Assessment of Interdependencies between Communication and Information Technology Infrastructure and other Critical Infrastructures from Public Failure Reports

Hafiz Abdur Rahman and Konstantin Beznosov
{rahmanha,beznosov}@ece.ubc.ca

Laboratory for Education and Research in Secure Systems Engineering
lersse.ece.ubc.ca

Technical report LERSSE-TR-2005-03*

February 24, 2006

Abstract

Failure in Communication and Information Technology Infrastructure (CITI) can disrupt the effective functionalities of many of the critical infrastructures. Conversely, failures in other infrastructures can also propagate to CITI and hence disrupt the operation of these interconnected systems. Understanding the origin of these failures, their propagation patterns and their impacts can give us important ideas about infrastructure interdependencies and can be used for secure and reliable infrastructure design and operation. In this research we have taken the approach to use public domain failure reports to identify these interdependencies. We have developed a methodology to collect and categorize these reports and defined a set of critical attributes to extract meaningful information from them. Using this approach we have analyzed 12 years of infrastructure failure reports from ACM's RISKS forum. Our results have shown interdependencies between CITI and other critical infrastructures in different dimensions, such as origin of failures, impacts of failures in spatial and temporal dimensions, how they have affected public safety; and how failures have propagated from one infrastructure to another. Results obtained from the analysis of real life failure cases, which happened over a considerable span of time, should be useful for infrastructure researchers and practitioners. This paper also discusses the difficulties while using public domain data in an academic research.

*This and other LERSSE publications can be found at lersse-dl.ece.ubc.ca

Contents

1 Introduction

Modern communication and information technology infrastructure (CITI) provides key links and services to many other critical infrastructures [1], such as telecommunication, 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. As such, failure in CITI, either due to an accident or caused by a malicious attack, can propagate to other infrastructures and can 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 disruption may led to huge disturbance in the public life of modern nation states. Volatile world situation increases these threats even further. As such, there are enormous concerns for secure and reliable operation of different critical infrastructures [1, 2]. One of the prerequisites for smooth operation of these interconnected infrastructures is to gain understanding of their interdependencies. By studying the origin of the infrastructure related failures and their propagation patterns, we can develop better understanding about their interdependencies. Such understanding can be useful for the decision makers and infrastructure operators for policy making and system design [3]. As CITI is at the core of many of these infrastructures' operation [2], it is very important to understand how faults propagate to and from CITI.

Since 1992, US telephone companies are required to submit major failure information to US Federal Communication Commission (FCC). Using FCC outage reports, a study was done on the failure pattern of Public Switch Telephone Networks (PSTN) [4]. According to 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 to develop good understanding of failure patterns and their interdependencies among different critical infrastructures. Data from the infrastructure service providers especially helpful, because they may give detailed information about systems' states, their control parameters, input and output specifications, operating assumptions, procedures and practices, back-up facilities and other physical and environmental constraints [5]. Both public and private infrastructure operators are quite reluctant to share these information with the research community [6]. The FCC Outage Report mentioned above is now accessible only to the users of FCC and US Department of Homeland Security (DHS) [5].

Given this reality, one possible alternative is to use public domain infrastructure failure reports, such as newspaper or mass media reports to develop an understanding of infrastructure interdependencies. There are two major difficulties in this approach. First, normally these failure reports have brief information content, second, they do not have any regular structure. Even though an individual report may not give much information about a specific failure, by studying large number of cases we can trace common trend among similar class of failures. Besides, to do analysis on these unstructured reports, we develop a methodology to classify these reports based on their failure type and extract meaningful information through some critical attributes. We collect 12 years (1994 to 2005) of failure data from Association for

Computing Machinery's (ACM) RISKS forum [7], which is the largest known public repository of these kinds of reports. Posting to this forum is moderated, which ensure certain level of quality and reliability. Even then, due to media biasness in reporting failure events and also biasness in the sampling of reported events to its actual population, there may have concern about usefulness of the statistics we get from a public forum like RISKS. Essentially, both of these concerns are related to public perception of risks. Research shows that despite partial information, public perception of risks are fairly accurate [8] and their judgement is not much different from that of experts [9]. As such, due to the open nature and substantial size of RISKS forum, we may assume accumulated reports represent similar statistical trend of real life events. To ensure authenticity of our selected cases, we give special importance to the verifiability of our selected report's sources. For instance, we give preference to the newspaper reports than to reports by private individuals. Our methodology is discussed in Approach and Method section. Doing so, we share our experience of using public domain data in academic research.

In this research we identify interdependencies between CITI and other infrastructures based on some key factors, such as, origin of failures, impact of failures in spatial and temporal dimensions, how these failures affect public safety and how failures propagate from CITI to other critical infrastructures and vice versa. More specifically we would like to get answer to the questions like, main causes of infrastructures' failures, nature of their impact; locality affected by them and their geographical locations, how their fatality changed over time; and how infrastructures are related to each other, etc. In the absence of any formal model between CITI and other critical infrastructures' interdependency, our findings should give useful ideas to the policy makers, practitioners and the researchers. In Section 2, we discuss previous works on CITI and other infrastructure failures. In Section 3, we give a brief overview of our own methodology. In Section 4, we summarize the results of our analysis. In Section 5, we discuss implications and usefulness of our findings in infrastructure safety and security analysis. Section 6 concludes this report discussing its contribution and with some future research directions.

2 Related Work

There have been some works to classify and interpret failures in computer based systems. One approach is to qualitatively understand risk and threat associated with the use of any computer system [10]. Some works focus on classifying vulnerabilities only within CITI [4, 11, 13]. Other approaches consider failure in infrastructure agnostic way [14]. Impact of failures on multiple infrastructures has also been discussed in [15]. Following section gives brief overview of these works.

Peter Neumann started Association for Computing Machinery's (ACM) RISKS forum [7] in 1985, to compile risks to the public in the use of computers and related systems. Later Neumann published a book (1994) named "Computer-Related Risks" [10]. In this book, he qualitatively analyzes a large collection of RISKS forum reports. Using these factual evidences he argues, origin of most of the failures are due to system design error, improper runtime conditions, human mistakes, natural causes

or due to deliberate malicious attacks. He draws attention to the safety and security risks associated with those failures. In his book he also discusses some risk reduction techniques, such as how to reduce system complexity using layered design technique, tamper resistance using trusted computing base, improve system reliability using system-engineering approach, etc.

Kuhn [4] proposes a failure classification scheme [4] to analyze public switched telephone network (PSTN) failure data. Using this scheme he analyzed two years of PSTN failure data (1992-1994) from US Federal Communication Commission (FCC). His analysis shows origin of failures by different categories and their impacts on PSTN operation. He explains high reliability of PSTN systems in terms of two factors - systems' interaction and coupling, where interaction refers to dependencies between components, and coupling refers to the flexibility.

Howard proposes a taxonomy based on attack types [11] in computer network and using that taxonomy he performs frequency analysis of more than 4000 security related incidents reported to Computer Emergency Readiness Team Coordination Center (CERT/CC). From the results of this analysis, he proposes a set of recommendations for government, vendors, CERT/CC, and individual users to improve security practices. Howard and Longstaff [12] extend this taxonomy to incorporate additional terms to include additional objects and attributes, such as site name, attack date and reporting time, etc.

Chakrabarti and Manimaran [13] propose a taxonomy to classify Internet infrastructure security failures. Their classification is based on 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 attack.

Scope of the above three works [4, 12, 13] are limited within CITI and as such give limited idea to classify and interpret failures which affects other infrastructures beyond CITI. Rinaldi et al [15] propose a taxonomy based on following six functional dimensions to address cross infrastructure interdependency issues.

- types of interdependencies (e.g., physical, cyber, logical)
- infrastructure environment (e.g., business, economic, health care)
- coupling and response behavior (e.g., adaptive, loose, tight)
- infrastructure characteristics (e.g., temporal, spatial, organizational)
- types of failures (e.g., common cause, cascading, escalating)
- state of operations. (e.g., normal, stressed, repaired)

Their failure based interdependency classification is very restrictive (common cause, cascading, escalating) and gives very limited number of options to analyze RISKS forum failure reports. Classification of failures in system agnostic way have been discussed by Avizienis et al [14]. We find many of their failure classifications [14] applicable to model failures between CITI and other critical infrastructures. We have used some of those failure categories in our work. However, Rinaldi et al [15] interdependency types have some similarity with our failure classification in three functional layers (Physical, Network and IT Service).

3 Approach and Methods

A four step methodology has been followed to collect and analyze failure reports. This started with systematically collecting failure cases from RISKS forum, categorizing those reports based on their failure type, extracting useful information from those classified reports and then performing analysis on the extracted information. Following sections discuss these steps in detail.

3.1 Data Collection: Collection of useful failure reports from the huge volume of RISKS forum data was the most important step to start with. Selection criteria was simple. Only those reports will be collected where failure was originated from CITI and affected other critical infrastructures (including CITI); or 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 specified performance requirements [16]. Failure may be the result of one or many faults. Fault is a defect in a hardware device or component; or an incorrect step, process or data definition in a computer program [16]. In our study, failure attributed to critical infrastructures. Following infrastructures (Table 1) are considered critical infrastructures based on a US Congress document [1] on critical infrastructure identification.

Critical Infrastructures
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

Table 1: List of Critical Infrastructures

This apparent simple task of selecting appropriate set of reports became complicated due to the subtleties associated with each of the reports. Following two examples explain this complicacy. A selected report is as follows:

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, RISKS (16, 58)

The above example clearly shows that fault in Electrical Power System due to human error caused a failure in Telecommunication Infrastructure. The report has a clear reference to a newspaper source. In contrast to the above report, following is an example of a report that was not selected:

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 this report, failure in Electrical Power System is not clearly related to CITI. There is no clear reference to place name and also there is an undefined term 17000 in the report.

3.2 Fault Classification: Our next step was to categorize collected reports based on their nature of failure. After analyzing the reports, we have found following thirteen types of faults (Table 2) that can best describe the origin of infrastructure failures. These fault types were derived from the taxonomies discussed in Section 2. Many of their definitions can be found in [14]. Further break down of each of these thirteen fault cases could be desirable. For instance, Kuhn [4] subdivides Hardware failures into Cable component, Power supplies, Facility Component and Clock or clock synchronization. Such granularity in fault analysis is possible if someone has access to official version of failure reports, which might have much detailed information. However, we were limited by the description of public domain failure reports where many detail are often missing and there was no mechanism to get any additional information.

Fault Type	**Meaning**
Natural Fault	Physical faults that are caused by natural phenomena without human participation.
Hardware Fault	All fault classes that affect hardware.
Vandalism	Sabotage or other intentional damage.
Overload	Service demand exceed the designed system capacity.
Malicious Logic Fault	These include Torojan horses, logic or timing bombs, viruses, worms, or zombies.
Intrusion Attempt	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.
Content Failure	Content of the information delivered at the service interface deviates from implementing the system function.
Timing Failure	Time of arrival or the duration of the information delivered at the service interface deviates from implementing the system function.
Halt Failure	Service is halted.
Erratic Failure	Service is delivered, but is erratic.
Human Error	Non-deliberate faults introduced due to mistake.
Configuration Fault	Fault caused by inappropriate configuration of hardware or software.
Software Fault	Fault caused due to the error in software system.

Table 2: Fault Types

Careful observation shows that these fault types (Table 2) could be grouped into three layers. Natural Fault, Hardware Fault and Vandalism are related to the failure in physical systems. We named them physical layer faults (Class A). Physical layer is at the bottom of our classification hierarchy. Overload, Malicious Logic Fault and Intrusion Attempt faults are related to basic network services. We named them network layer faults (Class B). Network service layer is just above the physical layer. Content Failure, Timing Failure, Halt Failure and Erratic Failure are service level failures which effect the functional dependencies between infrastructures. We named them IT Service layer faults (Class C). IT Service Layer is the top most layer in our fault class hierarchy. Human Error, Configuration Fault and Software Fault are layer independent and can disrupt the functionality of physical system, network connectivity or functional dependency between infrastructures. Following table shows the layer based classification of faults (Table 3):

Potential benefits of layer based fault classification are, failure at each layer has their own characteristics and requires specific type of strategy and technology for fault detection and prevention. For instance, our findings have shown that physical layer faults are mostly related to reliability and survivability aspect of the infrastruc-

Fault Classes	**Fault Types**
Physical Layer	Natural Fault, Hardware Fault, Vandalism.
Network Layer	Overload, Malicious Logic Fault, Intrusion Attempt.
IT Service Layer	Content Failure, Timing Failure, Halt Failure, Erratic Failure.
Layer Independent	Human Error, Configuration Fault, Software Fault.

Table 3: Fault Classes

tures. General strategy to deal with these problems are redundant devices, backup physical channels, etc. [17]. Similarly, for network level, different kinds of protection techniques are used. These include secure routing protocols, secure domain name systems, authentication technologies, firewall, anti virus tools and intrusion detection systems [13]. IT Service level failure management requires more specialized tools and techniques. For instance, air transportation services require specialized hardware and software tools for their systems' reliability. [18].

This approach also enabled us to use many of the well understood conceptual terms to analyze failures related to each layer. Such as mean time between service failure, mean time to hardware repair, mean time between interruption of packet transmission, etc. For instance, Hagin [19] uses such layered approach for reliability and survivability analysis of X.25/X.75 switching network.

3.3 Feature Extraction: After categorizing a failure report to a particular failure class, we have extracted key features from each of these reports using a set of key attributes. Many of these features intent to capture extent of failures, impact of failures in spatial and temporal dimensions, how these failures affect public safety and how failures propagate from CITI to other critical infrastructures and vise versa. Table 4 group those in their respective category. Table 5 lists each feature with its meaning. There are few features who's values needs some more explanation which are discussed below.

Feature Dimension	**Feature Name**
Extent of failure	Fault Class, Degree of Impact, Fault Intent, Fault Type
Impact (spatial)	Country, Locality, Affected Sites
Impact (temporal)	Date, Duration
Public safety	Public Safety
Failure propagation	Source Infrastructure, Affected Infrastructures

Table 4: Feature those capture different failure dimensions

Degree of Impact is a feature intent to capture severity of a failure. Kuhn [4] uses customer minutes (product of average no of customer affected and average outage duration) to measure the severity of failure in PSTN network. This is possible only because, each FCC report has to include the date, time, duration and the number of affected customers (Page 33) [4]. Unlike FCC, our failure reports did not have such uniform and universal impact dimensions. As such, some of our impact analysis was judgemental (subjective). Reading the failure case we tried to understand how many people, system were impacted. Based on that understanding a Degree of Impact was assigned. "High" were those events, which massively affect the functionality of CITI other critical infrastructures, "Medium" were those, which moderately affect these infrastructures, and "Low" was for those failure which have small impact on the operation of these critical infrastructures. Clemen et al [20] show that in the absence of uniformity of data, expert judgement can predict risk with reasonable accuracy. Following three examples show assignment of Degree of Impact by subjective judgement (A.1, B.1 and C.3 are record numbers in our failure database).

Degree of Impact - High (A.1) - 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 (B.1) - 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 (C.3) - 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 4,901 other

messages. The system had to be shut down for five hours

Another feature is Report Accuracy, where based on the source type we subjectively assigned an accuracy rating on a scale of 10. For each of these reports, 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 detail was given, it had 7 or 8 points. All newspaper reports had 5 or 6 points. Reports from individuals, which had difficulty to verify, were normally given less than 5 points. Higher rating was given to a report of a particular class, if the report fulfilled most of our additional attribute criteria. For instance, if a newspaper report had most of the information like severity, duration, financial impact, description of fault origin, etc., then it was given 6 points. Otherwise, it was given 5 points.

All extracted data were kept in a failure database as discussed in Section 4.

3.4 Data Analysis: Objective of our analysis was to identify interdependencies between CITI and other infrastructures based on some key dimensions, such as, origin of failures, impact of failures in spatial and temporal dimensions, how these failures affect public safety, etc. We have used frequency based approach to quantify these results from the extracted features' database (Section 3.3, 4) and tried to get answer to the questions such as, most likely cause of infrastructure failures, types of locality affected by them, what was the recent trend of these failures and their implications on public safety. However, due to diversity of infrastructure types and incompleteness (missing attributes) of many of the failure reports, some of the results of our analysis were qualitative (judgemental). For instance, only few reports they have both Duration and Affect Site #. Similarly, Numbers of people affected are rarely mentioned in these reports. As such, it was not possible to use any uniform concept like "Customer minutes" [4] for all failure cases of our analysis. Instead, we have used "Degree of Impact" and have done frequency analysis on this field. Assignment of High, Medium and Low values to this field was not quantitative as discussed in the previous section. However, judgemental approach is a valid approach in these types of cases as discussed by Clemen et al [20].

Feature Name	Meaning
Title	Title of the report. Most often this is same as the original report.
Fault Class	Type of the fault class A, B or C.
Date	Date of the failure report.
Country	Where fault incident originated. For global fault it is World.
Locality	How much area was affected. Could be an Organization, a City, a Region (a big part of the country), a Country, a Continent or 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 NS2 [21] or similar network simulator.
Fault Intent	Fault could be Intentional, which occur due to deliberate and malicious attempts by any individual or groups, or Unintentional due to human error or system flaw
Fault Type	Fault type is one of the thirteen type mentioned in Table 2
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 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 we assign an accuracy rating on a scale of 10.
Fault Origin	An qualitative assessment about the origin of fault
Source Infrastructure	Is one of the critical infrastructure from Table 1
Affected Infrastructures	Is one of the critical infrastructure from Table 1
Affected Industry Sectors	Description of the industry sectors affected by the failure
Comment	comments to specify some interesting aspects of these faults.

Table 5: Extracted features and their meaning

4 Failure Database

Collected cases and their extracted features were compiled in a MS Excel database. Each record in this database has a code number. Code number is a sequential number coupled with its fault type (Class A, B or C). Data is organized in three groups based on their failure classes and sorted in ascending order based on report date. There are 121 collected cases over 8 years (from 1994 to 2001). Frequency of collected cases has changed over time as shown in the following graph. It shows linear rise in reported infrastructure failure cases over time. One possible reason for this is, critical infrastructures are increasingly becoming dependent on CITI for their critical functionalities.

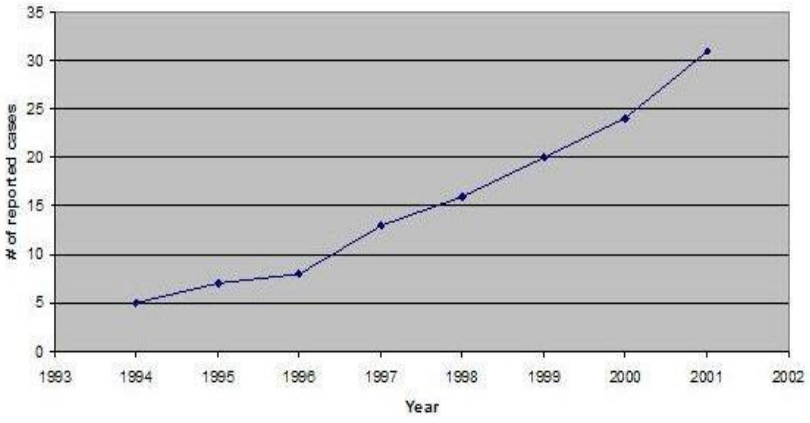

Figure 1: Reported failures over time

Each of the following attribute has its own legal value as shown in Table 6.

Field Name	**Legal Values**
Code	Code number is a sequential number coupled with its fault type (e.g., A.1)
Fault Class	A, B or C
Title	Text String
Date	MM/DD/YYYY
Country	Country Name / World
Locality	Organization / City / Region / Country / Continent / World
Degree of Impact	High / Medium / Low
Network Trace	Yes / No / Unknown
Simulation	Yes / No / Unsure
Fault Intent	Intentional / Unintentional / Unknown
Fault Type	One of the thirteen fault type (Table 2)
Duration	# Hour
Financial Impact	# Million USD
Public Safety	Yes / No / Unknown
Affected Sites	# / Unknown
Description	Text String
Report Source	Text String
Report Accuracy	#
Fault Origin	Text String
Source Infrastructure	One of the infrastructure from Table 1
Affected Infrastructures	One of the infrastructure from Table 1
Affected Industry Sectors	Text String
Comment	Text String

Table 6: Legal values for Failure Database

A sample failure report with the extracted feature values is shown below. All the analytical results are also kept within the same MS Excel spreadsheet. This spreadsheet is available from the authors of this paper.

A.1	**Ground-cable removal blows Iowa City phone system upgrade**			
Date	**Country**	**Locality**	**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 Type	Human Error			
Fault Origin	Fault in electrical system due to human error.			
Source Infrastructure	Electrical Power System			
Affected Infrastructures	Telecommunication Infrastructure			
Affected Industry Sectors	All kinds of industries of Iowa City			
Comment	Lack of detailed planning			

Table 7: A Sample Database Record

5 Results

In this section, we present our findings from failure database. Table 8 summarizes the number and percentage of failures based on their intention, impact scale, location and public safety concern.

	Physical Layer (A)		**Network Layer (B)**		**IT Service Layer (C)**		**Total**	
Total/category	46	38%	31	26%	44	36%	121	100%
Intention								
Intentional	2	4%	21	68%	0	0%	23	19%
Unintentional	39	85%	10	32%	41	93%	90	74%
Unknown	5	11%	0	0%	3	7%	8	7%
Impact Scale								
High	37	80%	21	68%	26	59%	84	69%
Medium	8	17%	8	26%	10	23%	26	21%
Low	1	2%	2	6%	8	18%	11	9%
Location								
US/Canada	26	57%	18	58%	30	68%	74	61%
Other	20	43%	13	42%	14	32%	47	39%
World	2	4%	3	10%	0	0%	5	4%
Public Safety								
Concern	15	33%	6	19%	8	18%	29	24%
No Concern	20	43%	18	58%	31	70%	69	57%
Unknown	11	24%	7	23%	5	11%	23	19%

Table 8: Effect of failures by intention, impact scale, location and public safety

Following figures show the results obtained from Table 8.

5.1 Failure Class Distribution: Figure 2 shows how failures are distributed in each layer. Most of the failures are found to be in the IT Service layer, then in Physical and in Network Layer.

Figure 3 shows percentage of failure based on fault origin. According to this classification, hardware fault is the predominant class of failure (26%), next to it is the software fault (13%). Human error is also responsible for significant number of failures (11%).

Figure 4 shows that origins of most of the CITI failures are unintentional (74%).

5.2 Impact of Failure: If we combine the findings of Figure 2 and Figure 4, we observe an interesting picture as shown in Figure 5. This shows intentional failures are mostly concentrated in the network layer. This is because "Malicious Logic Fault" and "Intrusion Attempt" are by definition intentional.

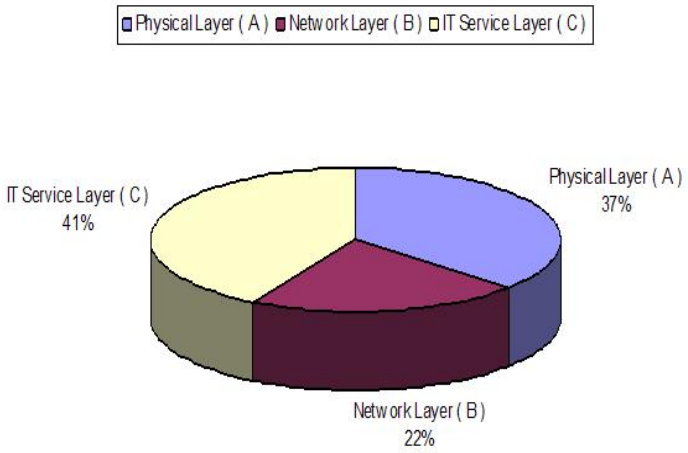

Figure 2: Failure class distribution

Figure 6 shows another interesting aspect of interdependencies between infrastructures. More number of high impact failures are in the lower layer than in the upper layer. One possible reason is, any failure in the lower layer may lead to severe consequences, than the failure in the upper layer. For instance, physical layer failure leads to both network and service layer failure. However, service level failure may not imply network or physical layer disruption.

5.3 Public Safety Concerns: Figure 7 shows that public safety concern is more for lower layer failures than for upper layer. For example, disconnected cable in the telephone network may lead to failure of 911 services, but disruption in air transportation service may not lead to failure of 911 services.

However, Figure 8 shows another interesting aspect of relation between public safety and failure impact. There are more numbers of High, Medium and Low failure impact cases where public safety is not a concern (pink line), than where it is a concern (black line). So, severe failure may not always imply great concern in public safety service.

5.4 Change of Degree of Impact over Time: Figure 9 shows degrees of

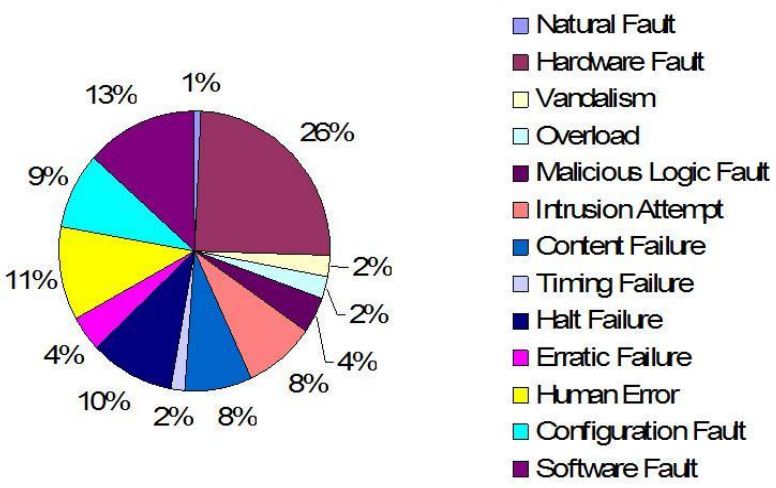

Figure 3: Faults those lead to Infrastructure failure

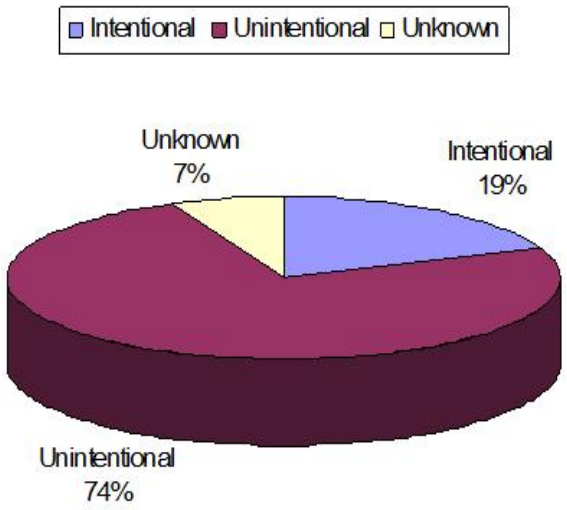

Figure 4: Failure type distribution

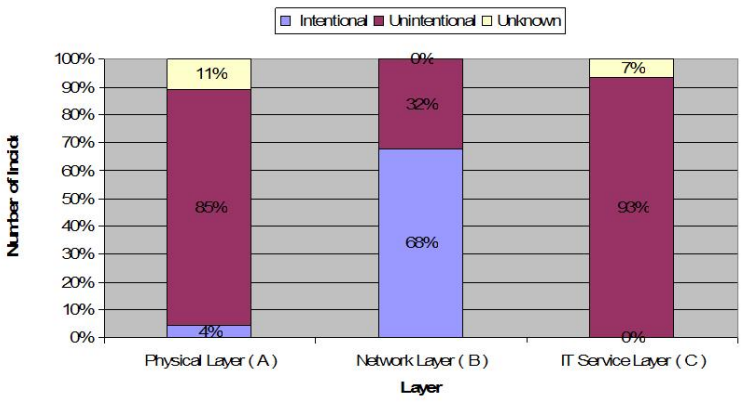

Figure 5: Failure types for each layer.

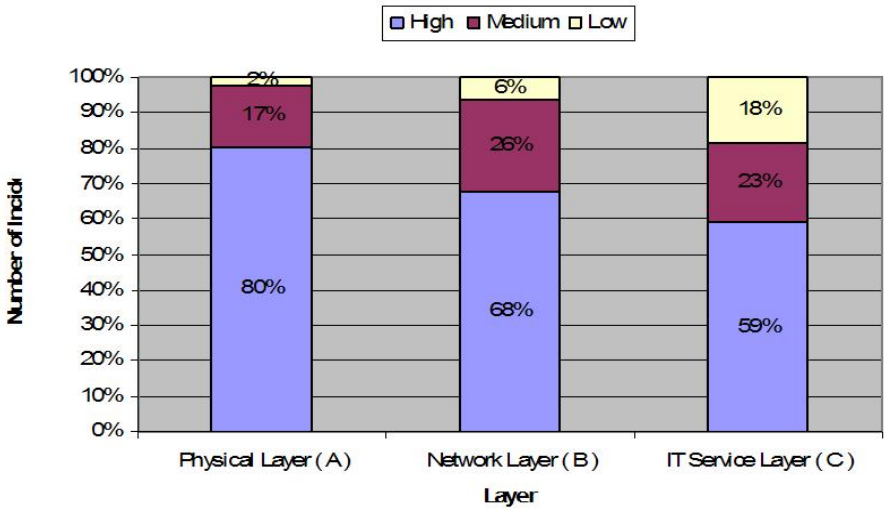

Figure 6: Impact of failure related to each layer.

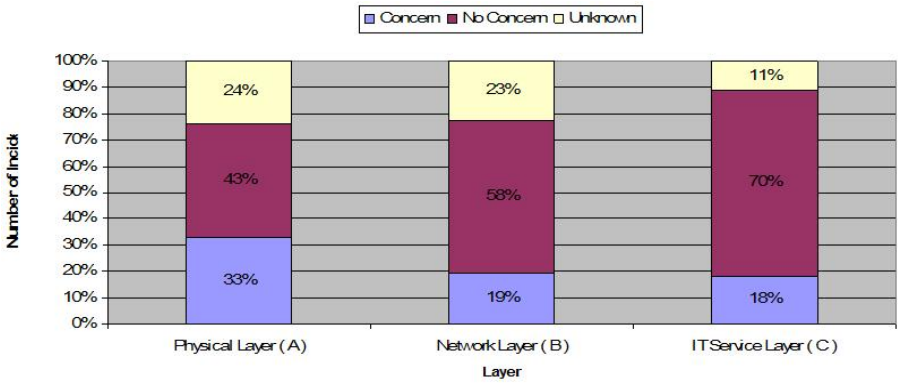

Figure 7: Public safety related to each layer.

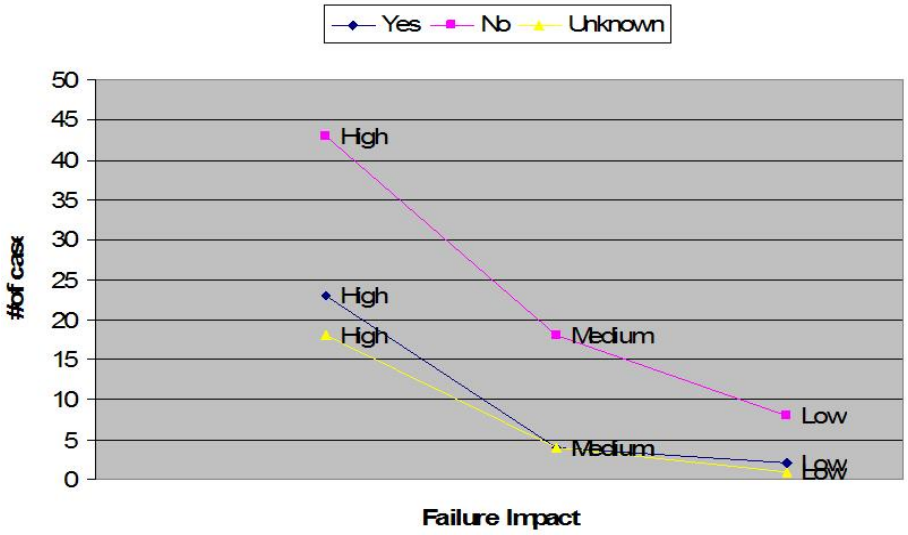

Figure 8: Public Safety vs. Failure Impact.

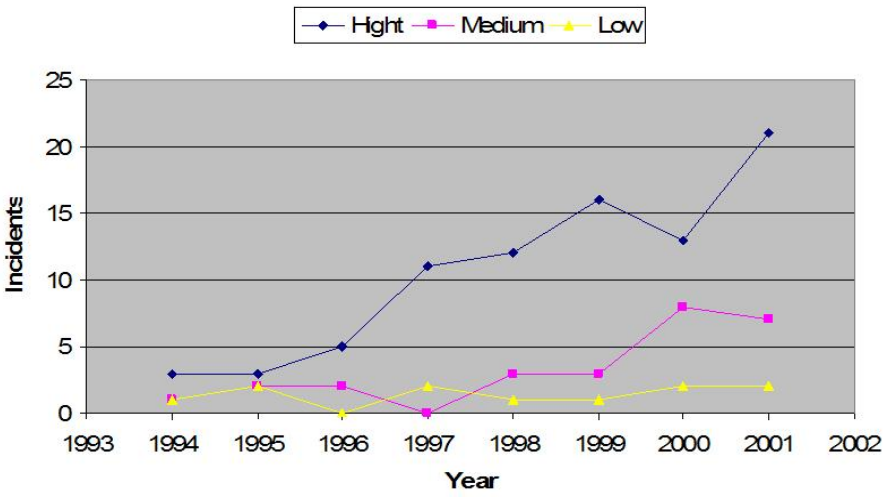

Figure 9: Change of degree of impact over time.

impact (severity) of infrastructure failure are on the rise.

However, this higher fatality may not be due to malicious attacks, because Figure 10 shows that this severity in failure is largely contributed by unintentional failures.

5.5 Locality affected by CITI Failures: Figure 11 shows that most of the time CITI failures ends up beyond organization boundary (65%) that affects the life of general public. Crossing the national boundary is very unlikely, unless a failure is targeted internationally.

Figure 12 and 13 show the infrastructure failures with respect to their concentration in geographical locations. Figure 12 includes international cases (eg., worm attack). Whereas, Figure 13 excludes those cases. In both figures, most of the reported failures (above 60%) have taken place in North America (US/Canada). One possible explanation is, this region has highest number of computer uses than any other parts in the world.

5.6 Interdependency between Infrastructures: Figure 14 shows, most of the time CITI failures are originated from within CITI infrastructure. Figure 15 shows, CITI failures mostly affect CITI itself, Banking and Finance and Multiple

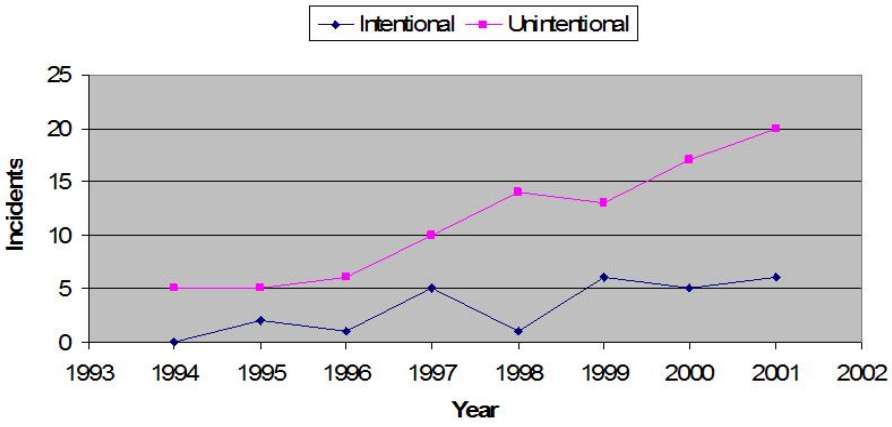

Figure 10: Change of intentional and unintentional failure over time.

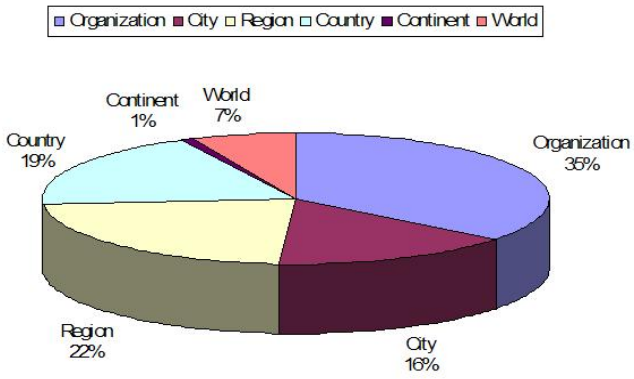

Figure 11: Localities affected by infrastructure failures.

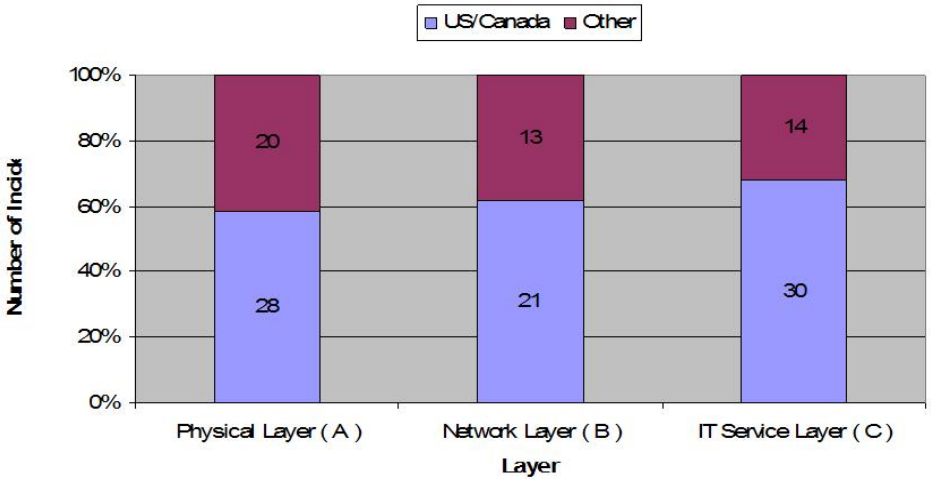

Figure 12: Failure location (includes International).

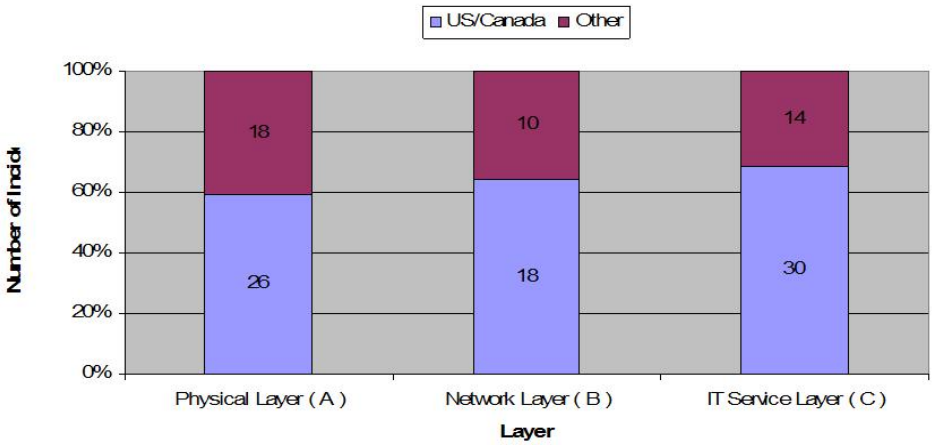

Figure 13: Failure Location (excludes International).

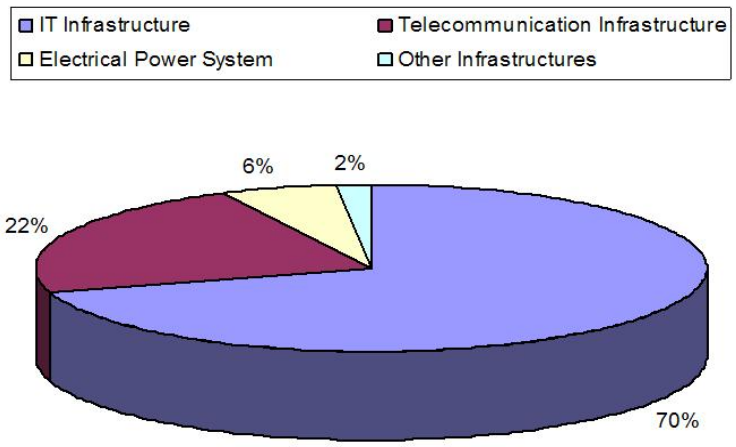

Figure 14: Failure those affect CITI.

Infrastructures(more than one infrastructures).

How failures propagate from one infrastructure to another is shown in the following Figure 16. (Under Construction)

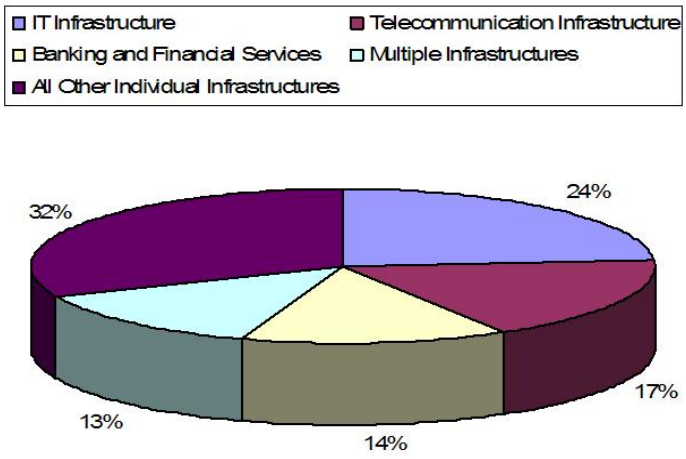

Figure 15: Infrastructure affected by CITI failures.

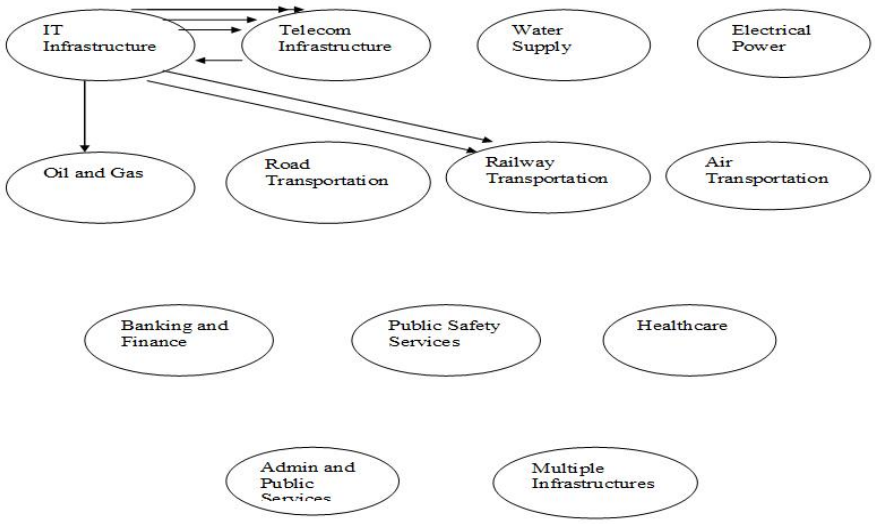

Figure 16: Propagation of failure from one infrastructure to another. (Under Construction).

6 Discussions

Results obtained in Section 5 presents some important facts about critical infrastructures and their interdependencies during their normal course of operation (we did not find enough cases related to catastrophic disaster). Following are some of the observations:

Figure 4 shows that nearly three-fourth of failure cases are due to accidental cause, Figure 3 shows that almost half of these failures are due to system failure (Hardware 26% + Software 13% + Configuration 9% = 48%) and Figure 5 shows that most of the system failures (hardware, software) are unintentional. This implies that, infrastructures are mostly vulnerable to accidental causes (non malicious). Accidental reasons include hardware or software fault, configuration problem, human error, etc. In contrast to that, Malicious Logic Fault, Intrusion Attempt and Vandalism are only accounted for less than 15%. This leads us to the conclusion that more focus has to be paid on system reliability.

Figure 9 shows that numbers of high impact failures are rising and Figure 10 shows that they are mostly contributed by unintentional failures. This implies poor system design, implementation or operations are also on the rise.

Figure 7 shows that public safety is largely related to system failure (lower layer failures have higher impact). As system failure is on the rise, there are increasing concerns for public safety.

Figure 11 shows that failures are most likely to affect beyond their organization boundary (65%) and as such, may affect the life of others. This implies infrastructure failures are growing concern for general people. Figure 12 and Figure 13 shows that North America (US/Canada) is especially vulnerable region (60%) due to growing dependency on critical infrastructure related services.

Figure 14 shows CITI failures are most likely to originate (92%) within CITI and impact CITI itself (Figure 15). As such, improving the technique of CITI infrastructure design, implementation and management can ensure greater reliability and safety.

7 Conclusions

Functionality of modern states largely depends on the smooth operation of CITI and other critical infrastructures. Any disruption in their operation may result in greater disturbance in public life. Therefore, understanding of their failure patterns and interdependencies is a requirement for their smooth operation. Such understanding may help infrastructure operator, researcher and public decision maker in general. In this research, we have explored a less studied proposition of using public domain data to understand CITI and other critical infrastructures' interdependencies. Using this approach we have studied public domain infrastructure failure report for considerable span of time (12 years). To our knowledge this is the first attempt to understand relation between CITI and other critical infrastructures from the failure cases (using either public or protected data sources). Results obtained from this analysis of real life failure cases should be useful for future research on critical infrastructures.

We have identified empirical pattern for the origin of infrastructure failures, their propagation pattern, their impacts on public life, and their historical trends. We have developed a failure database during our research. This database will also be useful reference for the validation of any model related to critical infrastructures. Also, the method we have used for public domain data collection, classification and analysis should be useful for any research involving such sources. Our future plan is to get access to the infrastructure failure data from the provider sources and compare results obtained from those sources to the results from pubic data sources. We would also like to do some infrastructure simulation and compare those results with the findings of this paper.

References

[1] Critical Infrastructure and Key Assets: Definition and Identification
http://www.fas.org/sgp/crs/RL32631.pdf

[2] Executive Order on Critical Infrastructure Protection
http://www.whitehouse.gov/news/releases/2001/10/20011016-12.html

[3] Joint Infrastructure Interdependencies Research Program (JIIRP)
http://www.nserc.ca/programs/jiirp_e.htm

[4] D. R. Kuhn, "Sources of failure in the public switched telephone network," IEEE Computer, v 30, n 4, April 1997, p 31-36

[5] FCC Network Outage Reporting System - User Manual
http://www.fcc.gov/oet/outage/nors_manual.pdf

[6] Eugene H. Spafford, Congressional Testimony, 10 October 2001
http://www.house.gov/science/full/oct10/spafford.htm

[7] The RISKS Forum:
http://catless.ncl.ac.uk/Risks

[8] Baruch Fischhoff, Paul Slovic, Sarah Lichtenstein, "Lay Foibles and Expert Fables in Judgments about Risk", The American Statistician, Vol. 36, No. 3, Part 2: Proceedings of the Sixth Symposium on Statistics and the Environment. (Aug., 1982), pp. 240-255.

[9] Gene Rowe, George Wright, "Differences in Expert and Lay Judgments of Risk: Myth or Reality?", Risk Analysis, Volume 21, Number 2, April 2001, pp. 341-356

[10] Peter G. Neumann, Computer-Related Risks, Addison-Wesley Professional, October 18, 1994, ISBN: 020155805X

[11] John D. Howard,"An analysis of security incidents on the Internet 1989-1995", Ph.D. Thesis, Carnegie Mellon University, 1997

[12] John D. Howard, Thomas A. Longstaff, "A Common Language for Computer Security Incidents", Sandia National Laboratories techinical report SAND98-8997, 1998.

[13] Anirban Chakrabarti, G. Manimaran, "Internet infrastructure security: A Taxonomy", IEEE Network, v 16, n 6,November/December, 2002, p 13-21

[14] A. Avizienis, J.-C. Laprie, B. Randell, C. Landwehr, "Basic Concepts and Taxonomy of Dependable and Secure Computing", IEEE Transactions on Dependable and Secure Computing, v 1, n 1, Jan 2004, p 11-33

[15] Steven M. Rinaldi, James P. Peerenboom, Terrence K. Kelly,"Identifying, understanding, and analyzing critical infrastructure interdependencies", IEEE Control Systems Magazine, v 21, n 6, December, 2001, p 11-25

[16] IEEE Standard Glossary of Software Engineering Terminology. IEEE Standard 610.12-1990, 1990

[17] Edward E. Balkovich, Robert H. Anderson, "Critical Infrastructures Will Remain Vulnerable: Neighbourhoods Must Fend for Themselves", International Journal of Critical Infrastructures, 2004 - Vol. 1, No.1 p. 8 - 19

[18] Barry Kirwan, "The role of the controller in the accelerating industry of air traffic management", Safety Science, v 37, n 2-3, 2001, p 151-185

[19] Alexander A. Hagin, "Performability, Reliability, and Survivability of Communication Networks: System of Methods and Models for Evaluation", Proceedings of International Conference on Distributed Computing Systems, 1994, p 562-573

[20] Robert T. Clemen, Gregory W. Fischer, Robert L. Winkler, "Assessing dependence: some experimental results", Management Science, v 46, n 8, Aug 2000, p 1100-1115

[21] NS2 - Network Simulator
http://www.isi.edu/nsnam/ns/