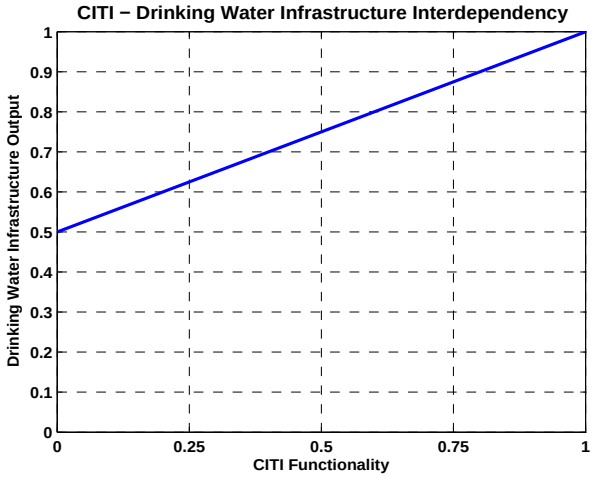


Figure 5.11: CITI-Water Infrastructure Interdependency

5.4.4.1 Scenario Description

- t=t_0 + 1300 (min) [10:00 AM 14 November 2002]
  - Water station's control network suffered congestion
- t=t_0 + 1450 (min) [12:30 PM 14 November 2002]
  - Water station's control network resumes normal operation.

5.4.4.2 Results Analysis

This event was also simulated by reducing the data network bandwidth for the water station in the FFM simulator (between node 2 and 5 of Figure 5.8). The change in the network bandwidth is shown in Figure 5.12. Due to the linear relationship between the CITI input relative to the water station's output (Figure 5.11), we have an identical reduction in water supply. The reduced water input to the hospital decreases the hospital's service capacity as shown in the Figure 5.13. The figure shows quite dramatically that the control system data network outage had more impact on the hospital's output than the outage of hospital's own data network.

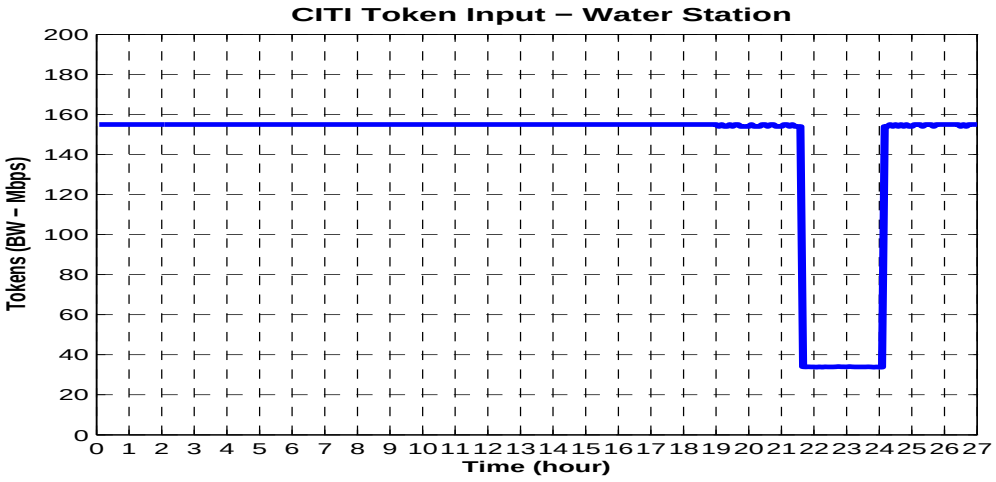


Figure 5.12: CITI Token Input to the Water Station

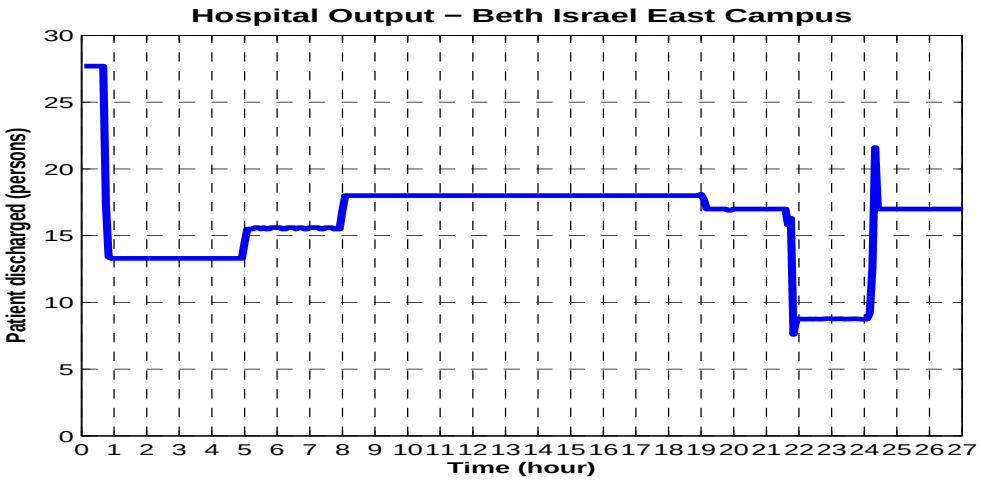


Figure 5.13: Hospital (East Campus) Output Capacity while Water Supply Decreased

5.5 Chapter Summary

In this chapter we described an approach to the simulation of cyber interdependencies, which are one of the fundamental kinds of interdependencies that exist in critical infrastructure networks. Our approach was implemented in a fully functional critical infrastructure simulator I2Sim and was validated through the simulation of real life events. The proposed hybrid model extends the capability of I2Sim such that it can capture domain-specific events. The cyber interdependencies simulation is a special case of hybrid simulation. Since cyber interdependencies have a significant impact on many critical infrastructures, our work is an important contribution in this field. In this work, we have simulated a systemic failure when data network resources become unavailable. Our CITI failure source identification (Chapter 2 and [11]) shows other types of CITI failures, that we would like to simulate in future work.

Chapter 6

Conclusion

The growing interdependency between CITI and other critical infrastructures has created new challenges for the reliable and secure operation of these multiple infrastructure networks. Many of these problems emerge from the vulnerability of the CITI infrastructure itself [11]. As critical infrastructures are the lifelines of modern societies, their disruption has major consequences that deserve serious attention from the research community. The development of appropriate modelling and simulation techniques can be very useful to understand these vulnerabilities and can be important tools for researchers and industry practitioners. Our present research is a pioneering attempt to develop such tools.

In this thesis, we have presented some techniques to model and simulate cyber interdependencies. This chapter summarizes the main contributions of this work and suggests some ideas for future extensions. We accomplished our research goals as a series of steps. First, we identified the origin of different types of CITI failures and their impacts on critical infrastructure networks. Then, we developed a novel technique to estimate interdependencies between CITI and other critical infrastructures. Finally, we developed techniques to simulate cyber interdependency in a critical infrastructure simulator. These techniques are new approaches and are important contribution to this field.

6.1 Research Contributions

In this thesis, we made following contributions to the cyber interdependency modelling and simulation for critical infrastructures:

1. **Classification of CITI faults into orthogonal types**: We have classified CITI faults into eight orthogonal fault classes and generic fault types that belong to each of these cla**sses**. This orthogonal fault classification can help to distinguish different sources of failure and can be useful for selecting an appropriate recovery procedure based on the origin of the failure. The orthogonal fault classification may have an additional advantage in resource allocation, tool development and for a variety of statistical analyses.

2. **Identification of CITI related failure trends:** We have identified different trends and statistics related to CITI failures. These findings include, failures due to malicious attack are rising, software related faults constitute the majority of the CITI failures, most of the CITI failure cases are unintentional, high impact failure cases are rising, etc. We have also identified that banking and financial services are mostly affected by CITI failure. These are some of the findings from our analysis of historical data. Understanding the trend of these failures may help governments and corporations in the allocation of budgets, and assist in the prioritization of research funds and to mobilization of work forces.

3. **Method to estimate CITI interdependency function:** We have developed a novel technique to estimate interdependency between CITI and other critical infrastructure. Following our method, we can represent interdependency in a functional form. This functional form of interdependency is an essential component of computational modelling and simulation. We have applied this method to estimate CITI interdependencies for some important critical infrastructures, such as, electrical networks, financial services, healthcare, water supply, road, railway, and air transportation, etc. The functional description of interdependency between these infrastructure networks is rooted in system theory and is required in a representation of the Cell-Channel model [10] based interdependency relationship.

4. **Development of MATE based I2Sim critical infrastructure simulator:** We have used the MATE approach to implement an efficient version of the I2Sim critical infrastructures simulator [5]. Our implementation of I2Sim [20] is based on matrix-partitioning techniques named Multi-Area Thevenin Equivalent (MATE) [25]. The MATE model has been used for large-scale real-time power system simulation and is an efficient alternative to agent-based critical infrastructure simulation techniques.

We have extended and generalized the MATE model for critical infrastructure cases and have shown its potential to solve any kind of physical infrastructure network. The critical infrastructures in I2Sim are represented according to the Cell-Channel model [10], where interdependencies among different infrastructures are represented through an extension of Loentief's input-output model. The accuracy of the I2Sim simulator has been confirmed through various benchmark tests.

5. **Hybrid systems model for cyber interdependency simulation:** We have adopted a hybrid approach to extend the capability of I2Sim simulator to capture domain specific events. The technique is based on simulating smaller time-stepped and larger time-stepped infrastructure components in parallel. The results from both these simulations are periodically combined using latency techniques. In the hybrid model, we have extended Moreira and Martí's latency techniques [26, 27] by applying it to a more generalized solution model. The cyber interdependency simulation is a special case of hybrid simulation, where the domain-specific part is represented by a data-communication simulator. We have implemented our hybrid model in I2Sim and the output was validated through a simulation of real life events.

6.2 Recommendations for Future Work

The work presented in this thesis can be extended in different directions. Some of these are:

1. **Validating trends in CITI-related failures based upon other failure-related databases:** The validity of CITI failure related statistics that we have compiled comes from the archive of the RISKS forum [38] of the Association for Computing Machinery (ACM). Even though, the RISKS forum archive has multiple contributing sources that makes it a close representation of actual scenarios, it is still a good idea to crosscheck our findings with some other failure-report repositories. However, those repositories should be bounded, similar to the RISKS forum (unlike the Internet, which is unbounded).

2. **Improving and extending interdependency estimates for different infrastructures:** The interdependency estimates (functions) we have calculated are based on

important and commonly used CITI services. As such, these empirical functions can be considered as the lower bound of the possible cyber interdependencies. These estimates can be improved through more comprehensive system studies. Besides, our interdependency estimate is limited to a few important cell types. There are different cell types for which we could not develop functions due to a lack of information. Calculating interdependencies for those infrastructures can be an important extension of this research.

3. **Extension of the hybrid model for other critical infrastructures:** The hybrid model we developed can be used to plug-in different types of domain-specific simulators in I2Sim, such as an electrical network simulator or road network simulator, etc. This can be a useful research to find different kinds of domain-specific models and their implementation techniques in the I2Sim simulator. This will also demonstrate the flexibility and accuracy of I2Sim simulator for any such extension.

4. **Information redundancy and availability during disaster:** Experience shows that having redundant physical infrastructures (e.g., backup channel, power supply, etc.) is always helpful [49]. A similar approach has to be explored for CITI services. For instance, it is important to study an improved information-caching mechanism in a hospital, so that in the case of a disaster, important information can still be available to continue hospital operations.

6.3 Final Remarks

Several new concepts related to cyber interdependency simulation have been introduced and implemented in this research. The techniques presented in this thesis can give us the ability to model and simulate cyber interdependencies and can give reasonable answers to the research questions that we raised in the introduction of this thesis. Our approaches were validated against real-life test cases and will be useful in future to predict impacts for different CITI failure scenarios. As there are many new ideas appearing in this growing discipline, it will be interesting to see how this field evolves in coming days.

Bibliography

[1] S. M. Rinaldi, J. P. Peerenboom, and T. K. Kelly, "Identifying, Understanding, and Analyzing Critical Infrastructure Interdependencies," *IEEE Control Systems Magazine*, vol. 21, no. 6, pp. 11–25, December 2001.

[2] P. Kilbridge, "Computer Crash – Lessons from a System Failure," *New England Journal of Medicine*, vol. 348, no. 10, pp. 881–882, 2003. [Online]. Available: http://content.nejm.org/cgi/reprint/348/10/881.pdf

[3] A. Barnard, "Beth Israel Deaconess Copes with a Massive Computer Crash," Boston Globe, November 2002. [Online]. Available: http://ieee802.org/secmail/msg02970.html

[4] S. Berinato, "All Systems Down," *CIO Magazine. Feb*, vol. 15, pp. 46–53, 2003. [Online]. Available: http://www.cio.com.au/article/65115/all_systems_down

[5] J. Marti, C. Ventura, J. Hollman, K. Srivastava, and H. Juárez, "I2Sim Modelling and Simulation Framework for Scenario Development, Training, and Real-Time Decision Support of Multiple Interdependent Critical Infrastructures during Large Emergencies," in *How to Modelling and Simulation Meeting the Defence Challenges out to 2015?* Vancouver: NATO (OTN) MSG-060, October 7-8 2008.

[6] Public Safety Canada, "Critical Infrastructure Sectors," 2008. [Online]. Available: http://www.ps-sp.gc.ca/prg/em/nciap/about-eng.aspx

[7] "Joint Infrastructure Interdependencies Research Program (JIIRP)," 2004, Government of Canada. [Online]. Available: http://www.nserc.ca/programs/jiirp_e.htm

[8] "Infrastructure Interdependencies Simulation Team," 2005, University of British Columbia. [Online]. Available: http://www.ece.ubc.ca/~jiirp/

[9] J. Marti, J. Hollman, C. Ventura, and J. Jatskevich, "Design for Survival. Real-Time Infrastructures Coordination," in *Proceedings of the International Workshop on Complex Network and Infrastructure Protection*. CNIP2006 Rome, March 2006.

[10] ——, "Dynamic Recovery of Critical Ifrastructures: Real-time Temporal Coordination," *International Journal of Critical Infrastructures*, vol. 4, no. 1, pp. 17–31, 2008.

[11] H. A. Rahman, K. Beznosov, and J. R. Marti, "Identification of Sources of Failures and their Propagation in Critical Infrastructures from 12 Years of Public Failure Reports," *International Journal of Critical Infrastructures*, vol. 5, no. 3, pp. 220–244, 2009. [Online]. Available: http://www.ece.ubc.ca/~rahmanha/cris2006_CS2_paper.pdf

[12] ——, "Identification of Sources of Failures and their Propagation in Critical Infrastructures from 12 Years of Public Failure Reports," in *Third International Conference on Critical Infrastructures, Alexandria, VA*. CRIS 2006, September 2006.

[13] P. Pederson, D. Dudenhoeffer, S. Hartley, and M. Permann, "Critical Infrastructure Interdependency Modeling: A Survey of US and International Research," Idaho National Laboratory (INL), Tech. Rep., 1 August 2006.

[14] T. Brown, W. Beyeler, and D. Barton, "Assessing Infrastructure Interdependencies: the Challenge of Risk Analysis for Complex Adaptive Systems," *International Journal of Critical Infrastructures*, vol. 1, no. 1, pp. 108–117, 2004.

[15] G. Goth, "Infrastructure Simulation Effort Has High Hopes, Faces High Hurdles," *Computing in Science and Engineering*, vol. 4, no. 3, pp. 4–7, September 2002.

[16] S. Conrad, R. LeClaire, G. O'Reilly, and H. Uzunalioglu, "Critical National Infrastructure Reliability Modeling and Analysis," *Bell Labs Technical Journal*, vol. 11, no. 3, pp. 57–71, 2006.

[17] G. OŔeilly, H. Uzunalioglu, D. Houck, and T. Morawski, "A Dynamic Simulation Approach to Business Continuity of Wireline and Wireless Networks With Cross-Industry Infrastructures," in *International Conference on Communications*. IEEE, 16-20 May 2005.

[18] A. Jrad, H. Uzunalioglu, D. Houck, G. OŔeilly, S. Conrad, and W. Beyeler, "Wireless and Wireline Network Interactions in Disaster Scenarios," in *Military Communications Conference*, ser. Military Communications Conference. IEEE, 17-20 Oct 2005.

[19] G. OŔeilly, A. Jrad, R. Nagarajan, T. Brown, and S. Conrad, "Critical Infrastructure Analysis of Telecom for Natural Disasters," in *12th International Telecommunications Network Strategy and Planning Symposium*. IEEE, November 2006.

[20] H. A. Rahman, M. Armstrong, D. Mao, and J. R. Martí, "I2Sim: A Matrix-partition based Framework for Critical Infrastructure Interdependencies Simulation," in *8th*

IEEE Electrical Power & Energy Conference, Vancouver, Canada. EPEC 2008, 6-7 October 2008. [Online]. Available: http://www.ece.ubc.ca/~rahmanha/I2Sim_paper_EPEC2008.pdf

[21] H. A. Rahman and J. R. Martí, "Quantitative Estimates of Critical Infrastructure Interdependencies on Communication and Information Technology Infrastructure," University of British Columbia, Tech. Rep. I2SIM-TR-2008-02, 22 December 2008.

[22] S. Panzieri, R. Setola, and G. Ulivi, "An Agent Based Simulator for Critical Interdependent Infrastructures," in *2nd International Conference on Critical Infrastructures*, ser. Securing Critical Infrastructures, October 25-27 2004.

[23] S. De Porcellinis, R. Setola, S. Panzieri, and G. Ulivi, "Simulation of Heterogeneous and Interdependent Critical Infrastructures," *International Journal of Critical Infrastructures*, vol. 4, no. 1, pp. 110–128, 2008.

[24] K. Hopkinson, X. Wang, R. Giovanini, J. Thorp, K. Birman, and D. Coury, "EPOCHS: A Platform for Agent-based Electric Power and Communication Simulation Built from Commercial Off-the-shelf Components," *IEEE Transactions on Power Systems*, vol. 21, no. 2, pp. 548–558, 2006.

[25] J. Marti, L. Linares, J. Hollman, and F. Moreira, "OVNI: Integrated Software/Hardware Solution for Real-Time Simulation of Large Power Systems," in *Proceedings of the 14th Power Systems Computation Conference, Seville, Spain*. CNIP2006 Rome, June 24th - 28th 2002.

[26] F. A. Moreira, "Latency Techniques in Power System Transients Simulation," PhD Thesis, University of British Columbia, October 2002.

[27] F. Moreira and J. Marti, "Latency Techniques for Time-domain Power System Transients Simulation," *Power Systems, IEEE Transactions on*, vol. 20, no. 1, pp. 246–253, 2005.

[28] E. F. Bedell, *The Computer Solution: Strategies for Success in the Information Age*. New York, NY, USA: Dow Jones-Irwin, Homewood, 1984.

[29] V. Misra, W.-B. Gong, and D. Towsley, "Fluid-based Analysis of a Network of AQM Routers Supporting TCP Flows with an Application to RED," *SIGCOMM Comput. Commun. Rev.*, vol. 30, no. 4, pp. 151–160, 2000.

[30] Y. Liu, F. Presti, V. Misra, D. Towsley, and Y. Gu, "Scalable Fluid Models and Simulations for Large-Scale IP Networks," *ACM Transactions on Modeling and Computer Simulation*, vol. 14, no. 3, pp. 305–324, 2004.

[31] M. Armstrong, J. R. Marti, L. R. Linares, and P. Kundur, "Multilevel MATE for Efficient Simultaneous Solution of Control Systems and Nonlinearities in the OVNI

Simulator," *IEEE Transactions on Power Systems*, vol. 21, no. 3, pp. 1100–1115, August 2006.

[32] J. Moteff and P. Parfomak, *Critical Infrastructure and Key Assets Definition and Identification*. Congressional Research Service, Library of Congress, 2004.

[33] G. Bush, "Executive Order on Critical Infrastructure Protection," in *Proceedings of the 12th annual conference on Computers, Freedom and Privacy*. ACM Press New York, NY, USA, 2002, pp. 1–10.

[34] S. Gorman, L. Schintler, R. Kulkarni, and R. Stough, "The Revenge of Distance: Vulnerability Analysis of Critical Information Infrastructure," *Journal of Contingencies and Crisis Management*, vol. 12, no. 2, pp. 48–63, 2004.

[35] D. R. Kuhn, "Sources of Failure in The Public Switched Telephone Network," *IEEE Computer*, vol. 4, no. 30, pp. 31–36, April 1997.

[36] "FCC Network Outage Reporting System - User Manual," 2005, http://www.fcc.gov/oet/outage/nors_manual.pdf. [Online]. Available: http://www.fcc.gov/oet/outage/nors_manual.pdf

[37] E. H. Spafford, "Congressional Testimony," 10 October 2001, http://www.house.gov/science/full/oct10/spafford.htm. [Online]. Available: http://www.house.gov/science/full/oct10/spafford.htm

[38] "The RISKS Forum," 1985, Association for Computing Machinery. [Online]. Available: http://catless.ncl.ac.uk/Risks

[39] B. Fischhoff, P. Slovic, and S. Lichtenstein, "Lay Foibles and Expert Fables in Judgments about Risk," *American Statistician*, vol. 36, no. 3, pp. 240–255, August 1982.

[40] G. Rowe and G. Wright, "Differences in Expert and Lay Judgments of Risk: Myth or Reality?" *Risk Analysis*, vol. 21, no. 2, pp. 341–356, April 2001.

[41] S. Chang, T. McDaniels, J. Mikawoz, and K. Peterson, "Infrastructure Failure Interdependencies in Extreme Events: Power Outage Consequences in the 1998 Ice Storm," *Natural Hazards*, vol. 41, no. 2, pp. 337–358, 2007.

[42] J. D. Howard, "An Analysis of Security Incidents on the Internet 1989-1995," Ph.D. Thesis, Carnegie Mellon University, 1997.

[43] A. Chakrabarti and G. Manimaran, "Internet Infrastructure Security: A Taxonomy," *IEEE Network*, vol. 16, no. 6, pp. 13–21, November/December 2002.

[44] P. G. Neumann, *Computer-Related Risks*. Addison-Wesley Professional, 18 October 1994.

[45] A. Avizienis, J. C. Laprie, B. Randell, and C. Landwehr, "Basic Concepts and Taxonomy of Dependable and Secure Computing," *IEEE Transactions on Dependable and Secure Computing*, vol. 1, no. 1, pp. 11–33, January 2004.

[46] J. D. Howard and T. A. Longstaff, "A Common Language for Computer Security Incidents," Sandia National Laboratories, Tech. Rep. SAND98-8997, 1998.

[47] IEEE, *IEEE Standard Glossary of Software Engineering Terminology. IEEE Standard 610.12-1990*, Institute of Electrical and Electronics Engineers, Inc., 1990.

[48] R. Chillarege, I. Bhandari, J. Chaar, M. Halliday, D. Moebus, B. Ray, M. Wong, I. Center, and Y. Heights, "Orthogonal Defect Classification - A Concept for In-process Measurements," *IEEE Transactions on Software Engineering*, vol. 18, no. 11, pp. 943–956, 1992.

[49] E. E. Balkovich and R. H. Anderson, "Critical Infrastructures will Remain Vulnerable: Neighbourhoods Must Fend for Themselves," *International Journal of Critical Infrastructures*, vol. 1, no. 1, pp. 8–19, 2004.

[50] B. Kirwan, "The Role of the Controller in the Accelerating Industry of Air Traffic Management," *Safety Science*, vol. 37, no. 2-3, pp. 151–185, 2001.

[51] R. Clemen, G. Fischer, and R. Winkler, "Assessing Dependence: Some Experimental Results," *Management Science*, vol. 46, no. 8, pp. 1100–1115, 2000.

[52] A. Sage, *Methodology for Large-Scale Systems*. McGraw-Hill New York, 1977.

[53] L. Liu, "Prototyping and Cells Modelling of the Infrastructures Interdependencies Simulator I2Sim," Master Thesis, University of British Columbia, August 2007.

[54] T. Renkema and E. Berghout, "Methodologies for Information Systems Investment Evaluation at the Proposal Stage: a Comparative Review," *Information and Software Technology*, vol. 39, no. 1, pp. 1–13, 1997.

[55] P. Sassone, "Cost Benefit Analysis of Information Systems: A Survey of Methodologies," in *Proceedings of the ACM SIGOIS and IEEECS TC-OA 1988 Conference on Office information systems*. ACM New York, NY, USA, 1988, pp. 126–133.

[56] T. Chou, S. Chou, and G. Tzeng, "Evaluating IT/IS Investments: A Fuzzy Multicriteria Decision Model Approach," *European Journal of Operational Research*, vol. 173, no. 3, pp. 1026–1046, 2006.

[57] S. Smithson and R. Hirschheim, "Analysing Information Systems Evaluation: Another Look at an Old Problem," *European Journal of Information Systems*, vol. 7, pp. 158–174, 1998.

[58] L. Hitt and E. Brynjolfsson, "Productivity, Business Profitability, and Consumer Surplus: Three Different Measures of Information Technology Value," *MIS Quarterly*, vol. 20, no. 2, pp. 121–142, 1996.

[59] A. de Moor, "A Practical Method for Courseware Evaluation," in *Proceedings of the 2nd International Conference on Pragmatic Web*. ACM New York, NY, USA, 2007, pp. 57–63.

[60] ——, "The Pragmatic Evaluation of Tool System Interoperability," in *Proceedings of the Second Conceptual Structures Tool Interoperability Workshop (CSTIW 2007)*, 2007, pp. 1–19.

[61] B. Brinton Anderson, A. Bajaj, and W. Gorr, "An Estimation of the Decision Models of Senior IS Managers when Evaluating the External Quality of Organizational Software," *The Journal of Systems & Software*, vol. 61, no. 1, pp. 59–75, 2002.

[62] D. Karlsson, M. Hemmingsson, and S. Lindahl, "Wide-area System Monitoring and Control-terminology, Phenomena, and Solution Implementation Strategies," *Power and Energy Magazine, IEEE*, vol. 2, no. 5, pp. 68–76, 2004.

[63] F. Wu, K. Moslehi, and A. Bose, "Power System Control Centers: Past, Present, and Future," *Proceedings of the IEEE*, vol. 93, no. 11, pp. 1890–1908, 2005.

[64] P. Pourbeik, P. Kundur, and C. Taylor, "The Anatomy of a Power Grid Blackout - Root Causes and Dynamics of Recent Major Blackouts," *Power and Energy Magazine, IEEE*, vol. 4, no. 5, pp. 22–29, 2006.

[65] S. Johnsen, R. Ask, and R. Roisli, "Reducing Risk in Oil and Gas Production Operations," *International Federation For Information Processing-Publications-IFIP*, vol. 253, p. 83, 2008.

[66] S. Bajpai and J. Gupta, "Securing Oil and Gas Infrastructure," *Journal of Petroleum Science and Engineering*, vol. 55, no. 1-2, pp. 174–186, 2007.

[67] S. BV and M. Publishers, *Electronic Banking: The Ultimate Guide to Business and Technology of Online Banking*. Springer, 2001.

[68] R. Haux, "Health Information Systems–Past, Present, Future," *International Journal of Medical Informatics*, vol. 75, no. 3-4, pp. 268–281, 2006.

[69] P. Reichertz, "Hospital Information Systems – Past, Present, Future," *International Journal of Medical Informatics*, vol. 75, no. 3-4, pp. 282–299, 2006.

[70] R. Koppel, J. Metlay, A. Cohen, B. Abaluck, A. Localio, S. Kimmel, and B. Strom, "Role of Computerized Physician Order Entry Systems in Facilitating Medication Errors," *The Journal of the American Medical Association*, vol. 293, no. 10, pp. 1197–1203, 2005.

[71] E. Ammenwerth, A. Buchauer, and R. Haux, "A Requirements Index for Information Processing in Hospitals," *Methods of Information in Medicine*, vol. 41, no. 4, pp. 282–288, 2002.

[72] S. M. Maviglia, G. J. Kuperman, and B. Middleton, *Hospital Information Systems*, 1, Ed. Lippincott Williams & Wilkins, 2005.

[73] I. Hunt, B. Wall, and H. Jadgev, "Applying the Concepts of Extended Products and Extended Enterprises to Support the Activities of Dynamic Supply Networks in the Agri-food Industry," *Journal of Food Engineering*, vol. 70, no. 3, pp. 393–402, 2005.

[74] J. Van der Vorst, A. Beulens, and P. van Beek, "Innovations in Logistics and ICT in Food Supply Chain Networks," *Innovation in Agri-Food Systems*, vol. 1, pp. 245–292, 2005.

[75] E. Hargesheimer, O. Conio, J. Popovicova, and C. PROAQUA, *Online Monitoring for Drinking Water Utilities*. American Water Works Association, 2002.

[76] M. Schütze, A. Campisano, H. Colas, P. Vanrolleghem, and W. Schilling, "Real-Time Control of Urban Water Systems," in *Proceedings of the International Conference on Pumps, Electromechanical Devices and Systems Applied to Urban Water Management*, 2003, pp. 1–19.

[77] A. Ostfeld and E. Salomons, "Optimal Layout of Early Warning Detection Stations for Water Distribution Systems Security," *Journal of Water Resources Planning and Management*, vol. 130, p. 377, 2004.

[78] G. Giannopoulos, "The Application of Information and Communication Technologies in Transport," *European Journal of Operational Research*, vol. 152, no. 2, pp. 302–320, 2004.

[79] G. Giannopoulos and M. McDonald, "Developments in Transport Telematics Applications in Japan: Traffic Management, Freight and Public Transport," *Transport Reviews*, vol. 17, no. 1, pp. 37–59, 1997.

[80] Z. Lin, "Network OAM Requirements for the New York City Transit Network," *Communications Magazine, IEEE*, vol. 42, no. 10, pp. 112–116, 2004.

[81] M. Clarke, "Irregular Airline Operations: A Review of the State-of-the-practice in Airline Operations Control Centers," *Journal of Air Transport Management*, vol. 4, no. 2, pp. 67–76, 1998.

[82] B. Fields, P. Amaldi, and A. Tassi, "Representing Collaborative Work: the Airport as Common Information Space," *Cognition, Technology & Work*, vol. 7, no. 2, pp. 119–133, 2005.

[83] H. Jasso, T. Fountain, C. Baru, W. Hodgkiss, D. Reich, and K. Warner, "Prediction of 9-1-1 Call Volumes for Emergency Event Detection," in *Proceedings of the 8th annual international conference on Digital government research: bridging disciplines & domains*. Digital Government Research Center, 2007, pp. 148–154.

[84] N. Netten and M. van Someren, "Automated Support for Dynamic Information Distribution in Incident Management," in *Proceedings of the 3 International ISCRAM Conference (B. Van de Walle and M. Turoff, eds.)*, 2006.

[85] E. Gomez and M. Turoff, "Community Crisis Response Teams: Leveraging Local Resources through ICT E-Readiness," in *Hawaii International Conference on System Sciences*, vol. 40, no. 1. IEEE, 2007, p. 398.

[86] C. Williams, M. Markus, M. Tyworth, S. Sawyer, M. Dias, S. Vilvovsky, J. Fedorowicz, and D. Jacobson, "Mapping Theory to Practice: a Cartographic Analysis of Public Safety Networks," in *Proceedings of the 2008 international conference on Digital government research*. Digital Government Research Center, 2008, pp. 171–180.

[87] I. Shin, "Adoption of Enterprise Application Software and Firm Performance," *Small Business Economics*, vol. 26, no. 3, pp. 241–256, 2006.

[88] T. OŔourke, A. Lembo, and L. Nozick, "Lessons Learned from the World Trade Center Disaster about Critical Utility Systems," *Beyond September 11 th: An Account of Post-Disaster Research*, vol. 1, pp. 269–290, 2003.

[89] M. Amin, *National Infrastructures as Complex Interactive Networks. Chapter 14 in Automation, Control, and Complexity: New Developments and Directions*, T. Samad and J. Weyrauch, Eds. John Wiley and Sons, 2000.

[90] J. Burgess, "Social Values and Material Threat: the European Programme for Critical Infrastructure Protection," *International Journal of Critical Infrastructures*, vol. 3, no. 3, pp. 471–487, 2007.

[91] A. Ghorbani and E. Bagheri, "The State of the Art in Critical Infrastructure Protection: a Framework for Convergence," *International Journal of Critical Infrastructures*, vol. 4, no. 3, pp. 215–244, 2008.

[92] L. Lee, H. Nwana, D. Ndumu, and P. D. Wilde, "The Stability, Scalability and Performance of Multi-Agent Systems," *BT Technology Journal*, vol. 16(3), pp. 94–103, 1998.

[93] M. Dunn, "Understanding Critical Information Infrastructures: An Elusive Quest," *International CIIP Handbook 2006*, vol. 1, pp. 27–40, 2006.

[94] "Simulink - MATLAB," 2007, simulation and Model-Based Design. [Online]. Available: http://www.mathworks.com/products/simulink/

[95] J. Hollman and J. Marti, "Real-time Network Simulation with PC-Cluster," *Power Systems, IEEE Transactions on*, vol. 18, no. 2, pp. 563–569, 2003.

[96] M. Tomim, "Parallel Computation of Large Power System Networks Using the Multi-Area Thévenin Equivalents," PhD Thesis, University of British Columbia, August 2009.

[97] H. Min, W. Beyeler, T. Brown, Y. Son, and A. Jones, "Toward Modeling and Simulation of Critical National Infrastructure Interdependencies," *IIE Transactions*, vol. 39, no. 1, pp. 57–71, 2007.

[98] D. D. Dudenhoeffer, M. R. Permann, and M. Manic, "CIMS: a Framework for Infrastructure Interdependency Modeling and Analysis," *Proceedings of the 37th conference on Winter simulation*, vol. 1, pp. 478–485, 2006.

[99] K. Hopkinson, K. Birman, R. Giovanini, D. Coury, X. Wang, and J. Thorp, "EPOCHS: Integrated Commercial Off-the-shelf Software for Agent-based Electric Power and Communication Simulation," in *Proceedings of the 2003 Winter Simulation Conference, 2003*, vol. 2, 2003.

[100] A. Sulistio, C. Yeo, and R. Buyya, "A Taxonomy of Computer-based Simulations and its Mapping to Parallel and Distributed Systems Simulation Tools," *Software-Practice and Experience*, vol. 34, no. 7, pp. 653–673, 2004.

[101] K. Thibert, "A Methodology for Assessing the Seismic Risk of Buildings," Master Thesis, University of British Columbia, April 2008.

[102] M. DeTao, "Modeling Hospital: Construction and Fitting the Multi-Dimension I/O Function with Incomplete Information," University of British Columbia, Tech. Rep., September 2007.

[103] L. Linares-Rojas, "OVNI: (Object Virtual Network Integrator) a New Fast Algorithm for the Simulation of Very Large Electric Networks in Real Time," PhD Thesis, University of British Columbia, August 2000.

[104] H. A. Rahman, "I2Sim User Guide," University of British Columbia, Tech. Rep., 12 October 2007. [Online]. Available: http://www.ece.ubc.ca/~jiirp/publications.html

[105] "Boost C++ Libraries," 1999, free peer-reviewed portable C++ source libraries. [Online]. Available: http://www.boost.org/

[106] L. Garshol, "BNF and EBNF: What are They and How do They Work?" 2004. [Online]. Available: http://www.garshol.priv.no/download/text/bnf.html

[107] H. Schwetman, "Hybrid Simulation Models of Computer Systems," *Communications of the ACM*, vol. 21, no. 9, pp. 718–723, 1978.

[108] S. Bohacek, J. Hespanha, J. Lee, and K. Obraczka, "A Hybrid Systems Modeling Framework for Fast and Accurate Simulation of Data Communication Networks," in *Proceedings of the 2003 ACM SIGMETRICS international conference on Measurement and modeling of computer systems*. ACM New York, NY, USA, 2003, pp. 58–69.

[109] Y. Guo, W. Gong, D. Towsley, and M. Amherst, "Time-stepped Hybrid Simulation (TSHS) for Large Scale Networks," in *IEEE INFOCOM 2000. Nineteenth Annual Joint Conference of the IEEE Computer and Communications Societies. Proceedings*, vol. 2, 2000.

[110] A. Kavimandan, W. Lee, M. Thottan, A. Gokhale, and R. Viswanathan, "Network Simulation via Hybrid System Modeling: A Time-stepped Approach," in *Computer Communications and Networks, 2005. ICCCN 2005. Proceedings. 14th International Conference on*, 2005, pp. 531–536.

[111] Y. Gu, Y. Liu, and D. Towsley, "On Integrating Fluid Models with Packet Simulation," in *IEEE INFOCOM*, vol. 4, 2004, pp. 2856–2865. [Online]. Available: http://www-net.cs.umass.edu/fluid/ffm.html

[112] H. Dommel, "Digital Computer Solution of Electromagnetic Transients in Single and Multiphase Networks," *IEEE Transactions on Power Apparatus and Systems*, vol. PAS-88, pp. 388–399, 1969.

[113] A. Adas, "Traffic Models in Broadband Networks," *IEEE Communications Magazine*, vol. 35, no. 7, pp. 82–89, 1997.

[114] I. Lee and A. Fapojuwo, "Stochastic Processes for Computer Network Traffic Modeling," *Computer Communications*, vol. 29, no. 1, pp. 1–23, 2005.

[115] L. Breslau, D. Fall, K. Floyd, S. Heidemann, J. Helmy, A. Huang, P. McCanne, S. Varadhan, and K. Yu, "Advances in Network Simulation," *Computer*, vol. 33, no. 5, pp. 59–67, 2000.

[116] S. Low, F. Paganini, and J. Doyle, "Internet Congestion Control," *Control Systems Magazine, IEEE*, vol. 22, no. 1, pp. 28–43, 2002.

[117] D. Lee and N. Brownlee, "Passive Measurement of One-way and Two-way Flow Lifetimes," *SIGCOMM Comput. Commun. Rev.*, vol. 37, no. 3, pp. 17–28, 2007.

[118] A. Botta, A. Pescapé, and G. Ventre, "Quality of Service Statistics Over Heterogeneous Networks: Analysis and Applications," *European Journal of Operational Research*, vol. 191, no. 3, pp. 1075–1088, 2008.

[119] Y. Wu and W. Gong, "Accuracy Study of Time-stepped Simulation of High Speed Networks," in *IEEE International Conference on Communications, 2003. ICC'03*, vol. 3, 2003.

[120] J. D. Halamka, "The CareGroup Network Outage," March 2008, Dr. Halamka's Blog. [Online]. Available: http://geekdoctor.blogspot.com/2008/03/caregroup-network-outage.html

[121] H. Zimmermann, "OSI Reference Model – The ISO Model of Architecture for Open Systems Interconnection," *Communications, IEEE Transactions on [legacy, pre-1988]*, vol. 28, no. 4, pp. 425–432, 1980.

[122] "Beth Israel Deaconess Medical Center's Stats and Facts," Website, May 2009. [Online]. Available: http://www.bidmc.org/AboutBIDMC/StatsandFacts.aspx

Appendices

Appendix A

A Sample CITI Interdependency Assessment Questionnaire:

Infrastructure Cell Name: .....
Person Interviewed: .....
Date of Interview: .....

Instructions: Please answer following questions based on the available data (e.g., design documentation, service request, failure log, etc.) or based on your own knowledge and experience. Quantitative figures are preferred, if that applies. However, if such information is not available, descriptive answers will be useful as well. Please use additional papers, if the given space is not sufficient.

1. Please name CITI systems and applications (hardware and software) according to their importance to manage your facility. This may include monitoring, controlling, coordination and decision-making, data archiving, etc. Please specify the importance in consideration to the output of your facility.

i. ..... [ ] very important [ ] important [ ] moderate [ ] not much
ii. ..... [ ] very important [ ] important [ ] moderate [ ] not much
iii. ..... [ ] very important [ ] important [ ] moderate [ ] not much
iv. ..... [ ] very important [ ] important [ ] moderate [ ] not much

2. Please specify how effective are of the above systems to manage your facility:

i. ..... [ ] high [ ] medium [ ] low [ ] not much
ii. ..... [ ] high [ ] medium [ ] low [ ] not much
ii. ..... [ ] high [ ] medium [ ] low [ ] not much
iv. ..... [ ] high [ ] medium [ ] low [ ] not much

3. Please identify the possible sources of failure of the above systems.

i. ..... [ ] hardware [ ] software [ ] human error [ ] overload [] malicious attack [] authorization violation
ii. ..... [ ] hardware [ ] software [ ] human error [ ] overload [] malicious attack [] authorization violation

iii. ..... [ ] hardware [ ] software [ ] human error [ ] overload [] malicious attack [] authorization violation

iv. ..... [ ] hardware [ ] software [ ] human error [ ] overload [] malicious attack [] authorization violation

4. If you had failure of these systems in the past, please give some information on impact scale (e.g. few users, department, campus), chronology of the events (if available), how failure affected the output of your facility and how much time it took for recovery?

.....

.....

5. Please give other additional detail of these failures if possible (exact failure source, how it was detected, any known financial losses, etc.)?

.....

.....

6. Do you have backups for each of the CITI systems mentioned?

.....

.....

7. Do you have network layout, device specifications, uses statistics and throughput measurements available for research studies? [ ] Yes [ ] No

8. What kinds of malicious attacks do you have for your systems? Do you know the sources of these attacks? What kinds of remedial actions have you taken? What are the possible consequences?

.....

.....

9. Do you have any strategy for CITI failures? How and when such plans kick in?

.....

.....

10. Please give some detail about the possible obstacles for reliable Telecoms and IT services in your facility.

.....

.....

11. Please specify the failures which may have public safety concerns (such as failure of 911 service, medical emergency service, fire rescue service or police service).

.....

.....

Appendix B

The I2Sim Input File for Five Cell UBC Test Case:

```
main { 900, 5, 2, 4}
cluster BLK1 (26, 12, 4, 1) {
   channel CHE1 (0, 1, 1);
   channel CHE2 (0, 2, 2);
   substation SB1 (1, 3, 1, 1, 1, 0.4765, 0.4765, 0.047);
   channel CH1 (7, 8, 8);
   power_house PW1 (8, 9, 1, 1, 0, 0.286, 0.714);
   channel CH2 (12, 14, 14);
   channel CHE3 (0, 15, 15);
   water_station WS1 (14, 16, 1, 0.635, 0.365);
   channel CH3 (13, 19, 19);
   channel CHE4 (0, 20, 20);
   channel CH4 (18, 21, 21);
   steam_station SS1 (19, 22, 1, 1, 0, 0, 1);
   token_source I1 (1, 0, 1, 40000, 0, 0);
   token_source I2 (2, 0, 2, 40000, 0, 0);
   token_source I3 (15, 0, 15, 375, 0, 0);
   token_source I4 (20, 0, 20, 160, 0, 0);
}
cluster BLK2 (34, 6, 1, 27) {
   channel CHE1 (0, 27, 27);
   channel CH1 (0, 28, 28);
   channel CH2 (0, 29, 29);
   channel CH3 (0, 30, 30);
   channel CH4 (0, 31, 31);
   hospital H1 (27, 32, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0
   token_source I5 (27, 0, 27, 3.3, 0, 0);
}
link LNK1 (5, 28, 0, 0) { null }
link LNK2 (6, 31, 0, 0) { null }
link LNK3 (17, 29, 0, 0) { null }
link LNK4 (26, 30, 0, 0) { null }
```

The I2Sim Input File for Beth Israel Deaconess Medical Center Test Case:

```
main {1720, 5, 2, 6}
cluster BLK1 (40, 20, 6, 1) {
```

```
    channel CHE1 (0, 1, 1);
    channel CHE2 (0, 2, 2);
    citilink FML1 (17, 3, 3, 7, 1, 155, 107.6);
    substation SB1 (1, 4, 1, 1, 1, 0.47, 0.47, 0.046, 0.0164);

    channel CHE3 (0, 10, 10);
    channel CHE3 (0, 11, 11);
    cntrlcntr CT1 (10, 12, 1, 0.249, 0.249, 0.249, 0.249, 0.249);

    citilink FML2 (13, 18, 18, 3, 1, 155, 15.96);
    channel CH3 (8, 19, 19);
    citilink FML3 (16, 20, 20, 6, 1, 155, 107.6);
    power_house PW1 (19, 21, 1, 1, 0, 0.286, 0.714);
    channel CH5 (24, 26, 26);
    channel CHE4 (0, 27, 27);
    citilink FML4 (15, 28, 28, 5, 1, 155, 107.6);
    water_station WS1 (26, 29, 1, 0.635, 0.365);
    channel CH7 (25, 32, 32);
    channel CHE5 (0, 33, 33);
    channel CH8 (31, 34, 34);
    citilink FML5 (14, 35, 35, 4, 1, 155, 107.6);
    steam_station SS1 (32, 36, 1, 1, 0, 0, 1);
    token_source I1 (1, 0, 1, 40000, 0, 0);
    token_source I2 (2, 0, 2, 40000, 0, 0);
    token_source I3 (10, 0, 10, 622, 0, 0);
    token_source I3 (11, 0, 11, 50, 0, 0);
    token_source I4 (27, 0, 27, 375, 0, 0);
    token_source I5 (33, 0, 33, 160, 0, 0);
}
cluster BLK2 (40, 10, 1, 41) {
    channel CHE1 (0, 41, 41);
    channel CH1 (0, 42, 42);
    channel CH2 (0, 43, 43);
    channel CH3 (0, 44, 44);
    channel CH4 (0, 45, 45);
    citilink FML1 (78, 46, 46, 9, 1, 155, 2.44);
    hospital H1 (41, 47, 1, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 1, 0, 0, 0
    channel CH7 (0, 76, 76);
    channel CH8 (0, 77, 77);
    admincntr ADM1 (76, 78, 1, 1, 1);

    token_source I5 (41, 0, 41, 3.3, 0, 0);
}
```

```
link LNK1 (6, 42, 0, 0) { null }
link LNK2 (7, 45, 0, 0) { null }
link LNK3 (30, 43, 0, 0) { null }
link LNK4 (40, 44, 0, 0) { null }
link LNK5 (18, 76, 0, 0) { null }
link LNK6 (9, 77, 0, 0) { null }
```

The Fluid Flow Model (FFM) Input File for Beth Israel Deaconess Medical Center Data Network:

```
Nodes: 13
Queues: 24
Weight: 1
QMin: 10
QMax: 300
QCap: 1000
Pmax: 0.2
DuplexLinks: 12
FlowClasses: 17
Src     Dst     BW      Delay
0       2       62200000  0.02
1       2       15500000  0.05
2       3       15500000  0.05
2       4       15500000  0.05
2       5       15500000  0.05
2       6       15500000  0.05
2       7       15500000  0.05
3       8       150000000  0.01
3       9       15500000  0.05
9       10      200000000  0.01
9       11      200000000  0.01
9       12      15500000  0.05
Src     Dst     Flows
0       1       500
0       4       100
0       5       100
0       6       100
0       7       100
0       8       2000
0       10      1000
0       11      1000
0       12      500
```

```
1       4       100
1       5       100
1       6       100
1       7       100
8       10      1000
8       11      1000
8       12      1000
10      11      5000
```