Vulnerability analysis for multiple critical infrastructure sectors using interdependencies modeling and network analysis

Lin, Jiwei; Tai, Kang; Tiong, Robert Lee Kong; Sim, Mong Soon

2016

Lin, J., Tai, K., Tiong, R. L. K., & Sim, M. S. (2016). Vulnerability analysis for multiple critical infrastructure sectors using interdependencies modeling and network analysis. 6th Asian-Pacific Symposium on Structural Reliability and its Applications (APSSRA6).

https://hdl.handle.net/10356/80968

Vulnerability analysis for multiple critical infrastructure sectors using interdependencies modeling and network analysis

J. Lin[1*], K. Tai[2], R.L.K. Tiong[3] & M.S. Sim[4]

[1]*Institute of Catastrophe Risk Management (ICRM), Interdisciplinary Graduate School, Nanyang Technological University, 50 Nanyang Ave, Singapore 639798, Singapore. linj0068@e.ntu.edu.sg;*
[2]*School of Mechanical and Aerospace Engineering, Nanyang Technological University, 50 Nanyang Ave, Singapore 639798, Singapore. mktai@ntu.edu.sg;*
[3]*School of Civil and Environmental Engineering, Nanyang Technological University, 50 Nanyang Ave, Singapore 639798, Singapore. clktiong@ntu.edu.sg;*
[4]*DSO National Laboratories, 20 Science Park Drive, Singapore 118230, Singapore. smongsoo@dso.org.sg;*
**Corresponding author.*

Abstract: Critical infrastructures like electric power plants, oil and gas supply, telecommunication, and the internet are essential systems for the continuous functioning of modern society. They are highly interconnected networks that are highly interdependent and any failure can cascade through the network, escalating to devastating impact on the economy. In this study, a combination of input-output (I/O) model and network analysis will be used for analyzing the vulnerabilities in critical infrastructure networks. The objective is to show how the critical infrastructure interdependencies model can be applied to disruptions/failures in two or more critical infrastructures and to evaluate the resulting impact on a country based on linkages established from its economic input-output table. The worst case scenario of a possible catastrophic event can therefore be simulated using this model. The methodology demonstrates how partial failure of two or more critical infrastructures contributes to higher overall impact on an economy and may provide useful insights for infrastructure stakeholders and policymakers.

Keywords: critical infrastructure, interdependencies, network analysis, input-output, perfect storm.

1. Introduction

In the modern world, critical infrastructure is a network of independent, man-made systems and processes that function collaboratively and synergistically to produce and distribute a continuous flow of essential goods and services (The President’s Commission on Critical Infrastructure Protection, 1997) essential to the defense, economic security and the smooth functioning of government and the society as a whole.

Critical infrastructure (which includes electrical, water, transportation, financial, etc) has been in the spotlight in recent years due to the immense amount of attention captured through different events ranging from natural disasters (2011 Tohoku Earthquake) to man-made disasters (terrorist attack in 11th September 2001) to potential cyber attacks like “Stuxnet”, a malicious computer worm that was discovered to target PLCs of machinery systems in nuclear power stations to collect information and cause disruptions. These events usually set off a domino effect propagated to other infrastructures. A disruption in telecommunication network may cause disruptions in the ticketing system in the airport (transportation), delays in providing ATM or bank services (finance) and loss in telephone communication (telecommunication). This is an example of a single disruption affecting different infrastructures due to their interconnected nature. Infrastructure networks do not work independently as they require interaction with some other infrastructure networks for themselves to be operational.

Critical infrastructure analysis is usually performed by analyzing interdependencies according to historical data and expert knowledge of the critical infrastructure. It can identify significant failure patterns and quantify interdependency strength in the network. This will lead to an informed decision making process for infrastructure owner and to provide risk analysis on infrastructure that might be affected by possible disasters. It is usually very difficult to identify all interdependencies among critical infrastructures under normal operation, because some interdependent relationships are undetectable by using standard data collection (Laefer, 2006) or emerge only after the occurrence of a disruptive event (Krimgold 2006). Therefore, in this paper, the authors construct the interdependencies of critical infrastructures using an I/O model and incorporate two physical infrastructure networks to simulate disruption. The I/O model serves to describe the interdependencies among critical infrastructure systems while the physical infrastructure network describe the network structure of the actual critical infrastructure.

2. Methodology

2.1 Input-output model as critical infrastructure interdependencies

The Leontief I/O model is a matrix of raw economic data collected by companies or governments to study the relationships between suppliers and producers to meet consumer demand. The I/O model is usually constructed from an observed set of data (e.g National I/O table) for one economic area (for example, a country's economy) (Miller et al. 2009). They can also be used to represent the interdependencies between different sectors of a national economy or different regional economies (William & Thijs, 2009).

The nature of I/O model provides a possible way to analyse the economy as an interconnected system of industries that can be directly or indirectly affected by one industry sector or another. This enables the estimation of economic impacts of any changes to the economy based on these inter-industry transactions. The ability to analyse and capture the economy's direct and indirect reaction to change in the economic environment (any disruption) makes I/O unique. This advantage is helpful in the current research as critical infrastructure network analysis requires a model that is able to link infrastructures together, according to interdependencies (Rinaldi et al. 2001, Lam et al. 2013, Kizhakkedath et al 2013, and Tai et al. 2013) and will be especially useful for researchers, critical infrastructure stakeholders, and governmental agencies of concern.

The I/O model mathematical formulation is based on demand-pull I/O quantity model (Miller et al. 2009). If x_i represents the total output (production) of sector i, z_{ij} represents the monetary values of the transaction between pairs of sectors (from sector *i* to sector *j*) and final demand f_i represents total final demand for sector *i*'s product, we may write a simple equation accounting for the way in which sector *i* sells its product as a form of input to other sectors and to the final demand as follows:

$$x_i = z_{i1} + z_{i2} + \cdots + z_{in} + f_i = \sum_{j=1}^{n} z_{ij} + f_i \quad (1)$$

In the I/O model, the technical coefficient a_{ij} is defined as the value of product *i* required as input to produce a unit value of product *j*, and can be expressed as:

$$a_{ij} = \frac{z_{ij}}{x_j} \quad (2)$$

Substituting equation (2) into (1), the sales of the output of each of the *n* sectors can be written in matrix form as in equation (3), simplified to equation (4) and rearranged to equation (5):

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_n \end{pmatrix} \quad (3)$$

$$\mathrm{X} = \mathrm{AX} + \mathrm{F} \quad (4)$$

$$(\mathrm{I} - \mathrm{A})\mathrm{X} = \mathrm{F} \quad (5)$$

The matrix A is usually called the direct requirement matrix, which represents the technological state of the economy for producing products for all sectors. And finally, applying standard matrix algebra to solve for X (the output for all sectors) in equation (5):

$$\mathrm{X} = (\mathrm{I} - \mathrm{A})^{-1}\mathrm{F} \quad (6)$$

Equation (6) represents the solution of the I/O model. X represents the output of all sectors due to the effect of the final demand F. $(I - A)^{-1}$ is known as the Leontief inverse or total requirement matrix. This inverse will also be referred to as the interdependency matrix that controls the direct and indirect effects due to any change in final demand F. The interdependency matrix will be illustrated as I/O interdependency model as shown in Fig. 1. Since equation (6) is a linear equation, it can also be represented by

$$\Delta X = (I - A)^{-1} \Delta F \qquad (7)$$

where the change in output X is directly proportional to the change in final demand F (Santos 2006). ΔF is very important for our study as I/O models are based on the assumption that export demand (or the ability of industries to sell to the external economy) is the engine that generates activity in the regional economy. Changes in final demand, ΔF, infuse local industries with new funds (in the case of positive change in demand), which increase output X (Liu et al. 2004). However, a negative change in final demand F will vice versa lead to decrease in output X. This is the main attribute of I/O model where a disruption, ΔF, to the critical infrastructure will result in impact in the form of decreased output X to all critical infrastructure connected within the I/O model.

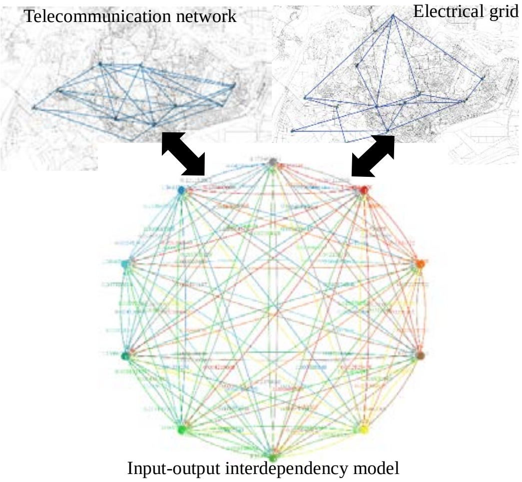

Fig. 1. Interfacing Physical infrastructure networks with I/O interdependency model

The analysis of economic impact (or losses) is made possible through the use of economic I/O data published. In this study, we will be using Singapore economic I/O table (Singapore National I/O table 2007), and aggregating the 136 sectors into 9 main critical infrastructure sectors and 1 sector that contains the remaining miscellaneous sectors. The new aggregated I/O model is then used to generate the I/O interdependency model that depicts the links between critical infrastructure sectors as shown in Fig. 1. Through the interfacing of physical infrastructure network and I/O interdependency model, analysing multiple critical infrastructure in I/O interdependency model becomes possible. The advantage would be that the use of physical infrastructure network would be more potent in modelling real world events which lead to cascading impact to other sectors represented within the I/O interdependency model. Thus, with the interfacing, the effect and extent of impact to other sectors can be established if an event that causes damage to an actual physical network occurs. The I/O interdependency model serves the purpose of describing physical linkage of the 10 (9+1) grouped sectors. The details on the model and aggregation of I/O table are found in Lin et al. (2015).

2.2 Combining of interdependencies and critical infrastructure network

The simplified network topology of two actual physical infrastructure networks are modelled and chosen for network analysis (Fig. 1). The networks correspond to the main linkages of the telecommunication network and electrical grid in Singapore respectively. The network comprises of two sets of elements: connection nodes and transmission links. The authors will consider a disruption to a node as node failure to simplify the study. Upon any node failure in the physical infrastructure network, the network performance for the physical infrastructure will be computed. A node failure causes all links that are connected to that node to be disabled and the node itself to be removed as well. Two network performance metrics will be used in this study. The two metrics are (1) efficiency of network (Crucitti et al. 2004, Wang et al. 2013) and (2) network average clustering coefficient (Mao et al. 2009, Brust & Rothkugel 2007, Holmgren 2006). There have not been any studies to determine which of the two metrics is more appropriate as a measure of the network performance, hence both metrics are used in this work.

The efficiency of a network is given by

$$\eta = \frac{1}{n(n-1)} \sum_{i,j \in V} \frac{1}{d_{ij}} \qquad (8)$$

where n represents the number of nodes in a network, and d_{ij} represents the shortest path length between nodes i and j.

Average clustering coefficient of a network (Kemper 2009) is defined as the average of the clustering coefficient of all individual nodes in the network, given by

$$\eta = \frac{1}{n}\sum_{i=1}^{n} \frac{\text{number of triangles connected to node } i}{\text{number of triplets centered on node } i} \quad (9)$$

where n represents the number of nodes in a network.

The disruption (e.g cyber attack) will be simulated and the percentage losses of network performance, e_{loss}, upon node failure in physical infrastructure network can be calculated as

$$e_{loss} = \left(1 - \frac{\eta_{new}}{\eta_{initial}}\right) \times 100\% \quad (10)$$

where $\eta_{initial}$ is the initial network performance and η_{new} is the new network performance. Both network performance must be calculated according to the selected network performance metric. If the performance of the network after a node failure is lower, the removed node is critical (e_{loss} will be positive).

When a node failure in a network happens, it will result in failure to other nodes of the same network. The failure of physical infrastructure is being captured within a fixed time frame (e.g one day) and the losses of the particular infrastructure due to node failure can be captured in the form of lower contribution to the economy as a whole in a day. We assume that for one day of failure of a physical infrastructure, it will incur losses in the form of:

$$w_{loss} = \frac{x_i}{365} e_{loss} \quad (11)$$

where w_{loss} represents the monetary loss for an infrastructure sector, x_i represents the output of the sector i in a year. The w_{loss} will be used as ΔF in the I/O model. The losses of network performance will be able to propagate across to the other industry sectors and the total damage to all industry sectors can then be evaluated. A complete exhaustive search for all cases of node failure can be performed to find the worst case scenario for a predefined number of failed nodes.

The network performance for multiple physical network will have to be calculated individually as they are physically not connected but are connected through I/O interdependency model. The e_{loss} and w_{loss} for each of the two infrastructure networks in this study will be calculated individually and used as ΔF in the I/O model to get the overall economic impact over to all industry sectors.

2.3 Assumption of input-output interdependency model

This model follows the original I/O model assumption. Due to the equilibrium assumption of the Leontief model, the economic losses are typically estimated on an annual basis. Hence, for smaller time resolutions, it is assumed that the losses are evenly distributed throughout the year (Anderson et al. 2007).

Each element in I/O table gives a constant, value that represents the contributions to one sector, say *j*, from any other sector, say *i*, which is proportional to the output of sector *j*. There are obvious examples where this assumption is valid. For example, if sector *i* is water production and sector *j* is agriculture production, then the value of water used by sector j would naturally increase linearly with the value of agriculture production. However, the linearity assumption may not be always valid. For example, sector *j* which is agriculture production may seek to improvise by using better technology through the use of water recycling or other technology that can reduce usage of water to increase their agriculture production.

3. Economic impact analysis and discussion

NetLogo is a multi-agent programming language and integrated modelling environment (Wilensky 1999) which is used for the network analysis in this work. In the Netlogo programme, any failure in the critical infrastructure network will disconnect all direct links to the failed node. This failure scenario is implemented under our framework to simulate the different combination of node failures and their impact on all critical infrastructure sectors linked in the I/O interdependency model. Fig. 2 shows the topology of the two physical infrastructure networks developed in Netlogo. The simulations include one-node failure, two-node failure and three-node failure. These failure scenarios attempt to simulate node failures in a particular time frame, for this case, the time frame is a one full day scenario. e_{loss} will be calculated based on the two different network performance metrics and thereafter, w_{loss} can be calculated, and will be used as ΔF for I/O model to assess the impact to all sectors, ΔX, due to the node failures. For clarity, node 0 to node 8 are the nodes in the telecommunication network while node 9 to node 20 are the nodes in the electrical grid.

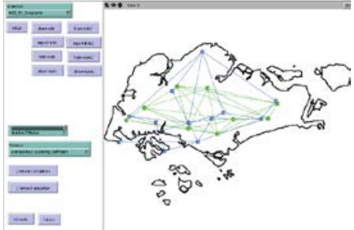

Fig. 2. Netlogo programme interface for modelling and simulation of node failure of two critical infrastructure networks

Tables 1, 2 and 3. show the results of the top 3 worst case scenarios for one-node, two-node and three-node, failures respectively. Different network performance metrics will result in different top 3 worst case scenarios due to the different nature of the metrics being used. $e_{loss\,(1)}$ and $e_{loss\,(2)}$ represent the percentage losses of network performance in the telecommunication network and electrical grid respectively.

Node 0 is found to be the most critical node in one-node failure based on the efficiency of network while node 14 is the most critical node in one-node failure based on the network average clustering coefficient metric.

From the two-node failure results in Table 2, the node pairs in the top 3 scenarios for efficiency of network scheme contain node 0 which is also the most critical node for one-node failure in Table 1. Based on network average clustering coefficient, nodes [14 17] and nodes [14 19] cause 100% losses in network performance in $e_{loss\,(2)}$. These two combinations of node failures cause the electrical grid to have clustering coefficient of zero for all nodes, and therefore zero network average clustering coefficient for the network.

Table 1 Results of the top 3 worst case scenarios for one-node failure

Network performance	Failure scenario	$e_{loss\,(1)}$	$e_{loss\,(2)}$	Overall impact, ΔX ($ M)
Efficiency of network	[0]	7.76	0.00	6.94
	[7]	2.17	0.00	1.94
	[8]	2.17	0.00	1.94
Network average clustering coefficient	[14]	0.00	65.3	22.0
	[16]	0.00	26.8	9.04
	[5]	9.64	0.00	8.62

Table 2 Results of the top 3 worst case scenarios for two-node failure

Network performance	Failure scenario	$e_{loss\,(1)}$	$e_{loss\,(2)}$	Overall impact, ΔX ($ M)
Efficiency of network	[0 1]	10.6	0.00	9.44
	[0 3]	10.6	0.00	9.44
	[0 7]	10.6	0.00	9.44
Network average clustering coefficient	[14 17]	0.00	100	33.7
	[14 19]	0.00	100	33.7
	[5 14]	9.64	65.3	30.6

Table 3 Results of the top 3 worst case scenarios for three-node failure

Network performance	Failure scenario	$e_{loss\,(1)}$	$e_{loss\,(2)}$	Overall impact, ΔX ($ M)
Efficiency of network	[0 1 3]	16.5	0.00	14.8
	[0 7 8]	16.5	0.00	14.8
	[1 3 5]	13.9	0.00	12.4
Network average clustering coefficient	[1 4 5]	58.0	0.00	51.8
	[3 4 5]	58.0	0.00	51.8
	[4 5 8]	28.0	65.3	47.0

From the three-node failure results in Table 3, nodes [0 1 3] and nodes [0 7 8] will cause the greatest overall impact based on the network efficiency metric while nodes [1 4 5] and nodes [3 4 5] will cause the greatest overall impact based on the average clustering coefficient. The node pairs [14 17] and [14 19] do not appear again in three node failures because other combinations of node failure will cause higher overall impact to the economy. From here, we can see that the node that causes the greatest damages in one-node or two-node failure does not always appear in the worst three-node failures. However, with the capability of the simulation, this framework will be significantly useful for researchers and stakeholders as they can run through and test out all scenarios before making judgement and decision.

4. Conclusion

This work is interesting because the combination of failures across nodes in different physical infrastructure networks may cause higher than expected impact than the same number of failed nodes within a single network. This is a good example of what may be called a perfect storm (Pate Cornell, 2012) whereby a combination of unfavorable circumstances happening simultaneously can cause unexpected and unimaginable damages. This simulation helps to provide insights to the fragility of the economic system to various forms of disruption The authors hope that with the sharing of information and complete data on critical infrastructures, black swan events like the 9/11 terrorist attack or some cyber attack targeting critical infrastructures can be mitigated, and stakeholders can be made aware of the potential risks and be prepared.

References

Anderson, C.W., Santos J.R. and Haimes Y.Y. (2007). A Risk-based Input–Output Methodology for Measuring the Effects of the August 2003 Northeast Blackout. *Economic Modeling For Disaster Impact Analysis 19(2)*, 183-204.

Brust, M. and Rothkugel,S. (2007). “Small-worlds: Strong clustering in wireless networks,” in First International Workshop on Localized Algorithms and Protocols for Wireless Sensor Networks, USA.

Crucitti, P., Latora, V. and Marchiori, M. (2004) "Model for cascading failures in complex networks," *Physical Review E,* vol. 69, p. 045104.

Kemper, Andreas (2009). Valuation of Network Effects in Software Markets: A Complex Networks Approach. Springer. p.142.

Kizhakkedath, A., Tai, K., Sim, M.S., Tiong, R.L.K. and Lin, J. (2013) An Agent-Based Modeling and Evolutionary Optimization Approach for Vulnerability Analysis of Critical Infrastructure Networks, *Communications in Computer and Information Science Vol.402*, Springer, pp.176-187

Krimgold F, B. J. (2006). Power Systems, Water, Transportation and Communications Lifeline Interdependencies for American Lifelines Alliance. Advanced Research Institute, Virginia Polytechnic Institute and State University.

Laefer, D. K. (2006). The Need for Baseline Data Characteristics for GIS-Based Disaster Management Systems. *J. Urban Plann. Dev., 132(3)*, pp. 115-119.

Lam, C.Y., Lin, J., Sim, M.S. and Tai, K. (2013) "Identifying Vulnerabilities in Critical Infrastructures by Network Analysis", *International Journal of Critical Infrastructures, Vol.9, No.3*, pp.190-210

Lin, J , Tai, K., Tiong, R.L.K. and Sim, M.S. (2015). A general framework for critical infrastructure interdependencies modeling using economic input-output model and network analysis. *Proceeding of the 2nd Asia-Pacific Conference on Complex Systems Design & Management, Singapore*. (In print)

Liu, Z., Ribeiro, R., & Warner, M. (2004). Comparing child care multipliers in the regional economy: Analysis from 50 states. Cornell University, Department of City and Regional Planning.

Mao, Z., Hong, L., Fei, Q., Ouyang, M. (2009). Interdependency Analysis of Infrastructures. *Proceedings of the 6th International Symposium on Neural Networks: Advances in Neural Networks*

Miller, R. E., & Blair, P. D. (2009). Input-Output Analysis: Foundations and Extensions. New York: Cambridge University Press.

Paté-Cornell, E. (2012), On “Black Swans” and “Perfect Storms”: Risk Analysis and Management When Statistics Are Not Enough. *Risk Analysis, 32*

Rinaldi, S.M.; Peerenboom, J.P.; Kelly, T.K. (2001). Identifying, understanding, and analyzing critical infrastructure interdependencies, in *Control Systems, IEEE , vol.21, no.6*, pp.11-25.

Santos, J. R. (2006), Inoperability input-output modeling of disruptions to interdependent economic systems. *Syst. Engin., 9*, pp. 20–34.

Tai, K., Kizhakkedath, A., Lin, J., Tiong, R.L.K. and Sim, M.S. (2013) Identifying Extreme Risks in Critical Infrastructure Interdependencies, *in ISNGI 2013*, Wollongong, Australia

The President’s Commission on Critical Infrastructure Protection. (1997). Critical Foudation: Protecting America's Infrastructure. USA.

Wang, S., Hong, L., Ouyang, M., Zhang, J., and Chen, X. (2013). Vulnerability analysis of interdependent infrastructure systems under edge attack strategies, *Safety science, vol. 51*, pp. 328-337.

Wilensky, U. (1999). NetLogo. Center for Connected Learning and Computer-Based Modeling, Northwestern University, Evanston, IL.

William Baumol & Thijs ten Raa, (2009). Wassily Leontief: In appreciation, The European Journal of the History of Economic Thought, T&F Journals, vol. 16(3), pp. 511-522.