NPS-OR-20-003

NAVAL POSTGRADUATE SCHOOL

MONTEREY, CALIFORNIA

OPERATIONAL RESILIENCE OF COMMAND AND CONTROL SYSTEMS TO MAINTAIN MULTILAYERED NETWORK FUNCTIONALITY IN RESPONSE TO LARGE-SCALE DISRUPTIVE EVENTS

by

David L. Alderson
Daniel Eisenberg
Alexander Ganin
Maksim Kitsak
Igor Linkov

December 2020

Approved for public release. Distribution is unlimited

Prepared for: Defense Threat Reduction Agency

THIS PAGE INTENTIONALLY LEFT BLANK

REPORT DOCUMENTATION PAGE

1. REPORT DATE *(DD–MM–YYYY)*	2. REPORT TYPE	3. DATES COVERED *(From — To)*
December 2020	Technical Report	17 November 2016 - 30 September 2020

4. TITLE AND SUBTITLE

OPERATIONAL RESILIENCE OF COMMAND AND CONTROL SYSTEMS TO MAINTAIN MULTILAYERED NETWORK FUNCTIONALITY IN RESPONSE TO LARGE-SCALE DISRUPTIVE EVENTS

5a. CONTRACT NUMBER

5b. GRANT NUMBER

DTRA-18681-M

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)

David L. Alderson, Daniel Eisenberg, Alexander Ganin, Maksim Kitsak, Igor Linkov

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Naval Postgraduate School
Monterey, CA 93943

8. PERFORMING ORGANIZATION REPORT NUMBER

NPS-OR-20-003

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

Defense Threat Reduction Agency

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release. Distribution is unlimited

13. SUPPLEMENTARY NOTES

The views expressed in this document are those of the authors and do not reflect the official policy or position of the Department of Defense or the U.S. Government.

14. ABSTRACT

This technical report summarizes the work conducted under Award DTRA-18681-M during the performance period 2016-2020. Specifically, this research project seeks to develop new theory, models and algorithms for optimal design of interdependent networks, including command and control (C2) systems, with the objective of making them resilient to both targeted (intentional) and random (natural) attacks. We study both large-scale WMD attacks and small-scale attacks on the most important nodes in the system. "System" refers here to a collection of interdependent networks such as the physical network comprised of sensors, satellite intelligence installations, and transmission and communications stations; information network, comprised of information processing and routing devices (e.g., servers) and software components (databases, encryption and transmission software); and social and cognitive domain comprised of the personnel involved in operation and control of the underlying information and physical domains. We study hierarchical and distributed organization of these domains and apply the resilience quantification methodology in order to find the best configuration of the system as a whole.

15. SUBJECT TERMS

command and control, resilience, network, disruption

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			David L. Alderson
Unclassified	Unclassified	Unclassified	UU	241	19b. TELEPHONE NUMBER *(include area code)* 831-656-1814

THIS PAGE INTENTIONALLY LEFT BLANK

NAVAL POSTGRADUATE SCHOOL
Monterey, California 93943-5000

Ann E. Rondeau   Robert F. Dell
President   Acting Provost

The report entitled "Operational Resilience of Command and Control Systems to Maintain Multilayered Network Functionality in Response to Large-Scale Disruptive Events" was prepared for and funded by the Defense Threat Reduction Agency.

Further distribution of all or part of this report is authorized.

This report was prepared by:

_____   _____

David L. Alderson   Daniel Eisenberg

_____   _____

Alexander Ganin   Maksim Kitsak

_____

Igor Linkov

Reviewed by: **Released by:**

_____   _____

W. Matthew Carlyle, Chairman   Jeffrey D. Paduan
Department of Operations Research   Dean of Research

THIS PAGE INTENTIONALLY LEFT BLANK

ABSTRACT

This technical report summarizes the work conducted under Award DTRA-18681-M during the performance period 2016-2020. Specifically, this research project seeks to develop new theory, models and algorithms for optimal design of interdependent networks, including command and control (C2) systems, with the objective of making them resilient to both targeted (intentional) and random (natural) attacks. We study both large-scale WMD attacks and small-scale attacks on the most important nodes in the system. "System" refers here to a collection of interdependent networks such as the physical network comprised of sensors, satellite intelligence installations, and transmission and communications stations; information network, comprised of information processing and routing devices (e.g., servers) and software components (databases, encryption and transmission software); and social and cognitive domain comprised of the personnel involved in operation and control of the underlying information and physical domains. We study hierarchical and distributed organization of these domains and apply the resilience quantification methodology in order to find the best configuration of the system as a whole.

THIS PAGE INTENTIONALLY LEFT BLANK

Table of Contents

THIS PAGE INTENTIONALLY LEFT BLANK

List of Figures

THIS PAGE INTENTIONALLY LEFT BLANK

List of Tables

THIS PAGE INTENTIONALLY LEFT BLANK

List of Acronyms and Abbreviations

AQI   Al-Qaeda in Iraq

AS   Autonomous System

AUPR   Area Under the Precision-Recall characteristic

BGP   Border Gateway Protocol

C2   Command and Control

CACC   Cooperative Adaptive Cruise Control

CCRP   Command and Control Research Program

CF   Critical Functionality

CFMCC   Coalition Force Maritime Component Commander

CGE   Computable General Equilibrium

CLM   Crucitti-Latora-Marchiori model

GRHG   Complementary-based Random Hyperbolic Graph

CMRCMPSP   Combinatorial Multi-mode Resource Constrained Multi-Project Scheduling Problem

CPM   Critical Path Method

DAG   Directed Acyclic Graph

DHS   Department of Homeland Security

DRAASTK   Dynamic Resourcing Allocation Analysis Simulation Tool Kit

DTRA   Defense Threat Reduction Agency

DoD   Department of Defense

ER	Erdos-Renyi
GC	Giant Component
GCC	Giant Connected Component
GDP	Gross Domestic Product
GLEAM	Global Epidemic and Mobility Model
GPS	Global Positioning System
GW	Gigawatts
HHTS	Household Travel Survey
ICS-CERT	Industrial Control Systems Cyber Emergency Response Team
ICT	Information and Communication Technology
ITS	Intelligent Transportation System
JFMCC	Joint Force Maritime Commander
JSOTF	Joint Special Operations Task Force
JTF	Joint Task Force
KEPCO	KoreanElectric Power Corporation
KPG	South Korean Power Grid
KPX	Korea Power Exchange
LA	Los Angeles
Los	Level of Service
MET	Mission Essential Task
MHQ	Maritime Headquarters
MOC	Maritime Operations Center

MPTE	Manpower, Personnel, Training, and Education
MRCMPSP	Multi-mode Resource Constrained Multi-Project Scheduling Problem
MRCPSP	Multi-mode Resource Constrained Project Scheduling Problem
NAS	National Academy of Sciences
NATO	North Atlantic Treaty Organization
NAVFOR	Naval Force
NCW	Network Centric Warfare
NEC	Network Enabled Capability
NHTS	National Household Travel Survey
NPS	Naval Postgraduate School
PERT	Project Evaluation and Review Technique
PPI	protein-protein interactions
REMI	Regional Economic Models Inc.
RCMPSP	Resource Constrained Multi-Project Scheduling Problem
RCPSP	Resource Constrained Project Scheduling Problem
RHG	Random Hyperbolic Graph
SCADA	Sensory Control and Data Acquisition
SEDAC	Socioeconomic Data and Application Center
SEIR	Susceptible Exposed Infected Recovered
SF	Scale-Free
SF	San Francisco
SME	Subject Matter Expert

USN	U.S. Navy
USG	United States government
VHT	Vehicle Hours Traveled
VMT	Vehicle Miles Traveled
VOR	Value of Reliability
VOT	Value of Time
WMD	Weapons of Mass Destruction

Executive Summary

This research project seeks to develop new theory, models and algorithms for optimal design of interdependent networks, including command and control (C2) systems, with the objective of making them resilient to both targeted (intentional) and random (natural) attacks. We study both large-scale WMD attacks and small-scale attacks on the most important nodes in the system. "System" refers here to a collection of interdependent networks such as the physical network comprised of sensors, satellite intelligence installations, and transmission and communications stations; information network, comprised of information processing and routing devices (e.g. servers) and software components (databases, encryption and transmission software); and social and cognitive domain comprised of the personnel involved in operation and control of the underlying information and physical domains. We study hierarchical and distributed organization of these domains and apply the resilience quantification methodology in order to find the best configuration of the system as a whole.

We have focused on four primary research thrusts:

(a) A first goal is to develop and test a universal resilience quantification method, applicable to any complex system based on the underlying network (possible heterogeneous and comprised of many layers and domains).

(b) A second goal is to develop appropriate abstract small-scale models for the command and control networks. In particular, during this stage, we will investigate the relevant properties of the nodes and links in C2 networks, and most importantly their underlying dynamics.

(c) A third goal is to develop large-scale abstract models of the command and control network and study various types of the network organization (e.g. hierarchically directed links between the nodes or distributed undirected organization of the nodes interconnectedness) and characteristics of the layers (e.g. node degree distribution, centrality, capacity and other relevant properties). We also shall study the network robustness depending on the type of damage it encounters (e.g. random or targeted).

(d) A fourth goal is focused on the resilience of the real military and civilian C2 systems. We apply either small-scale or large-scale models as appropriate for the data we obtain for real systems.

THIS PAGE INTENTIONALLY LEFT BLANK

Acknowledgments

THIS PAGE INTENTIONALLY LEFT BLANK

1 What is Command and Control?

This introduction is adapted from Eisenberg et al. (2018a).

Command and control (C2) systems are important, pervasive systems in modern society. C2 is understood as the process taken by teams and organizations in order to complete a shared goal (Stanton et al. 2008; Walker et al. 2008), and C2 systems are sociotechnical collections of human interactions, social norms, and built technologies that enable this process (Walker et al. 2008, 2012). Historical understandings of C2 refer to the strict, bureaucratic, and hierarchical structures of military organizations where individual roles are clear and mission expectations are non-negotiable (Alberts and Hayes 2003, 2006). Contemporary notions of C2, in contrast, refer to any organizational structure that connects people to perform shared goals, and emphasizes processes that enable decision-making and information-sharing rather than the social context in which they occur (Alberts and Hayes 2003; Linkov et al. 2013a). This broad view on C2 includes military hierarchies alongside public utilities and civil infrastructure managers (Petrenj et al. 2013), loosely connected non-governmental organizations during emergency response (Alberts et al. 2010), and even virtual organizations formed on the internet that lack rules for establishing leadership, allocating workloads, or monitoring business activities (Grabowski and Roberts 2015). Definitions of a leader, the act of sharing intent, and completing a task or mission depend upon the organizational structure in which C2 is being performed. C2 organizations are often geographically separated in such a way that requires information and communication technology (ICT) to share skills and knowledge, yet still have long-term interests and shared goals that maintain interaction among participants (Grabowski and Roberts 2015). Linking the historical and contemporary, essentially any collaborative group of people using ICT to achieve a shared goal can be characterized by a form of C2.

The purpose of C2 research is to be *descriptive* for which social and technological relationships will succeed in performing shared goals and *prescriptive* for how to design better C2 systems. Descriptive research reveals underlying factors that enable successful completion of missions, often measured as the speed an organization can complete a task, the diversity of different tasks an organization can complete, the amount of shared information among organizational members, or a combination of all three (Alberts 2011). Some descriptive research is also explanatory of sociotechnical phenomena and uncovers drivers and causal forces that dictate why some human or technological arrangements are more effective than

1

others (Alderson et al. 2006; Willinger et al. 2002). Prescriptive research then builds on these results to rearrange existing social relationships and/or introduce disruptive technologies to improve existing C2 processes. Together, C2 research endeavors have broad impact on social systems of decision hierarchy, interpersonal interactions, knowledge sharing, training, and skills and technological systems that enable data collection and resource use to act upon changing situation context.

Despite the potential for C2 research to improve the design of real-world sociotechnical systems, the broad literature remains disorganized and isolated in publication. Only few seminal works offer an integrated understanding of C2 theory (e.g., Alberts and Hayes 2003; Alberts et al. 2010), and none organize applied studies to show how C2 is understood and evaluated across contexts. This allows almost any new study involving a shared goal, ICT systems, software development, or a social network to claim it advances C2 theory. This lack of organization produces a disconnected body of knowledge and creates confusion for newcomers as some C2 related topics like 'social networks' encompass thousands (if not millions) of research records. This also results in a lack of standard practices across experts that limits advances in C2. One consequence is that the majority of research remains descriptive rather than prescriptive in experimental design, modeling, and analysis. New studies slow the development of prescriptive studies by relying on novel methods for analysis rather than common frameworks. The few studies that do prescribe ways to improve system-level C2 processes are completed in a piecemeal fashion that is not re-integrated into a comprehensive theory. C2 literature is marred by a disconnected landscape of research that inhibits the possibility of advancing research to practice.

The purpose of this introductory section is to review and organize the C2 literature in a manner that makes it more understandable and more usable. C2 literature needs integration of existing knowledge to guide the design of systems that successfully adapt to changing circumstances. This work has two specific goals: (1) organize and integrate the disconnected C2 body of knowledge, and (2) offer guidance to support more informed descriptive and prescriptive research. We achieve these goals via a comprehensive literature review that organizes research with well-established C2 theory and network evaluation methods. The paper concludes with a network science perspective on C2 literature that presents design and modeling constraints considered in past C2 research and five considerations for future C2 research.

1.1 Network Centric Warfare as Guiding Theory

The military doctrine of Network Centric Warfare (NCW; c.f. Network Enabled Capability (NEC) (Alberts et al. 2010)) offers a basis for organizing advances in C2 research. C2 research is guided by few overarching theories that break down the structure and function of sociotechnical systems for analysis and design. NCW is one of the most widely cited C2 theories that emphasizes the relationship between successful C2 processes and networked system structure. NCW doctrine was developed by the US Department of Defense (DoD) Command and Control Research Program (CCRP)—see Alberts and Hayes (2003); Cebrowski and Garstka (1998); Dekker (2008)—for harnessing rapid, ubiquitous, "Information Age" technologies to improve C2 activities. NCW is centered on 21st century technological advancements like the internet, wireless networking, sensors, and satellites that brought a shift in society by making "information" a strategic asset (Alberts 2011). Information in this context refers to high volume, velocity, and variety data (Laney 2001) that enables automated and distributed systems to work together and now underlies most aspects of daily life. While society writ-large quickly adopted Information Age technologies, this transition was difficult in military organizations that have long-standing social hierarchies and strict information assurance and security requirements. NCW started in the late 1990s as a critique on the lack of adoption of Information Age technologies in military C2 processes alongside a shift in perspective from treating military units as independent platforms to networked systems (Fewell and Hazen 2003). The resulting doctrine developed by the CCRP provides a comprehensive theoretical overview of C2 systems, processes, and needs, while also identifying characteristics of successful C2 systems. Organizing C2 research with respect to NCW doctrine provides a useful way to understand the current state of knowledge and offers a basis for making recommendations for future work.

In particular, NCW doctrine establishes that successful C2 systems are agile to adapt to changing mission needs, such that reviewing how current research advances NCW theory may reveal how to design agile systems. Agility is defined in NCW literature as, "the ability to successfully effect, cope with, and/or exploit changes in circumstances" (Alberts 2011), which corresponds to similar definitions found in manufacturing (Sherehiy et al. 2007), management (Santos Bernardes and Hanna 2009), and infrastructure (Chester and Allenby 2019) contexts. Agility is comprised of both passive and active components such as responsiveness, flexibility, and resilience among others (Alberts 2011) that influence how

exploiting a situation may occur. Agility is studied in NCW via three dimensions that define the C2 approach space: (1) allocation of decision rights to the collective, (2) patterns of interaction, and (3) distribution of information (Alberts et al. 2010). NCW doctrine asserts that these three organizational dimensions can predict a system's agility, where C2 systems range from "Conflicted C2" with the least agility due to constrained decision rights, patterns of interaction, and distribution of information to "Edge C2" with the highest agility due to unconstrained decision rights, patterns of interaction, and distribution of information. NCW prescriptive recommendations focus on shifting existing systems from Conflicted C2 to Edge C2 by adopting novel policies and technologies that decentralize decision rights, increase the frequency of interactions, and increase the richness of information sharing. An important output of this review is identifying studies that describe and prescribe ways to achieve agile systems and Edge C2.

1.2 Multilayer Network Science as a Guiding Framework

Advancing NCW doctrine requires finding the right balance between generic systems analysis methods that apply to a breadth of C2 sub-systems and operational specificity to capture C2 context and dynamic processes that dictate agility. Considering this need, we also review the C2 literature to identify network models and analysis methods that are generic enough to capture a breadth of contexts across descriptive research, but specific enough to prescribe context-specific network designs.

Network science is a popular way to model C2 sub-systems and offers a basis for comparing evaluation methods across C2 research. A network is a system model comprised of nodes representing constituent parts and links representing nodal interactions. Network science methods are useful for studying the structure and function of C2 sub-systems by ranking nodes and links, where social and technological networks use the same methods to represent different constructs and dependencies (Eisenberg et al. 2017). For example, a social network of people linked by who-knows-whom can quickly reveal critical actors and relationships that support information flow and decision-making. This is evaluated at the component scale by ranking individuals via simple measures like the number of links a person has (Altman et al. 2017) and at the system scale by classifying network topology (Newman 2010). Studies also show relationships between system structure with network stability (Kitsak et al. 2018) and C2 processes like agility and resilience (Ganin et al. 2016). Thus, network science

models and analysis methods provide a consistent basis for comparing technical advances in the C2 literature, even when considering disparate models and application contests.

Recent advances in multilayer network science provide additional justification for using networks to compare evaluation methods found in the C2 literature. While research on a single C2 sub-system like ICT is important, advancing C2 theory requires new knowledge at the intersection of human and technological sub-systems. Growth in network science literature has brought with it the extension of methods for individual networks to the integrated analysis of multiple networks together. As network models provide a basis for comparing evaluation methods among disconnected sub-systems, multilayer network analysis (Kivelä et al. 2014; Boccaletti et al. 2014a; De Domenico et al. 2014) provides a consistent way to compare interactions among interdependent sub-systems. Examples of multilayer network studies that inform C2 theory include coupled cyber-physical networks (Buldyrev et al. 2010), sociotechnical networks (**?**), and cyber-physical-social networks (Barker et al. 2017). The breadth of these studies emphasizes that multilayer network science provides flexible way to organize literature across disparate application contexts. Moreover, comparing C2 literature through the lens of network models and analysis can set a baseline for the current state of C2 knowledge within and across sub-systems.

The review in Eisenberg et al. (2018a) links NCW doctrine and multilayer network science to identify which studies advance C2 agility. While the primary goal is to organize the field, a secondary goal of this work is to identify how well network science serves C2 research. The generic nature of network science methods is both advantageous and problematic for C2 theory, because network analysis of diverse human and technological systems is possible even when detailed understanding of either is lacking. Research that is overly specific to a single application context is not necessarily helpful for advancing C2 research. However, multilayer networks that reveal interactions across sub-systems often lack consideration of the physical principles or operational specifications of C2 systems making them inappropriate for real-world use. In more extreme cases, network science studies can even be counterproductive to C2 research when narrow applications ignore the breadth of sociotechnical interactions or results produced with unrealistic models lack operational detail (Doyle et al. 2005; Alderson and Doyle 2010; Alderson et al. 2013). Identifying how network science is applied in C2 literature can help overcome these issues to further both NCW doctrine and network evaluation methods.

1.3   C2 as Four Interacting Sub-systems

We draw upon NCW doctrine to define the social and technological sub-systems that comprise C2 networks. NCW doctrine defines four domains of warfare that dictate C2 agility, where each domain can be represented by network models. Alberts and Hayes (2003) and Atkinson and Moffat (2005) provide the first descriptions of C2 domains by linking complex systems science to the structure and implementation of military policies. The authors describe C2 as an interacting set of layers, including the physical world, the information exchange that occurs over physical systems, the interpretation of data by people, and a shared understanding generated by this process. Thus, a simple task like sending an email involves networks of technologies that the individual has access to (e.g., computer), software and digital services required to deliver an email, values and beliefs of the individuals writing and reading the email, and actions taken due to this exchange. Based on this definition, C2 systems are comprised of at least four interacting sub-systems representing physical, information, social, and cognitive systems (see Figure 1). These descriptions were formalized into four broad domains by the NATO research task group SAS-065 (Alberts et al. 2010) as:

- *Physical Domain:* sensors, facilities, and equipment;
- *Information Domain:* creation, manipulation, and storage of data;
- *Social Domain:* human organization and interactions; and,
- *Cognitive Domain:* mental models, preconceptions, biases, and values.
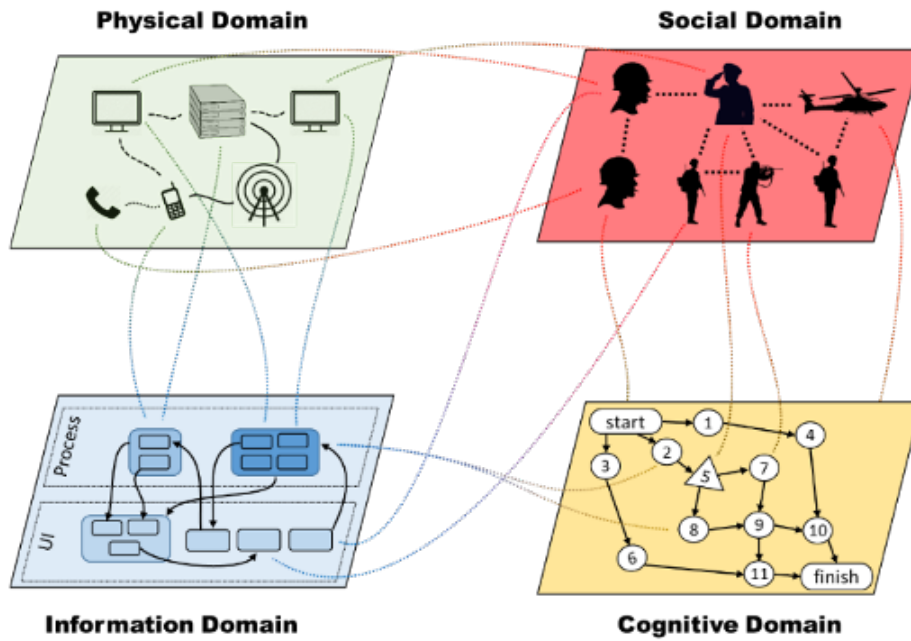
Figure 1. Multilayer network super structure of NCW domains and literature. NCW domains are comprised of distinct sub-systems that interact to form a four-layer network superstructure. C2 literature informs the drivers, models, and measures that describe and prescribe inter- and intra-network structure and function.

Each domain layer in Figure 1 embeds context-specific information that dictates the structure and function of underlying systems. Table 1 lists the model primitives commonly associated with each layer.

Table 1. Network primitives considered across the C2 literature

C2 Domain	Nodes	Links	Dynamics
Physical Domain: *Information & Communication Infrastructure*	- Sensors - Routers - Access Points	- Cables / Fiber - Wireless Signals - Containment	- Network Traffic
Information Domain: *Digital Services*	- Functions - Packages	- Data Streams - Dependencies	- Service Provision - Service Use
Social Domain: *Organizational Structure*	- Individuals - Teams - Organizations	- Interactions - Knowledge	- Shared Awareness - Communication
Cognitive Domain: *Mission & Goals*	- Activities - Decisions	- Preconditions - Iterations	- Needs Assignment

Columns present the different constructs that comprise C2 network models, where the combination of nodes, links, and dynamics across rows defines a network layer. Advances in C2 research may reveal new primitives, structures, and multilayer analysis methods that combine network models together into a single study.

Table 2 lists the types of technical information experts consider within reviewed articles when modelling physical, information, social, or cognitive systems. For example, physical domain includes connections among ICT hardware networks, the information domain by software architectures, service federation, calls, and pointers, the social domain by teams and organizational structure and the policies and beliefs that are embedded within them, and the cognitive domain by task networks that show the dependencies between actions in a logical sequence.

We also use complex systems theory that defines the architecture of each NCW domain to specify how to analyze C2 networks. Current C2 theory falls short of defining network primitives used to construct a C2 network model (i.e., nodes and links). Instead, we draw upon complex systems theory of system architecture to establish the constraints that dictate model structure and function. In particular, Alderson and Doyle (2010) define four design constraints that capture the breadth of system architectures found in physical, cyber, social, and cognitive domains, including:

- *Component constraints:* physical laws and requirements that dictate the capability of network nodes;
- *Protocols:* rules for the configuration and interaction of system components;
- *System-level constraints:* higher-level functional purpose of a single network layer including objectives and design criteria the system is meant to serve, e.g., maximizing radar signal; and,
- *Emergent constraints:* the laws that dictate physical limitations of real systems often expressed as needs and interactions across systems.

Table 2 lists common component constraints, protocols, and system-level constraints for each of the layers.

Table 2. Breadth of C2 domain architectures and sub-system considerations

C2 Domain	Component-level Constraints	Protocols	System-level Constraints
Physical Domain: *Information & Communication Infrastructure*	- Manufacturer - Energy Use - Size - Security - Reliability - Cost	- Industry Standards - Routing (e.g., TCP / IP) - Security (e.g., key generation) - Data Federation	- Latency - Access - Geographic Location
Information Domain: *Digital Services*	- Programming Language - Latency - Accuracy - Security - Reliability - Cost	- Software Architecture - Access (e.g., User Interface) - Discovery - Service Federation	- Availability - Security - Storage - Resource Budget - Documentation
Social Domain: *Organizational Structure*	- Roles & Duties - Psychological Traits - Beliefs & Practices - Demographics	- Standards - Administration - Policies - Partnerships	- Laws & Regulations - Politics - Geographic Location - Cultural Norms
Cognitive Domain: *Mission & Goals*	- Manpower - Funds - Time - Energy - Skills Required - Safety	- Execution Order - Control Logic	- Shared Objectives - Commander Intent - Stakeholder Needs - Budget

Advances in C2 research depend on the analysis of component, protocol, and system-level constraints within NCW domains with respect to past studies on similar and interdependent systems.

Table 3 additionally describes some of the emergent constraints that arise from the layered architecture in C2 systems.

Table 3. Emergent constraints across multilayer C2 systems

Needs for... / Constraints for ...	Software & Digital Services	Social & Organizational Structure	Mission & Goals
Information & Communication Infrastructure	Resource Use (processing, storage, connectivity, etc.)	Communication Traffic	Communication Activities
Software & Digital Services		Service Access	Human Factors & Capabilities
Social & Organizational Structure			Skills & Knowledge

System architectures within a single domain influence the dynamics of other domains, such that emergent, cross-layer may constraints dictate C2 agility. Advances in C2 research could benefit from an improved understanding of how emergent constraints influence the allocation of decision rights, patterns of interaction, and distribution of information.

The dimensions that influence C2 agility identified by NCW doctrine (i.e., allocation of decision rights, patterns of interaction, and distribution of information) are a result of the structure and function of integrated C2 systems. Ultimately, decisions on which constraints, primitives, and measures to use in representing and analyzing C2 network systems dictate the types of results one obtains. Together, the review in Eisenberg et al. (2018a) organizes the research literature to identify the state-of-the-art with respect to NCW doctrine and complex systems science.

2 Resilience Quantification

This project considered several techniques for quantifying resilience in the context of diverse applications. The material in this section is adapted from Ganin et al. (2016), Massaro et al. (2018), Ganin et al. (2017), Kurth et al. (2020), and Ganin et al. (2019).

2.1 Operational resilience: concepts, design and analysis

Building resilience into today's complex infrastructures is critical to the daily functioning of society and its ability to withstand and recover from natural disasters, epidemics, and cyber-threats. This study proposes quantitative measures that implement the definition of engineering resilience advanced by the National Academy of Sciences. The approach is applicable across physical, information, and social domains. It evaluates the critical functionality, defined as a performance function of time set by the stakeholders. Critical functionality is a source of valuable information, such as the integrated system resilience over a time interval, and its robustness. This section demonstrates the formulation on two classes of models: 1) multi-level directed acyclic graphs, and 2) interdependent coupled networks. For both models synthetic case studies are used to explore trends. For the first class, the approach is also applied to the Linux operating system. Results indicate that desired resilience and robustness levels are achievable by trading off different design parameters, such as redundancy, node recovery time, and backup supply available. The nonlinear relationship between network parameters and resilience levels confirms the utility of the proposed approach, which is of benefit to analysts and designers of complex systems and networks.

2.1.1 Introduction

The daily functioning of modern society is necessarily challenging, and traditional risk-based approaches to managing critical infrastructure are often criticized for their inability to address widely unknown and uncertain threats (Linkov et al. 2014c; Vespignani 2010; Forum 2015). Risk-based approaches require developing threat scenarios, evaluating system vulnerabilities and quantifying consequences associated with specific failures of system components. In the case of an unknown threat space, developing realistic scenarios proves to be an additional challenge. Moreover, it may be difficult to justify investing in hardening system components based on hypothetical and uncertain threats (Park et al. 2013a). Weak-

nesses and the potentially misleading nature of risk quantification approaches, for example in cyber systems, have been pointed out by a number of researchers (Jansen 2009; Bartol et al. 2009).

Building resilience into infrastructure networks (Holling 1973) has been proposed as the key to protecting against the deleterious effect of system disruption due to natural disasters (Cimellaro et al. 2010; Adger 2005) as well as infrastructure and engineering systems' failures (Ouyang et al. 2012; Kahan et al. 2009; Como et al. 2013; Vugrin et al. 2010). Not surprisingly, numerous interpretations of resilience have sprouted, testifying to the richness of the concept but also presenting challenges for its measurement and application1. Yet recent publications and guidance documents (U.S. Department of Homeland Security 2009; The White House 2009) coalesce around the definition of resilience provided by the National Academy of Sciences (NAS) (Press 2012): Resilience of a system is its ability "to plan and prepare for, absorb, respond to, and recover from disasters and adapt to new conditions". An important feature of resilience captured in this definition is the temporal dimension: the ability to recover and retain critical system functionality in response to a wide range of threats, both known and unknown. The assessment of resilience should therefore identify the critical functionality of a system and evaluate the temporal profile of system recovery in response to adverse events. Resilience management should comparatively evaluate cross-domain alternatives designed to enhance the system's ability to (i) plan for adverse events, (ii) absorb stress, (iii) recover and (iv) predict and prepare for future stressors in order to adapt to their potential threats.

Even though definitions of resilience as a system property are commonly reported in the literature (see details in Supplement S1, table S1.1), resilience assessment has been implemented in structured but largely qualitative or semi-quantitative ways (Barrett and Constas 2014). Insightful studies quantify resilience with metrics associated with different domains and subsequently integrate them into a risk-based evaluation or resilience index. For example, Bruneau et al. (2003) identified four dimensions of seismic community resilience: technical, organizational, social, and economic. Measures of resilience – robustness, rapidity, resourcefulness, and redundancy – were then aggregated in order to minimize a function of the probability of system failures, the consequences arising from such failures, and recovery time (Linkov et al. 2013a). Another qualitative but quantifiable approach sets forth a taxonomy for metrics that accommodates both change and interaction among

physical, informational, and human domains (Linkov et al. 2013c). The approach applies the taxonomy to cyber threats, energy systems and coastal infrastructure (Adger 2005; Linkov et al. 2013c; Carvalho et al. 2014). Such work provides insight and guidance for developing quantitative resilience measures that correspond to the qualitative identification of systemic issues and gaps. Unfortunately, it provides only limited insight into the management and control of the interconnected networks that constitute the entire system. Simultaneously, the field of network science has focused on the challenge of understanding the structure, dynamics and vulnerability of multi-layer systems across multilayer networks (Havlin et al. 2014; De Domenico et al. 2015; Brummitt et al. 2012; Massaro and Bagnoli 2014; Boccaletti et al. 2014c).

This section proposes a methodology for quantifying a system's resilience that captures the very concept of engineering resilience advanced by the NAS (Holling 1996; Pimm 1984) stated above. We make use of the critical functionality (CF) (which has been referred to before as functionality function (Cimellaro et al. 2010), performance (Ouyang and Wang 2015; Ouyang and Due nas-Osorio 2012), quality (Bruneau et al. 2003; Reed et al. 2009; Bocchini et al. 2014)), defined as a metric of system performance set by the stakeholders, to derive an integrated measure of resilience. One example for CF, among many possible ones, is the percentage of nodes that are functioning. Another is the ratio of a network's actual flow to its maximum capacity.

We note that, in addition to resilience, CF is rich in valuable information, and can be the source of many quantitative performance metrics, such as the robustness, which we also briefly discuss. For application domains, our focus in this paper are the following two classes of models: i) multi-level directed acyclic graphs (DAG) (Thulasiraman 1992), and ii) interdependent coupled networks (Buldyrev et al. 2010). While the second class of network is the subject of intense interest (D'Agostino and Scala 2014), the first class of networks, often overlooked by analysts, is of interest in many fields, from biology to computer science (Shapiro 2007; Yu and Gerstein 2006; Corominas-Murtra et al. 2013). As an application of this class, we also approximate the Linux, specifically Ubuntu 12.04, code system, as a DAG and estimate its resilience. Because obtaining analytical results is an intractable task, even for homogeneous networks, our approach is simulation based. We do, however, obtain analytical results in the Methods section for a simple yet illuminating special case of the first model, where only nodes without redundant active supply links may

13

be unable to supply service. This case sheds light on the relationship between redundancy and system resilience.

2.1.2 Resilience: an analytical definition

A network is modeled as a graph $G(N, L)$ with a set of nodes N connected by links L. Before considering network models, the proposed concept of resilience for complex networks is described generally. The specification of N and L includes characteristics relevant to resilience, such as capacity, location, and weight of each node and link. Let C be the set of temporal decision rules and strategies to be developed in order to improve the resilience of the system during its operation. From a computational viewpoint, the parameters and algorithms defined by C depend on the particular model being implemented.

Ultimately, the system must maintain its critical functionality K at each time step t, where K maps its states or parameters to a real value between 0 and 1. This mapping may, for instance, be linear

$$K(t; N, L, C) = \frac{\sum_{i \in (N,L)} w_i(t; C) \pi_i(t; C)}{\sum_{i \in (N,L)} w_i(t; C)} \tag{1}$$

where (N, L) is the set of all nodes and links, $w_i(t; C) \in [0, 1]$ is a measure of the relative importance of node or link i at time t, and $\pi_i(t; C) \in [0, 1]$ is the degree to which a node is still active in the presence of an attack. An alternate interpretation defines $\pi_i(t; C)$ as the probability that node or link i is fully functional. More complex, nonlinear and detailed definitions of critical functionality mappings are also possible. Finally, we introduce the class of adverse events (or potential attacks on targeted nodes) E. For instance, in the case of a random attack on two nodes, E is the set of all attacks on all possible node pairs.

Resilience, denoted by R, is a composite function of the network topological properties and their temporal evolution parameters defined for a certain critical functionality and a class of adverse events E:

$$R = f(N, L, C, E) \tag{2}$$

Note that not all targeted nodes are necessarily afflicted. For a non-afflicted attacked node, we thus have $\pi_i = 1$ over the entire time interval of interest. We evaluate R over a certain time interval $[0, T_C]$ where T_C is the control time (Majdandzic et al. 2013) which can be set a priori, for instance, by stakeholders or estimated as the mean time between adverse

events. In continuous time, we define R as

$$R \equiv R(K, E, [0, T_C]) = \frac{\frac{1}{|E|} \sum_E \int_{t=0}^{T_C} K(t; N, L, C)}{\int_{t=0}^{T_C} K^{nominal}(t; N, L, C)} \tag{3}$$

where $|E|$ is the cardinality of set E, and $K^{nominal}(t)$ is the critical functionality of the system in the case where no external events occur (Figure 2). Eq. 3 allows evaluation of the normalized dynamical performance of the system before (plan/prepare), during (absorption), and after an attack (recovery and adaptation); it intends to capture the definition advanced by the NAS given in the introduction. For computational convenience, the above equation is given in discrete time by

$$R \equiv R(K, E, [0, T_C]) = \frac{\frac{1}{|E|} \sum_E \sum_{t=0}^{T_C} K(t; N, L, C)}{\sum_{t=0}^{T_C} K^{nominal}(t; N, L, C)} \tag{4}$$

In most cases, we normalize to $K^{nominal}(t) = 1$. Consequently, a normalized measure of resilience may be given by

$$R = \frac{1}{T_C} \frac{1}{|E|} \sum_E \sum_{t=0}^{T_C} K(t; N, L, C) \tag{5}$$

We note that Eq. 1 embraces a large class of performance measures found in the literature. For instance, in addition to Eq. 5, we can also consider the measure

$$M = min_{t \in [0, T_C]} K(t; N, L, C) \tag{6}$$

The above measure is referred to as the Robustness (Cimellaro et al. 2010; Callaway et al. 2000). Alternatively one may define ($1 \check{} M$) as the Risk (Linkov et al. 2014c).

Due to the very complex nature of networked systems and the large number of variables defining their states, it is not possible to consider all events in the set E and obtain a closed-form expression for R, even if all design parameters are made homogeneous across nodes, links, and time. We therefore rely on a simulation based approach. Each simulation represents a possible scenario of the networked system's evolution. For each simulation, we calculate the average value of the critical functionality $K(t, N, L, C)$ at every time step
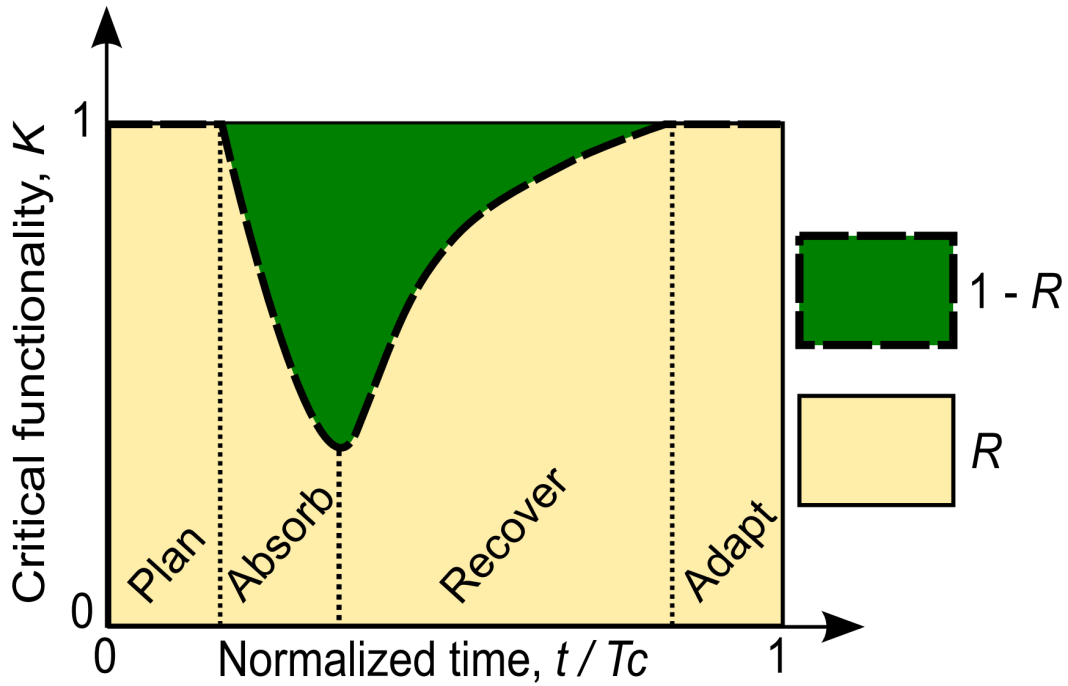
Figure 2. Resilience and critical functionality concepts as advanced by the NAS. The system's resilience is evaluated as the integral of the critical functionality's (K) dependency on time.

(Eq. 1), and from there, the resilience (Eq. 4 or Eq. 5) over the interval of interest. The approach proposed builds upon and extends the works of others (Cimellaro et al. 2010; Bruneau et al. 2003; Ouyang and Wang 2015; Ouyang and Due nas-Osorio 2012; Bocchini et al. 2014). The main issue encountered when dealing with the estimation of resilience based on the simulation of the system performance curves is that those curves in the general case vary depending on the adverse events modeled. The current approach to resolve this issue is to extend the techniques of probabilistic risk analysis to resilience analysis. This extension provides the weighted average performance curve with weights representing the probabilities of the adverse events.

By contrast, in our approach, we would like to argue that the resilience of the system should not be tied to the probabilities of the adverse events to occur. Again, according to the NAS resilience is the ability to plan, absorb, recover, and adapt. Inspired by this definition, we instead simulate the damage to the system from a certain adverse event (regardless of the probability for that event to occur) and define resilience for that particular damage. For

16

simplicity of decision making, however, we suggest considering a certain class of adverse events. For instance, in a networked system we might define one class of adverse events as a case when functionality of 4 to 5 nodes is reduced by 40-50% (instead of defining a particular adverse event reducing the functionality of specific 4 nodes by 50%).

We illustrate the approach with two simple models: multi-level DAGs and interdependent coupled networks. We assume that homogeneous nodes and links comprising the network have only two possible states: active and inactive, meaning $w_i(t), \pi_i(t) \in [0, 1]$ in Eq. 1. To simplify the explanation, we focus on node failures, though the concept may be extended to include links. If we denote the number of active nodes in the system at time t by $A(t)$ and the total number of nodes in the system by N, then the critical functionality simplifies to

$$K(t) = A(t)/N \tag{7}$$

We first consider a hierarchical multi-level DAG model (Figure 3) with Λ levels of nodes (Suzuki et al. 2003; Cho et al. 2012). We investigate how redundancy probability p_m, switching probability p_s, and the recovery time T_R, tradeoff parameters at the disposal of the system designer, influence the resilience of a supply-demand multi-level DAG across levels, nodes, links, and time, how they affect the absorption and recovery phases of a network's resilience profile, and how they address the optimization of network design, for a variety of attack scenarios. We also distinguish between cases where switching is instantaneous and delayed by one time step. Further description of the model is provided in the Methods. Two applications of the DAG model will serve to illustrate the quantitative resilience measure introduced, as well as the method for evaluating it: 1) synthetic random hierarchical multi-level supply-demand directed acyclic graphs, and 2) the Linux, specifically Ubuntu 12.04, software network. The first is a useful, if approximate, representation of networks found in many applications (Shapiro 2007; Yu and Gerstein 2006; Corominas-Murtra et al. 2013; Yan et al. 2010). The second realistically represents an existing and widely used network.

The second model is derived from the model introduced by Buldyrev et al. (2010) and developed by Parshani et al. (2010). They consider a system comprised of two coupled undirected networks (A and B). A certain fraction of nodes in network A depends on nodes in the network B (q_A) and vice versa (q_B). If node n in the network A depends on a node m in the network B then node m can only depend on node n (or not depend on nodes in the

17
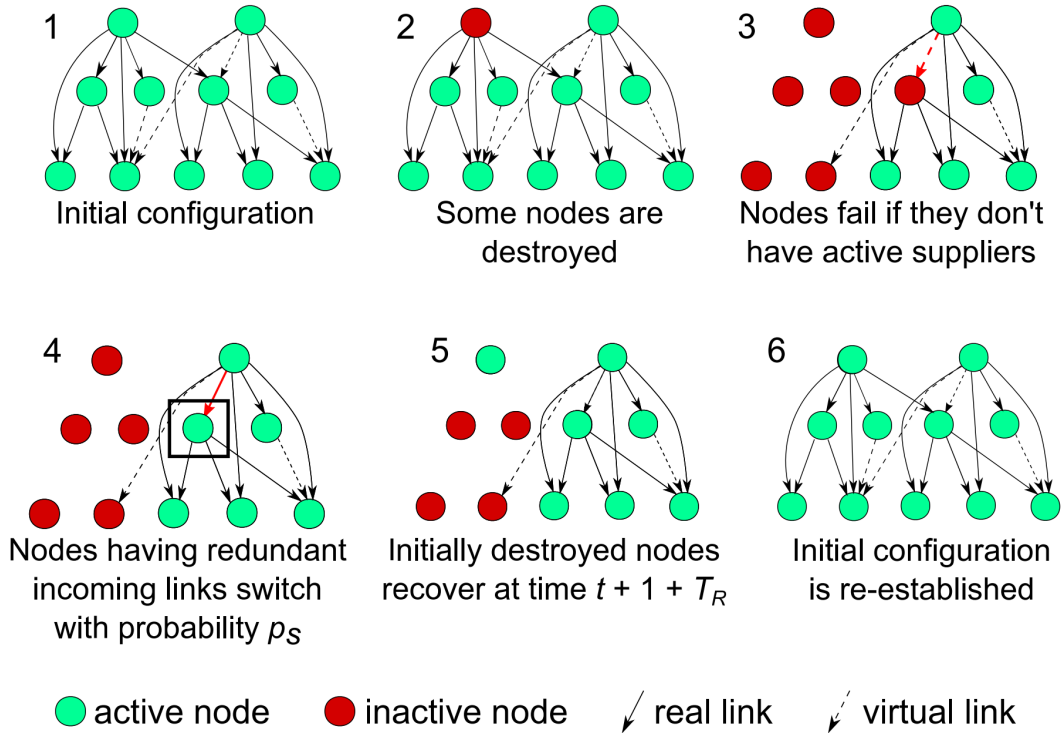
Figure 3. Network generation and modeling of an adverse event. The hierarchical area is first defined, then links are established according to the Bernoulli trial probability law, with parameter p_m. During the operation, if a node with a redundant link or links is made inactive, it can switch with probability p_s at each time step following attack time t. After the repair time period of T_R steps elapses following the attack, the initially destroyed nodes are restored.

network A at all) (see Methods). Without loss of generality we consider scenarios where networks A and B have the same node degree distribution. We present results for Erdos-Renyi and scale-free random networks with 800000 nodes (N) and average degree ($\langle k \rangle$) of 2.5, and slope factor (in the scale-free case) of 2.25. Networks are generated following the algorithm presented by Catanzaro et al. (2005).

We consider a case with a single adverse event that destroys a number of nodes in the network. For simplicity, the adverse event happens at the time step t = 0. We shall refer to the result of the adverse event as the initial damage. In case of the DAG model we denote the number of nodes that become inactive (i.e., are deactivated) in level i between time steps t and t + 1 as I_i^t. Thus, values I_i^0 represent the number of nodes made inactive upon

the occurrence of the adverse event. In case of the coupled networks model we denote the fraction of nodes rendered inactive in the network A as p_{destr} with the assumption that the adverse event doesn't affect the network B.

2.1.3 Results

Model 1 – Directed acyclic graphs. Synthetic graphs. We consider a network composed of N = 1000 nodes in four levels: $N_0 = 32, N_1 = 87, N_2 = 237, N_3 = 644$. We first look at the special case where the switching is instantaneous with probability $p_s = 1$. Assuming the overall damage to each level is small compared to the total number of nodes in the level, the approximation derived in the Methods section may be used. Figure 4 provides comparisons between the analytical calculations based on that approximation for this case and simulation results averaged over 2000 samples. We note that this case provides insight into the impact of link redundancy on both critical functionality and resilience over the time interval of interest, as we show in the first two scenarios of Figure 5. Examples of resilience profiles for different cases that vary in their initial damage, switching probability and recovery time are given in Figure 5. Case 1 is a scenario where only one node in the upper level is initially disabled. This scenario represents, for instance, an accident at a power plant. It follows, that the event set E consists of all possible one-node attacks in the upper level. Critical functionality suffers minimally; it reduces from 1 to 0.97 at its lowest. Its integral, resilience, consequently suffers minimally as well: R = 0.983. By contrast, for a more serious attack, such as in case 2, in which five nodes at every level are disabled (such an attack might represent a large earthquake in a certain area); both critical functionality and resilience suffer. Critical functionality can be as low as 0.8 (a considerably less robust system) for a protracted number of time steps, and resilience is reduced to 0.893. For case 3, 10 nodes are disabled, all from the top level, and the switching probability is reduced to $p_s = 0.25$. The robustness, or the critical functionality at its lowest, is more drastically reduced to 0.4, yielding an overall resilience value of 0.728. Case 4 is similar to case 3, except that the switching is delayed, i.e., if node i has become disabled at time t_*, then the first attempt to switch is made at time t* for case 3, and at time $t_* + 1$ for case 4.

The dependency of resilience on parameters p_m and p_s is given in Figure 6 (a) with the recovery time held constant at $T_R = 0.5T_C$. The figure shows that both parameters are compatible and combinable; they can be smoothly traded off to maintain a desired level of

19
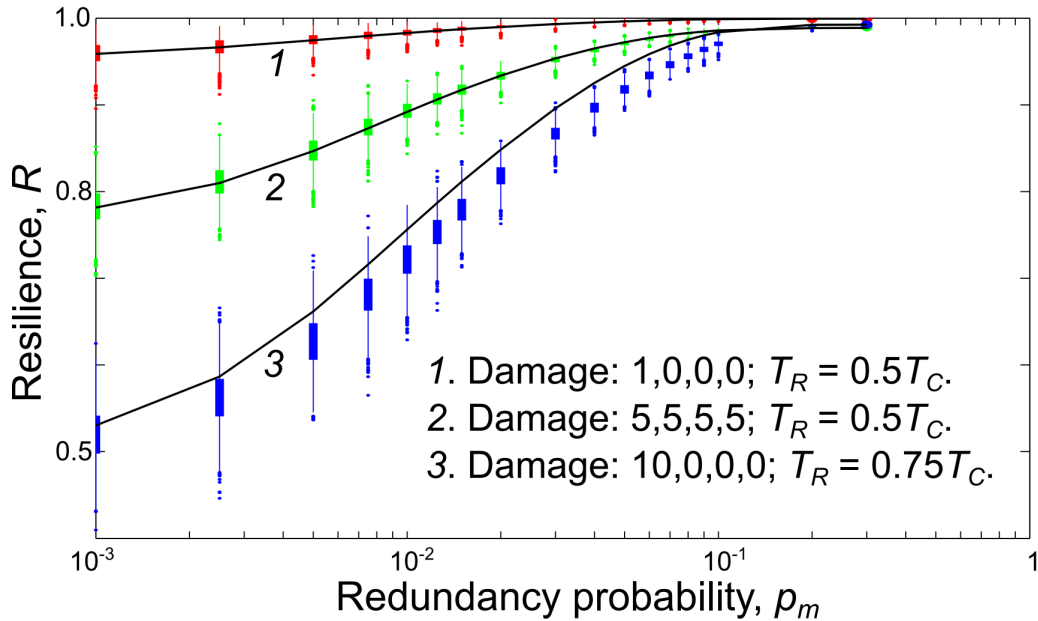
Figure 4. Special case: analytical vs computational solution. Comparison between the analytical solution and the computational experiments in the limiting but insightful special case in which switching is instantaneous when an additional link is available, meaning that $p_s = 1$, under three different scenarios. Initial damage numbers for each layer are ordered as follows: $I_0^0, I_1^0, I_2^0, I_3^0$. For instance, the initial damage in scenario 1 is: $I_0^0 = 1, I_1^0 = I_2^0 = I_3^0 = 0$. This special case reveals the impact on resilience of redundancy levels, as represented by the probability p_m. The cylinder represents the 25–75 percentile range.

resilience. The designer here has the opportunity to select the combination of p_m and p_s that is least costly. Additionally, increasing p_m and p_s simultaneously has an observable additive effect on resilience. Beyond certain level, however, investment in redundancy yields minimal return. For instance, as shown in Figure 6 (a), doubling the probability p_m from 0.1 to 0.2 leaves the resilience unchanged for $p_s > 0.3$. In addition, there is strong synergy between p_m and p_s; increasing both factors together produces a rapid increase in resilience, but increasing only one or the other variable will cause the resilience metric to plateau. This can be observed in Figure 6 (a) by regarding the resilience values shown across the phase diagram curves. Figure 6 (b) illustrates that similar tradeoffs can be made between the maximum node recovery time T_R, and the switching probability p_s. The redundancy parameter p_m is held constant at 0.01. When the recovery time is relatively
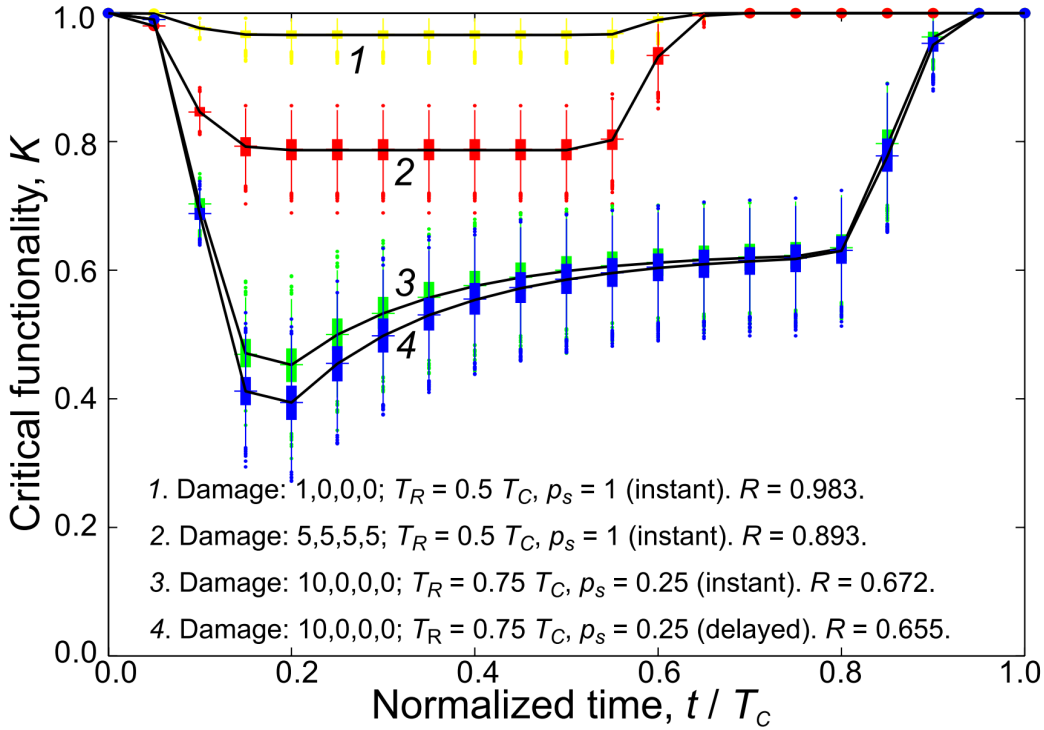
Figure 5. Resilience profiles for different scenarios in synthetic graphs. Results are shown for the redundancy probability parameter $p_m = 0.01$. Initial damage numbers for each layer are ordered as follows: $I_0^0, I_1^0, I_2^0, I_3^0$. For instance, the initial damage in scenario 1 is: $I_0^0 = 1, I_1^0 = I_2^0 = I_3^0 = 0$. The robustness (M) values for scenarios $1 - 4$ is the minimum value for each curve: 0.966, 0.787, 0.453, and 0.395 respectively.

short, $T_R < 0.1T_C$, resilience values close to 1 may be obtained even for values of p_s as small as 0.05. Resilience is strongly affected by the recovery time, T_R (Figure 6 (b)). This temporal factor determines the characteristics of the recovery phase and has a greater impact on the calculated resilience than does the potential increase in redundancy. This is particularly true when the switching probability p_s is low, as Figure 6 (b) demonstrates.

Cost and speed of design and implementation can now guide the ultimate choice from among the infinite possibilities of parameter combinations.

Model 1 – Directed acyclic graphs. Linux software network. The Linux software network exemplifies the structure of complex multilevel software systems and is important in its own right. This software operates in an estimated 95% of all supercomputing sys-
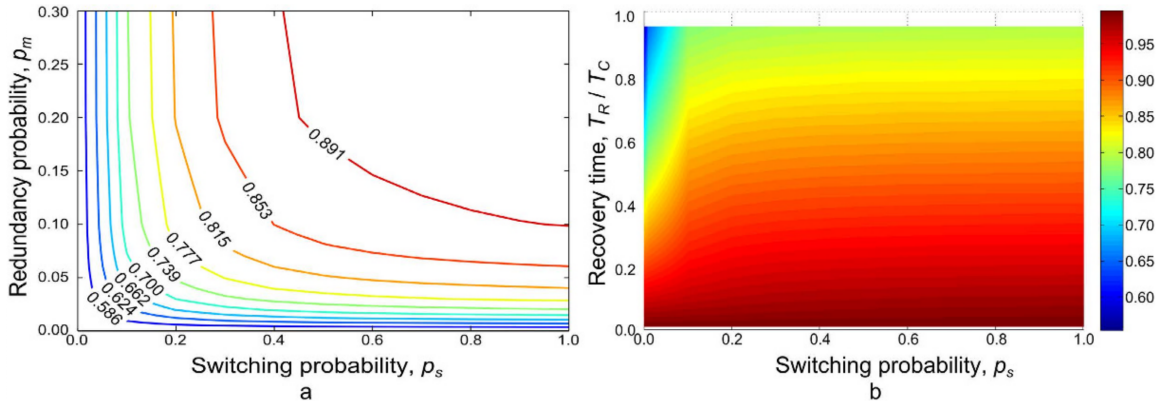
Figure 6. Resilience as a function of design parameters. (a) Resilience (value shown on curves, $\sigma \in [3.1E - 3; 2.0E - 2]$) dependencies on switching probability at each time step, or p_s, and redundancy parameter p_m, for a four level hierarchical network where the initial number of destroyed nodes at each level is $I_0^0 = 16, I_1^0 = I_2^0 = I_3^0 = 0$ respectively, and recovery time is held constant at $T_R = 0.5T_C$, and (b) resilience ($\sigma \in 5.9E - 4; 4.3E - 2]$) dependencies on p_s and T_R for $I_0^0 = I_1^0 = I_2^0 = I_3^0 = 5$ with constant $p_m = 0.01$ (color bar indicates the value of the resilience).

tems (Top500.org 2014), and the majority of the smartphones in use (in the form of the Android operating system). Packages in Linux are linked in a formally defined hierarchy of dependencies between individual software units. In this hierarchy, a package can only be installed if all required higher level packages have previously been installed. Some redundancy is possible when multiple packages provide the same functionality. Figure 7 shows a subnetwork of the packages network consisting of 117 nodes out of 36,902 possible nodes in the entire network. The graph data were obtained using Advanced Packaging Tool (Nussbaum 2014) on a standard installation of the Ubuntu 12.04 system.

Many modern cyber threats exploit vulnerabilities in software packages. Disabling a targeted software package leads to the failure of any services that are dependent on it. Even worse, the recovery might be protracted as a result of corrupted user data, thus requiring manual repair and cleanup. For example, an attack on the Apache web server might cause it to fail and subsequently send corrupted or maliciously designed data to backend databases (Kargl et al. 2001). Consequently, services dependent on Apache would experience data corruption, and if Apache crashes, it would be disabled as well. While the damaged server might be restarted relatively quickly, recovery from such an attack would involve checking
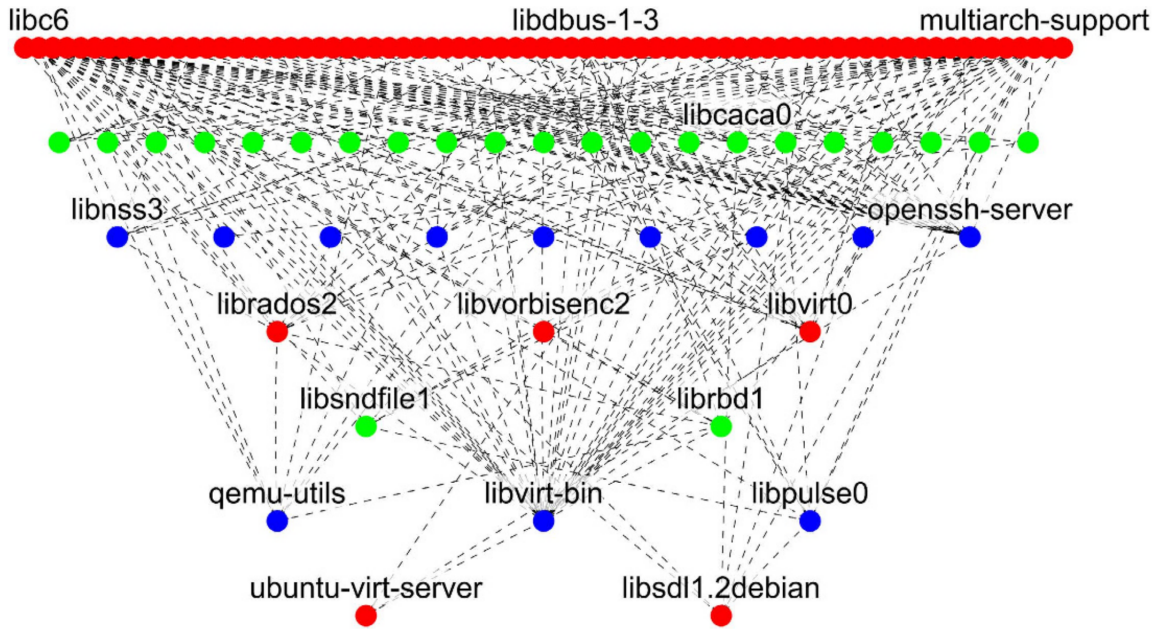
Figure 7. Subnetwork of the Linux hierarchical packages network.

the data, which gives rise to serious additional delays.

We evaluate the resilience of the Linux packages network in the presence of both random and guided attacks. Critical functionality and resilience profiles for guided attacks on several particularly important packages are given in Figure 8 (a). These packages are: xauth, libstdc++6, libc6, and gcc-4.6-base. Notably in these four cases there are four sets of adverse events E. Each of these sets contains only one event that successfully causes a particular node to be destructed. It is seen that the level of damage depends on which packages are targeted. For random attacks (Figure 8 (b)), we consider another set of adverse events, consisting of 36,902 events. In this case, resilience is significantly higher than in the case of guided attacks due to the low importance of many of the nodes that failed from the attacks, thus yielding R = 0.99975 and M = 0.999.

Model 2 – Interdependent networks. Synthetic graphs. We summarize the results for the second model in the Figure 9. Panels (a) and (b) show the dependency of the critical functionality of a system of two interdependent Erdos-Renyi (ER) networks on time for 2 distinct cases of the recovery resources available expressed as the number of the backup agents (N_b). As it is evidenced by the Figure 9 depending on the value N_b there is a
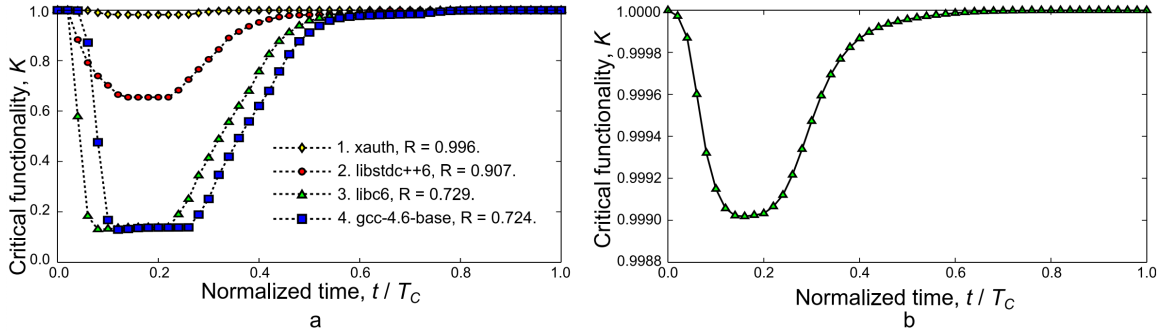
Figure 8. Resilience profiles for the Linux network. (a) Guided attacks, and (b) random attacks. It is clear that guided attacks are considerably more damaging. Moreover, not all guided attacks are equally damaging; as shown in (a), attacks on xauth are less damaging than on libstdc++6. Most damaging are attacks on libc6 and gcc-4.6-base. The robustness (M) values for scenarios 1 − 4 in the panel (a) are 0.982, 0.655, 0.130, and 0.129 respectively, while for the case in the panel (b) M = 0.999.

sharp distinction between two cases: in Figure 9 (a) the system is unable to return to the recovered state and the critical functionality oscillates between 0 and about 0.5; due to random duration of cascading recovery and failure processes eventually the amplitude of the oscillations of the mean value of CF decreases; in Figure 9 (b) the backup supply to 0.4N nodes in the network A allows reaching a stable state after the backup supply removed and further recovery of the system.

In scale-free networks (SF) the results for <k> = 2.5 and N = 800000 (Figure 9 (d), (e)) show a much larger dispersion than in Erdos-Renyi networks. In particular some of those networks suffer a much smaller drop in critical functionality in response to the adverse events modeled. This obviously is a consequence of the infinite dispersion in degrees distribution in the SF networks (though in our case the dispersion is finite due to the finite number of nodes). Another distinct specific trait of the SF networks with small value of the average degree is the fact that whether the network fully recovers or not is strongly dependent on the stochastic nature of the cascading failure process. Particularly it is obvious from Figure 9 (d) that the success of recovery is determined by whether the most important hubs were affected during the deactivation process. If those hubs are not affected the damage is relatively small, otherwise the damage results in large drop in critical functionality and no full recovery within the control time.

Finally panels (c) and (f) in the Figure 9 show the phase diagrams of the dependency of the resilience (calculated as the integral of the CF over the control time) on the model parameters p_{destr} and N_b. Notably the critical functionality practically drops to zero for only 0.2 of the nodes destroyed in the network A initially (p_{destr}). Parshani et al. (2010) has demonstrated that if p_{destr} is more than 0.2545 the network experiences the first order transition leading to a state with almost no active nodes. We have reproduced their results for the Erdos-Renyi networks and confirmed that if p_{destr} is less than the threshold of 0.2545 the transition doesn't occur. However due to minor modifications we made to the network generation algorithms aimed at connecting all the nodes in a single giant component (GC) in the beginning of the process, we observe decrease of the threshold value p_{destr}, causing the first order transition, to about $0.15 - 0.2$. After the drop of the critical functionality (due to the cascading failure) on the step T_R the recovery process starts. The recovery is successful only if N_b is about 0.4N or higher. Finally if the whole network A is destroyed as a result of the adverse event ($p_{destr} = 1$) then the recovery cannot start due to the absence of the GC (A). Results for the scale-free network show similar tendencies although notably are much more disperse in the region $N_b \in [0.1; 0.6]$, $p_{destr} \in [0.05; 0.85]$. We interpret this as the consequence of the divergence of the standard deviation of the degree distribution in the scale-free networks with $\gamma \in [2; 3]$.

2.1.4 Conclusion

We have presented a detailed approach for implementing the NAS definition of resilience as a function of design tradeoff parameters, as illustrated in the study with multi-level directed acyclic graphs and interdependent networks. The approach allows evaluation of resilience across time, and not just as a single quantity. Designers can thus analyze the effect of parameter choice and design emendations on overall network resilience and robustness. Focusing on multi-level directed acyclic graphs and interdependent networks, we have demonstrated how network parameters can be traded off to obtain a desired resilience and other performance measures' level. Future work will extend to multiplex systems, other real life networks. An important long term challenge is to model adaptation, which is part of the response cycle that follows restoration and includes all activities that enable the system to better resist similar adverse events in the future.
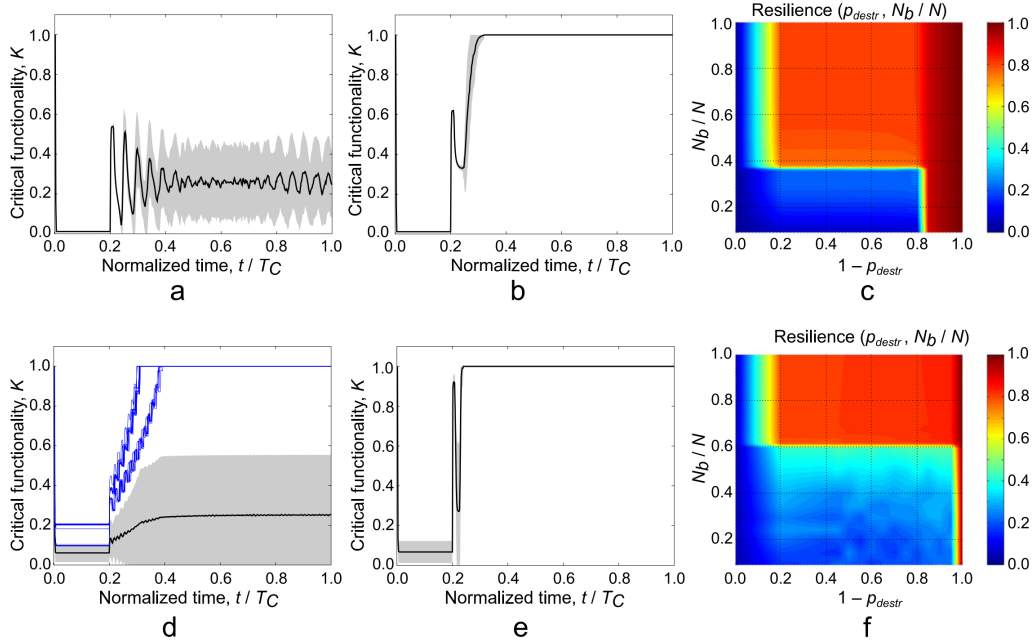
Figure 9. Representative profiles of the dynamics of K and resilience in networks with $N = 800000$. Panels (a) and (b) show results for ER networks with $p_{destr} = 0.5, Nb = 0.35N$ and $Nb = 0.4N$ respectively. Panels (d) and (e) show results for SF networks with $p_{destr} = 0.5, Nb = 0.1N$ and $Nb = 0.62N$ respectively. In the panels (a), (b), (d), (e) the solid line corresponds to the mean value of K over 100 simulations. Gray area corresponds to the region $K \pm \sigma(K)$ (where σ is the standard deviation). The plot in the panel (d) also shows simulations where critical functionality restores to 1. It follows that the success of the restoration algorithm depends for the most part only on the results of the cascading failure (rather than the random selection of the nodes with the backup supply): in cases when the important hubs are active after the cascading failure recovery is more likely. Finally panels (c) and (f) display phase diagrams of the resilience dependencies on both N_b and p_{destr} parameters. In both ER and SF networks the recovery process is stochastic and very sensitive to the backup supply available.

2.1.5 Methods

Absorption and recovery algorithms in the DAG model. A hierarchical multi-level DAG (Figure 3) has Λ levels of nodes (Suzuki et al. 2003; Cho et al. 2012). Each layer is comprised of N_i nodes ($i = 0, \ldots, \Lambda\check{~}1$). The links represent a supply–demand relationship. A link starts at a supplier node and ends at a demander node. Thus, every link in the network is directed. For every level, we identify a set of services that all nodes in a particular level supply. Then, for every service supplied in the network, we define whether or not it depends on other services and a (possibly empty) list of the dependencies. The levels are ordered in such a way that links only go from a higher level i to a lower level j. With this convention, i < j, or, the higher the level, the smaller the index. Additionally, no links can be formed between nodes in the same level. We also disallow cycles in the network by imposing the following constraint: a node cannot depend on any of the services provided by any of the other nodes in its level or on any of the services provided in any of the lower levels. Initially, all dependencies are resolved, and every node has one incoming link from one or more upper levels on which it depends for the supply of its services. Furthermore, for every dependent node and each of its required services, we introduce a list of potential suppliers. The probability that a node has a link from each of the potential suppliers is p_m. Said another way, a node has many links supplying a given service but only one of those links is enabled initially (known as real), while the others are contingent backup links (known as virtual), should they exist.

To model an adverse event, we introduce an ability to destruct a node for a time period T_R, as was recently done by Majdandzic et al. (2013). A destructed node is inactive and is therefore unable to supply services until it recovers. Another possible cause of deactivation is an unresolved dependency, that is, the absence of a real link to a node supplying a required service. This can happen when the only supply node available for a service is either destructed or its upstream supplier is destructed. We shall refer to the nodes with an unresolved dependency as disabled nodes. Note that a node can be disabled, destructed, or both at a given time. Once a node becomes inactive, all of its dependencies, connected through their real links, are subject to deactivation unless they have other real links providing all of their required services (Figure 3).

We assume that a node is eligible to switch links, that is, to turn a virtual or contingent link into a real one, if virtual links for all of the node's unresolved dependencies exist and

the node is not destructed. At every time step during which the node is both disabled and eligible to switch, switching happens with probability p_s. Switching can be either instant (the first attempt to switch is made at the same time step the node has become disabled) or delayed (the node with an unresolved dependency remains disabled for at least one time step).

Analytical approximation for the special case of the DAG model. In this section we derive equations describing the number of active nodes in the special case where the switching is instantaneous with probability $p_s = 1$ while the initial damage is small compared to the total number of nodes in the network.

Let us denote by Λ the number of levels in the network and by N_i the number of nodes in level $i(i = 0, \ldots, \Lambda\check{}1)$. We can find the probability that a node in level i has only one service provider in level j as follows:

$$f_j = (1 - p_m)^{(N_j - 1)} \tag{8}$$

For the case where the number of deactivated nodes at each time step is small enough or in which $p_m = 0$, we may assume that only the nodes with one link for a relevant service are disabled as a result of the inactivation of their supplier (thus neglecting the cases in which the node has more than one supplier of a service and all of them are deactivated).

The average number of active nodes in level i at time step $t(A_i^t)$ is given by the formula:

$$A_i^t = N_i - \sum_{s=0}^{t-1} I_i^s \tag{9}$$

These A_i^t nodes have A_j^t suppliers in each level $j < i$. We disabled I_j^{t-1} suppliers of the nodes in level j between steps $t - 1$ and t. The probability for a node in level i to become disabled between time steps t and t + 1 is therefore:

$$1 - \prod_{j=0}^{i-1} \left(1 - \frac{I_j^{t-1} f_i}{A_j^{t-1}}\right) \tag{10}$$

We approximate the distribution of $\frac{1}{A_j^{t-1}}$ to be linear in I_i^s, although the dependency itself is

28

not linear:

$$\frac{1}{A_j^{t-1}} = \frac{1}{N_j - \sum_{s=0}^{t-1} I_j^s} = \frac{1}{N_j} \frac{1}{1-x} \tag{11}$$

Where (under the assumption that the overall damage is small enough):

$$x = \frac{\sum_{s=0}^{t-1} I_j^s}{N_j} << 1 \tag{12}$$

Considering the Taylor expansion of $1 / (1-x)$, we have for small values of x: $1 / (1-x) \approx 1 + x$. Then, on average between steps t and t + 1, we disable the following number of nodes in level i:

$$I_i^t = \left[1 - \prod_{j=0}^{i-1} \left[1 - \left(\frac{I_j^{t-1} f_j}{\left(N_j - \sum_{s=0^{t-2}} I_j^s\right)}\right)\right]\right] \left(N_i - \sum_{s=0}^{t-1} I_i^s\right) \tag{13}$$

After the recovery period of T_R time steps has transpired, the initially destroyed nodes are rebuilt and become active unless they still lack sufficient supplies from the upper levels. Thus, assuming that $T_R > \Lambda$, $I_i^t = 0$ for $i = [0, ..., \Lambda]$ and $t = [\Lambda, ..., T_R - 1]$. The total number of nodes restored at step T_R in level i is given by the expression:

$$|I_i^{T_R}| = \frac{|I_i^0|}{N_i} \left(N_i - F_i^{T_R}\right) \tag{14}$$

Here, F_i^t represents the total number of nodes disabled in level i at time step t due to the fact that they do not have sufficient supplies from the upper levels:

$$F_i^t = \left[1 - \prod_{j=0}^{i-1} \left[1 - \frac{\left(N_j - A_j^{t-1}\right)}{N_j}\right]\right] N_i \tag{15}$$

And for the next steps, the formula is as follows:

$$|I_i^t| = \left[1 - \prod_{j=0}^{i-1} \left(1 - \frac{|I_i^{t-1}| f_j}{A_j^{t-1}}\right)\right] \left(N_i - F_i^t\right) \tag{16}$$

Using the formulae above, we may evaluate the average approximated resilience profiles and find the values of resilience.

Absorption and recovery algorithms in the coupled networks model.   The failure propagation algorithms are described in the original model of Buldyrev et al. (2010). Initial damage results in a certain fraction of nodes deactivated in the network A. Once those nodes are deactivated the network A is fractured in clusters. Nodes that do not belong to the largest cluster of the network A are also assumed deactivated. Then all the nodes in the network B that depend on the deactivated nodes in the network A are also deactivated. It results in fracturing of the network B, and the nodes that are not in the largest cluster of the network B are also assumed deactivated. In the second step of the process nodes in the network A depending on the deactivated nodes in the network B are deactivated and the process propagates in the same fashion until there are no more nodes to deactivate in any of the networks.

Recovery is accomplished by the backup supply agents replacing unresolved dependencies of the nodes in the first network (A). The number of those agents is denoted N_b. Each backup agent can serve only one node at a time. Nodes to provide the backup supply to are chosen randomly from those nodes in the network A that depend on a currently inactive node in the network B. Thus backup is provided either to all nodes in the network A with an unresolved dependency (in this case full recovery is guaranteed) or to N_b nodes only. If a node has backup supply and it is connected to its network's GC it is activated. Once a node is activated it is included in the network's GC. This causes eventual growth of the giant component of the network A. After that the nodes in the network B that depend on the activated nodes in the network A and are connected to the GC of the network B are also activated and the process propagates in a similar fashion. Once the process is complete the recovery phase finishes. After that the backup supply is removed meaning that all the nodes whose supplier in the network B is not active (after the recovery phase) are deactivated. This leads to a cascading failure propagating as described in the introduction section. Once the failure phase finishes the recovery phase is repeated until the full network recovery is established or the control time is reached.

Let us consider a simple two-network system (Figure 10). At the beginning of simulation (time 0) two nodes (A1 and A3) are assigned to be initially destroyed. Cascading process finishes before the recovery time (that is time to repair a node is more than the cascading failure time) T_R. Thus at time T_R the network is in the state it had after the cascading failure. After T_R steps have passed the recovery process starts. The case of $N_b = 0$ is shown in the

panel (a) of the Figure 10. In this case the only recoverable node is A3. After its recovery node B3 is also recovered, but further recovery is not possible. Nodes A1, A2, B1, and B2 can't be recovered as they have an unresolved dependency. In addition even if nodes A1, B1 were independent they still wouldn't have been recoverable due to the fact that they are not connected to the respective networks' GCs. Now consider the case $Nb = 1$. In this case two scenarios are possible: a) The node chosen for backup supply is A1 (Figure 10 (b)). Then no recovery can happen as this node is not connected to the network A GC (or GC(A)) and the recovery phase ends in 0 steps; b) The node chosen for backup supply is A2 (Figure 10 (c)). Then this node recovers, node B2 recovers in turn. During the second step of the recovery phase node A1 recovers and node B1 recovers in turn.
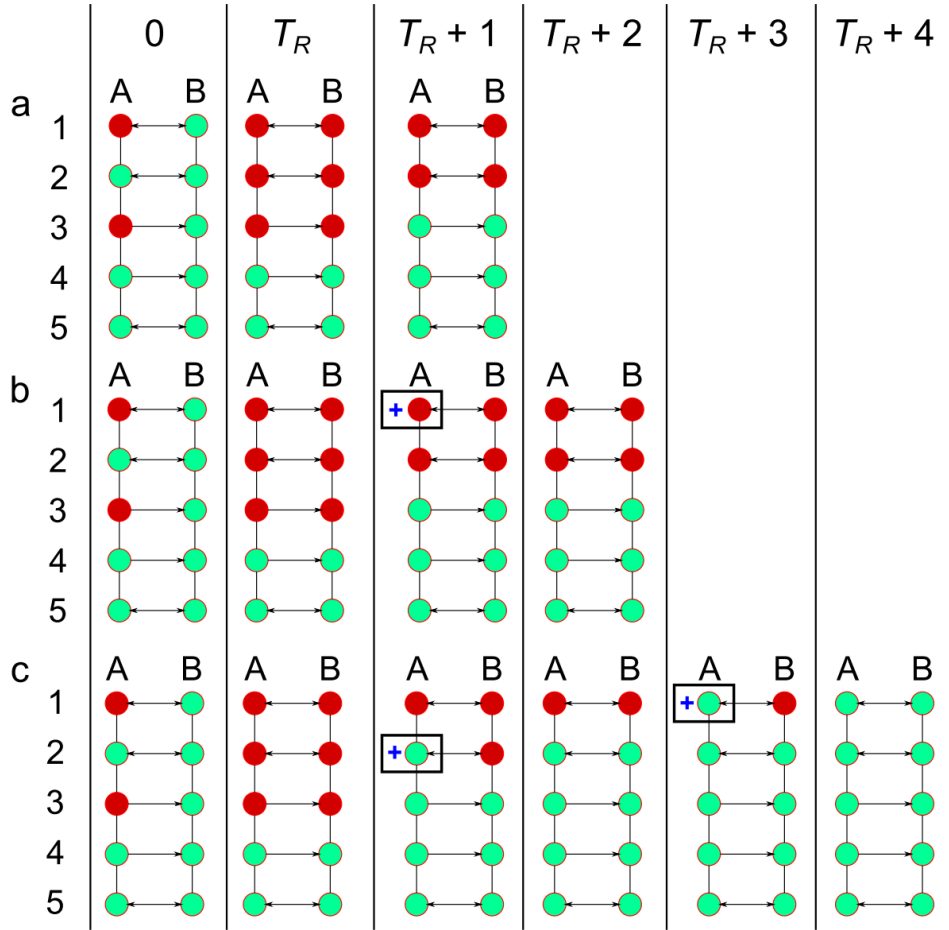
Figure 10. Recovery process in coupled networks. At time 0 nodes A1 and A3 are deactivated for T_R steps. The failure process finishes before T_R so at time T_R the network is in its state after the cascade. Panel (a) illustrates the case $N_b = 0$. Only independent node A3 can be recovered as it is connected to the network A GC. Once node A3 is activated its dependent node B3 is also activated. Panels (b) and (c) illustrate the stochastic nature of the recovery process. In these cases $N_b = 1$. At time $T_R + 1$ either node A1 (b) or node A2 (c) can have backup supply. In case (b) the recovery phase doesn't start. After that backup is removed but cascading failure occurs. On the next time step when backup is applied again to a randomly selected node the recovery cascade is possible if the node chosen is A2. This case is the same as the case (c) with time $T_R + 1$ (in the case (c)) corresponding to the time $T_R + 3$ (in the case (b)).

2.2 Resilience management during epidemic outbreaks

Assessing and managing the impact of large-scale epidemics considering only the individual risk and severity of the disease is exceedingly difficult and could be extremely expensive. Economic consequences, infrastructure and service disruption, as well as the recovery speed, are just a few of the many dimensions along which to quantify the effect of an epidemic on society's fabric. Here, we extend the concept of resilience to characterize epidemics in connected populations, by defining the system-wide critical functionality that combines an individual's risk of getting the disease (disease attack rate) and the disruption to the system's functionality (human mobility deterioration). By studying both conceptual and data-driven models, we show that the integrated consideration of individual risks and societal disruptions under resilience assessment framework provides an insightful picture of how an epidemic might impact society. In particular, containment interventions intended for a straightforward reduction of the risk may have net negative impact on the system by slowing down the recovery of basic societal functions. The presented study operationalizes the resilience framework, providing a more nuanced and comprehensive approach for optimizing containment schemes and mitigation policies in the case of epidemic outbreaks.

2.2.1 Introduction

Data-driven models of infectious diseases (Riley 2007; Marathe and Vullikanti 2013; Balcan et al. 2009b,a; Grais et al. 2003; Colizza et al. 2007; Hufnagel et al. 2004; Eubank et al. 2004; Merler and Ajelli 2010; Cooper et al. 2006; Ferguson et al. 2003, 2006; Germann et al. 2006; Epstein et al. 2007; Degli Atti et al. 2008) are increasingly used to provide real- or near-real-time situational awareness during disease outbreaks. Indeed, notwithstanding the limitations inherent to predictions in complex systems, mathematical and computational models have been used to forecast the size of epidemics (Team 2014; Meltzer et al. 2014; Rivers et al. 2014; Shaman et al. 2014), assess the risk of case importation across the world (Cooper et al. 2006; Epstein et al. 2007; Gomes et al. 2014), and communicate the risk associated to uncurbed epidemics outbreaks (Van Kerkhove and Ferguson 2012; Lipsitch et al. 2011; Van Kerkhove et al. 2010). Despite contrasting opinions on the use of modelling in epidemiology (Merler et al. 2015), in the last few years a large number of studies have employed them to evaluate disease containment and mitigation strategies as well as to inform contingency plans for pandemic preparedness (Ferguson et al. 2003;

Germann et al. 2006; Degli Atti et al. 2008; Merler et al. 2015; Wu et al. 2006). Model-based epidemic scenarios in most cases focus on the "how many and for how long?" questions. Furthermore, mitigation and containment policies are currently evaluated in the modelling community by the reduction they produce on the attack rate (number of cases) in the population. These studies aim at identifying best epidemic management strategies but typically neglect the epidemic and mitigation impact on the societal functions overall. The evaluation of vulnerabilities and consequences of epidemics is a multi dimensional complex problem that should consider societal issues such as infrastructures and services disruption, forgone output, inflated prices, crisis-induced fiscal deficits and poverty (World Health Organization 2009; S.N. et al. 2016). Therefore, it is important to broaden the model-based approach to epidemic analysis, expanding the purview by including measures able to assess the system resilience, i.e. response of the entire system to disturbances, their aftermath, the outcome of mitigation as well as the system's recovery and retention of functionality (Holling 1973; Linkov et al. 2013b, 2014c). Most important, operationalizing resilience (Linkov et al. 2013b, 2014c; Linkov and Palma-Oliveira 2017b) must include the temporal dimension; i.e. a system's recovery and retention of functionality in the face of adverse events (Linkov et al. 2014c; Ganin et al. 2016; Barrett and Constas 2014; Majdandzic et al. 2014; Gao et al. 2016). The assessment and management of system resilience to epidemics must, therefore, identify the critical functionalities of the system and evaluate the temporal profile of how they are maintained or recover in response to adverse events. Even though the assessment and management of adverse events resilience of complex systems is the subject of active research (Ganin et al. 2016; Barrett and Constas 2014; Majdandzic et al. 2014; Gao et al. 2016), its integration in the computational analysis of epidemic threats is still largely unexplored (S.N. et al. 2016; Linkov et al. 2014b; Lu et al. 2017). Here, we introduce a resilience framework to the analysis of the global spreading of an infectious disease in structured populations. We simulate the spread of infectious diseases across connected populations, and monitor the system–level response to the epidemic by introducing a definition of resilience that compounds both the disruption caused by the restricted travel and social distancing, and the incidence of the disease. We find that while intervention strategies, such as restricting travel and encouraging self-initiated social distancing, may reduce the risk to individuals of contracting the disease, they also progressively degrade population mobility and reduce the critical functionality thus making the system less resilient. Our numerical results show a transition point that

34

signals an abrupt change of the overall resilience in response to these mitigation policies. Consequently, containment measures that reduce risk may drive the system into a region associated with long-lasting overall disruption and low resilience. Interestingly, this region is in proximity of the global invasion threshold of the system, and it is related to the slowing down of the epidemic progression. Our study highlights that multiple dimensions of a socio-technical system must be considered in epidemic management and sets forward a new framework of potential interest in analyzing contingency plans at the national and international levels.

2.2.2 Results

We provide a general framework for the analysis of the system-level resilience to epidemics by initially considering a metapopulation network (Figure 11A). In this case we consider a system made of V distinct subpopulations. These form a network in which each subpopulation i is made of N_i individuals and is connected to a set k_i of other subpopulations. A complete description of the networked systems is given in the Methods section. The notation and the description of the parameters used in our simulations are reported in Table 4.
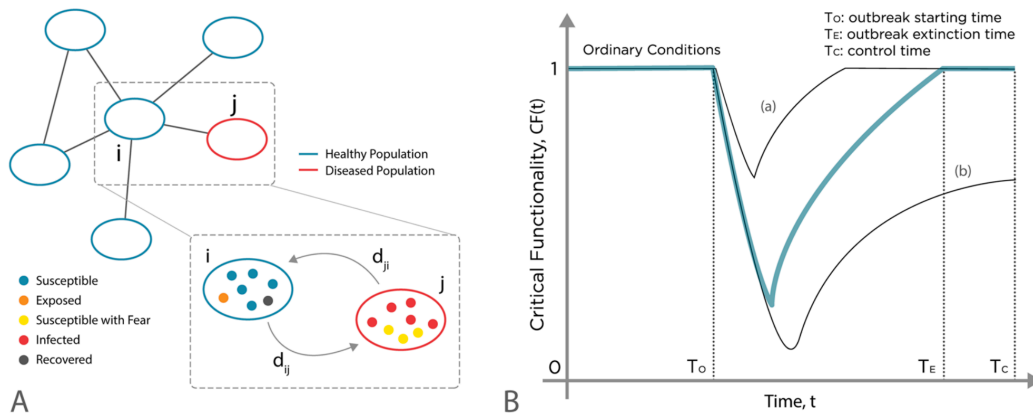
Figure 11. Schematic representation of the metapopulation model. (A) The system is composed of a network of subpopulations or patches, connected by diffusion processes. Each patch contains a population of individuals who are characterized with respect to their stage of the disease (e.g. susceptible, exposed, susceptible with fear, infected, removed), and identified with a different color in the picture. Individuals can move from a subpopulation to another on the network of connections among subpopulations. At each time step individuals move with a commuting rate c_{ij} from subpopulation i to subpopulation j. (B) Schematic illustration of the system's critical functionality. The system if fully functional ($CF(t) = 1$) during ordinary conditions when all the subpopulations are healthy and the number of real commuters is equal to the number of virtual commuters, i.e. $D(t) = 0$ and $C(t) = Z(t)$. After the outbreak takes place (T_0) the system's functionality decreases because of the disease propagation and the eventual travel reduction. Next the system starts to recover until the complete extinction of the epidemic (T_E) which corresponds to the time when no more infected individuals are in the system. The curves (a) and (b) represent the critical functionality of scenarios corresponding to high and low values of resilience.

36

Table 4. Notation and description of the parameters used in our simulations.

Notation	Description
V	Number of subpopulations in the metapopulation network
N	Total number of individuals in the system
$\langle k \rangle$	Average degree of the metapopulation network
D	Fraction of diseased populations
H	Fraction of healthy populations
A	Fraction of active travelers in the system
p	The parameter that regulates the system wide travel restrictions
r	System's resilience
CF	Critical Functionality
T_c	Resilience control time
S	Susceptible individuals
S^F	Susceptible with fear individuals
E	Exposed individuals
I	Infected individuals
R	Recovered individuals
R_0	Basic reproduction number
λ	The rate at which an 'exposed' person becomes 'infected'
μ	The rate at which an 'infected' recovers and moves into the 'recovered' compartment
β	The parameter controlling how often a 'susceptible'-'infected' contact results in a new 'exposed'
β^F	The parameter controlling how often a 'susceptible'-'infected' contact results in 'susceptible with fear individual'
α	The parameter controlling how often a 'susceptible'-'susceptible with fear individuals' contact results in a new 'susceptible with fear individual'
r_b	The parameter that modulates the level of self-induced behavioral change that leads to the reduction of the transmission rate
μ_F	The rate at which individual with fear moves back into the 'susceptible' compartment

Diffusion Processes. The edge connecting two subpopulations i and j indicates the presence of a flux of travelers i.e. diffusion, mobility of people. We assume that individuals in the subpopulation i will visit the subpopulations j with a per capita diffusion rate d_{ij} on any given edge (Colizza and Vespignani 2008) (see the Methods section for further details). We define the total number of travelers Z between the subpopulations i and j at time t as $Z_{ij}(t) = d_{ij}N_i(t)$, so that when the system is fully functional, the total number of travelers at time t from the node i is $Z_i(t) = \sum_{j \in k_i} Z_{ij}(t)$. Under these conditions, the total number of travelers in the metapopulation system at time t is simply $Z(t) = \sum_i Z_i(t)$. In the following we assume that infected individuals do not travel between subpopulations, thus reducing the actual number of travelers.

Reaction Processes. In analyzing contagion processes we extend the compartmental scheme of the basic SEIR model (Anderson et al. 1992; Keeling and Rohani 2008) (see Methods and Supplementary Information (SI) for a detailed description). Indeed an important element in the mitigation of epidemics is self-initiated behavioral changes triggered in the population by awareness/fear of the disease (Funk et al. 2010; Massaro and Bagnoli 2014). These generally reduce the transmissibility and spreading. Examples of behavioral changes include social distancing behaviors such as avoidance of public places, working from home, decrease of leisure and business travel etc. In order to include behavioral changes in our model, we consider a separate behavioral class within the population (Perra et al. 2011), defining a special compartment of susceptible individuals, S^F, where F stands for "fearful". In particular, individuals transition to this compartment depending on the prevalence of infected and other fearful individuals according to a rate β_F. This rate mimics the likelihood that individuals will adopt a different social behavior as a result of the increased awareness of the disease as perceived from the number of infected and fearful individuals present in the system. Clearly, spontaneous or more complex types of transitions (for example indirectly linked to the disease transmission due to mass media effects (Perra et al. 2011)) could be considered. However, they would require more parameters and introduce other non-trivial dynamics. We leave the study of other behavioral changes models for future works. It follows that in each subpopulation the total number of individuals is partitioned into the compartments $S(t), S^F(t), E(t), I(t), R(t)$ denoting the number of susceptible, fearful, exposed, infected, and removed individuals at time t, respectively. The transition processes are defined by the following scheme: $S + I \rightarrow E + I, S + I \rightarrow S^F + I, S + S^F \rightarrow 2S^F, S^F + I \rightarrow E + I, E \rightarrow I$ and $I \rightarrow R$ with their respective reaction rates, $\beta, \beta_F, \alpha\beta_F, r_b\beta, \lambda$ and μ. Analogously, individuals in the S^F compartment may transition back in the susceptible compartment with a rate $\mu_F, S^F + S \rightarrow S$. The model reverts to the classic SEIR if $\beta_F = 0$ (the detailed presentation of the dynamic is reported in the SI). The basic reproductive number of an SEIR model is $R_0 = \beta/\mu$. This quantity determines the average number of infections generated by one infected individual in a fully susceptible population. In each subpopulation the disease transmission is able to generate a number of infected individuals larger than those who recover only if $R_0 > 1$, yielding the classic result for the epidemic threshold (Colizza and Vespignani 2007); if the spreading rate is not large enough to allow a reproductive number larger than one (i.e., $\beta > \mu$), the epidemic outbreak will affect only a negligible portion of the population and

will quickly die out (the model details are reported in the Methods section).

System's resilience. Here, we introduce a quantitative measure that captures and implements the definition of resilience in epidemic modelling, similarly to what proposed in Ganin et al. (2016). Among the many possible elements defining the resilience of a system, we consider the system-wide critical functionality as a function of the individual's risk of getting the disease and the disruption to the system's functionality generated by the human mobility deterioration. For the sake of simplicity, in our model we assume that infected individuals do not travel. The extension to models in which a fraction of infected individuals are traveling is straightforward (Balcan et al. 2009a) with the only effect of decreasing the timescale for the disease spreading, but not altering the overall dynamic of the system. Furthermore, as discussed below, the system might be subject to other travel limitations. As a result, during the epidemic we have an overall decrease in the mobility flows with respect to a disease-free scenario. It follows that the number of travelers between subpopulations i and j at time t is $C_{ij}(t) = c_{ij} \tilde{N}_i(t)$, where c_{ij} is the adjusted diffusion rate, $\tilde{N}_i(t) = S_i(t) + E_i(t) + R_i(t)$, and the total number of commuters in the metapopulation system at time t is given by $C(t) = \sum C_i(t)$. Note that in general, $c_{ij} < d_{ij}$. This can be naturally related to a deterioration of the system-level critical functionality as it corresponds to economic and financial losses as well as logistic and infrastructural service disruption. In order to evaluate the system's loss of critical functionality related to the travel restrictions, we define the fraction of active travelers at time t as $A(t) = C(t)/Z(t)$. Analogously, we characterize the system's risk related to the disease propagation as the fraction of healthy subpopulations $H(t) = 1 - D(t)/V$, where $D(t)$ is the number of diseased subpopulations at time t and V is the total number of subpopulations in the system. The number of diseased subpopulations accounts for the amount of risk posed to individuals in the system, which we assume to be proportional to the overall attack rate and expresses the vulnerability of the networked system (Majdandzic et al. 2014; Buldyrev et al. 2010; Gao et al. 2011). Here, as the model assumes statistically equivalent subpopulations, the attack rate is proportional to the number of subpopulations affected by the epidemic. At time t, we define the critical functionality, $CF(t)$, (Figure 11B), as the product of the fraction of active travelers $A(t)$ and the fraction of healthy populations $H(t)$, i.e. $CF(t) = H(t).A(t)$. Per our earlier definition of resilience (Ganin et al. 2016) r, we evaluate it as the integral over time of the critical

functionality, normalized over the control time T_C so that $r \in [0, 1]$:

$$r = 1/T_C \int_0^{T_C} CF(t)dt \tag{17}$$

The control time T_C corresponds to the maximum extinction time T_E for different values of epidemic reproductive number R_0 (see the Supporting Information for further detail). Resilience, therefore, also includes the time dimension, in particular, the time to return to full functionality, as defined by the system's critical elements. In reference (Ganin et al. 2016) we provided an operational definition of resilience starting from the concepts advanced by the National Academy of Sciences in USA. In this paper, we apply such general framework to the case of disease spreading. Furthermore, we extend it to reaction-diffusion processes on metapopulations. In the following, we will quantitatively characterize different containment/mitigation interventions via a critical functionality analysis. Desirable (optimal) strategies correspond to high (maximum) value of r. It is worth remarking that, for the sake of simplicity, we use here a definition of critical functionality that weights equally the two components $A(t)$ and $H(t)$. Thus, our findings are constrained by such choice. The two contributions could be weighted differently, i.e. $CF(t) = H(t)^\alpha.A(t)^\beta$. However, our aim is to highlight the importance of going beyond "model-based" approach to epidemic analysis and move towards system resilience assessments. In this spirit, we opted for the simplest definition of critical functionality able to capture the two most used metrics in model-based approaches: epidemic risk and mobility. We used the multiplication of the two quantities because it makes the critical functionality more sensitive to small changes of the values. Furthermore, by multiplying two ratios we don't need to add a normalization factor (the critical functionality is defined in the interval [0,1]). In more realistic context, and depending on the precise cost-benefit analysis, the various terms may be weighted differently and more complex functional form for the critical functionality can be defined. Among other things, these type of analysis could consider: i) the details of the disease spreading in the population such as mortality, infectiousness, recovery time, and possible residual immunity ii) the preparedness, measured in terms of availability of vaccines, anti-virals, hospital beds, or intensive care units, iii) the socio-economical costs induced by a major outbreak and by interventions such as travel bans, school closures etc. iv) politics and public perception of risk.

Effects of system-wide travel restrictions. Epidemic containment measures, based on limiting or constraining human mobility, are often considered in the contingency planning and always re-emerge when there are new infectious disease threats (Riley 2007). The target of these control measures are travels to/from the areas affected by an epidemic outbreak and the corresponding decrease of infected individuals reaching areas not yet affected by the epidemic. At the same time, travel restrictions have a large impact on the economy and affect the delivery of services, including medical supplies and the deployment of specialized personnel to manage the epidemic. For this reason, travel restrictions must be carefully scrutinized to trade off the costs and benefits. We introduce the parameter $p \in [10^{-5}, 1]$ that allows us to simulate policy-induced system-wide travel restrictions. In our settings, such measures are active until the disease is circulating in the system, i.e. there is at least one infected individual across all subpopulations. In the case of no travel restrictions and/or after the disease dies out, we have $p = 1$. In the case of travel restrictions ($p < 1$), we rescale travel flow so that mobility is a fraction of that in the unaffected system; i.e. $c_{ij} = p.d_{ij}$. To better understand the effect of such mitigation strategy, let us consider the classic SEIR model by setting $\beta_F = 0$. In the presence of travel restrictions and depending on the level of mixing, each subpopulation may or may not transmit the infection or contagion process to another subpopulation it is in contact with. In other words, the mobility parameter p influences the probability that exposed individuals will export the contagion process to other regions of the metapopulation network. Further, it introduces a transition between a regime in which the contagion process may invade a macroscopic fraction of the network and a regime in which it is limited to a few subpopulations. The transition is mathematically characterized by the global invasion threshold R_* (Colizza and Vespignani 2007). This is the analogue of the basic reproductive number at the subpopulations level and defines the average number of infected subpopulations generated by one infected subpopulation in a fully susceptible metapopulation system. In general, R_* is a function of the basic epidemic parameters, including R_0, and the mobility parameter p. The invasion threshold occurs at the critical value p_c for which $R_* = 1$. In some cases, p_c can be evaluated analytically (see the Methods section). In general, it can be estimated numerically by measuring the number of infected subpopulations as a function of the parameter p. Risk, as measured in terms of attack rate, is, therefore, monotonically decreasing due to increasingly restricted travel, and falls to virtually zero for values of p below the invasion threshold. Thus, from a risk perspective, the best strategy during a disease outbreak is to reduce the mobility. However,

an inspection of the profile of resilience provides a different picture. In Figure 12 we report the value of r obtained by sampling the phase space of the model $p - R_0$ for different values of the travel diffusion parameter and the epidemic reproductive number in heterogeneous metapopulation systems (a comparison between homogenous and heterogeneous networks is reported in SI). Each point of the phase space is studied by performing 100 stochastic realizations. The 3D dimensional plot in the p, R_0, r space reported in Figure 12A indicates that the overall resilience profile is characterized by a sharp drop as we approach the invasion threshold, i.e. $p \rightarrow p_c$.

Figure 12B shows that, while the risk decreases, the reduction of the diffusion rate p causes a reduction of r until the global invasion threshold, after which the resilience value rapidly increases. This effect is mainly due to the critical slowing of the epidemic spreading near the invasion threshold. Indeed, close to the threshold, the epidemic is still in a supercritical state, but it takes increasingly longer time to invade the system as the threshold is approached. This can be simply related to the divergence of the invasion doubling time T_d, which is defined as the time until the number of infected subpopulations doubles, relative to that at some other time. The doubling time is related to the subpopulation reproductive number as T_d $(R_* - 1)^{-1}$, leading to a divergence of the doubling time as the invasion threshold is approached for $R_* \rightarrow 1$. Although the absolute risk is very low, the system remains in a state of deteriorated functionality (restrictions in travels) for longer and longer times (Fraser et al. 2004). The decrease of functionality is not offset by a corresponding decrease of risk, and the minimum in resilience is attained exactly at the global invasion threshold. The comparison between the theoretical values of the invasion threshold and the computed minimum values of resilience is reported in Figure 12C.

Effects of self-initiated behavioural changes. In order to isolate the effects of behavioural changes, in this section the travel parameter is kept constant with $p = 1$. Individuals in the S^F compartment adopt travel avoidance so that β_F plays a similar role to the travel restriction as reported in Figure 13. Furthermore, inside each subpopulation, individuals in the S^F compartment reduce their contacts, thus decreasing the likelihood to become infected. Overall, the presence of self-initiated behavioral changes in a population results in a reduction of the final epidemic size. In this setting, we have explored a phase space of parameters constituted of $R_0 \in [1.01, 3]$ and $\beta_F \in [0, 20]$ (see the Methods section for the other model parameters). In Figure 13A we quantify resilience for different values
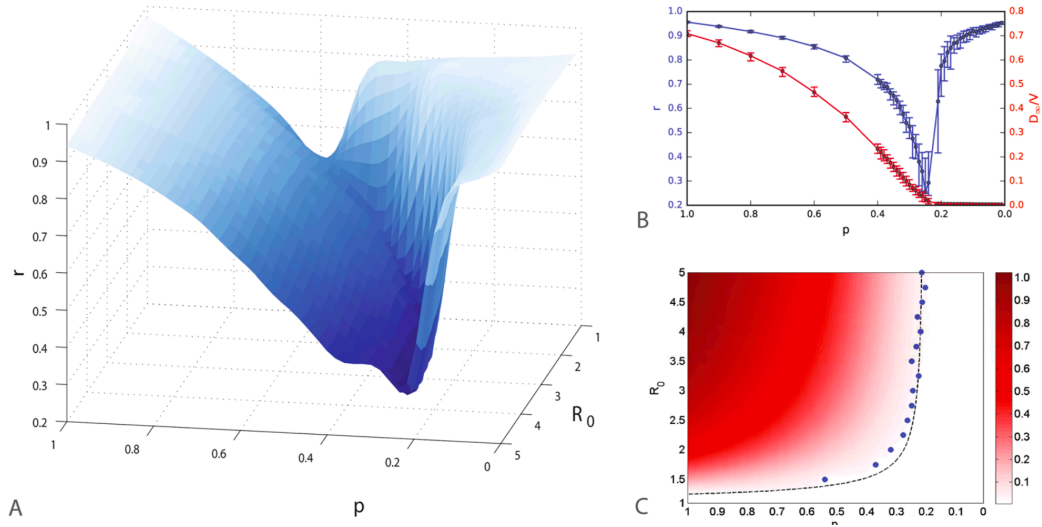
Figure 12. Resilience and final fraction of diseased populations in the hetero-geneous metapopulation system with traffic dependent diffusion rates. (A) 3D surface representing resilience in a homogeneous metapopulation system as a function of local threshold R_0 and the diffusion rate p: the minimum value of resilience separates two regions associated to values very close to the optimal case. (B) Cross-sections (blue) of the 3D plot for $R_0 = 3.5$ and its comparison with the final fraction of diseased populations (red): while the reduction of the diffusion rate p brings to a constant the fraction of diseased populations it also causes an initial decrease of resilience to a min-imum value after which it starts increasing and the system returns to its optimal conditions. (C) The map of the final fraction of diseased popula-tions D_∞/V is shown as a function of the local epidemic threshold R_0 and the travel diffusion p. We show that the minimum values of resilience (blue points) correspond to the theoretical value of the final fraction of diseased subpopulations D_∞/V at the end of the global epidemic (black line).

of the fear parameter β_F in heterogeneous metapopulation systems. The 3D dimensional plot in the β_F, R_0, r space shows a clear similarity with the travel restrictions scenario. Figure 13B shows that, while increasing β_F leads to a decrease in risk, it also induces a reduction of resilience. It is possible to observe that, even in this case, the minimum values of r are related to the invasion threshold. In Figure 3C the phase diagram of the fraction of diseased populations at the end of the simulations D_∞/V is reported in the β_F, R_0 space. This picture shows that there is a critical value of the fear transmissibility

43
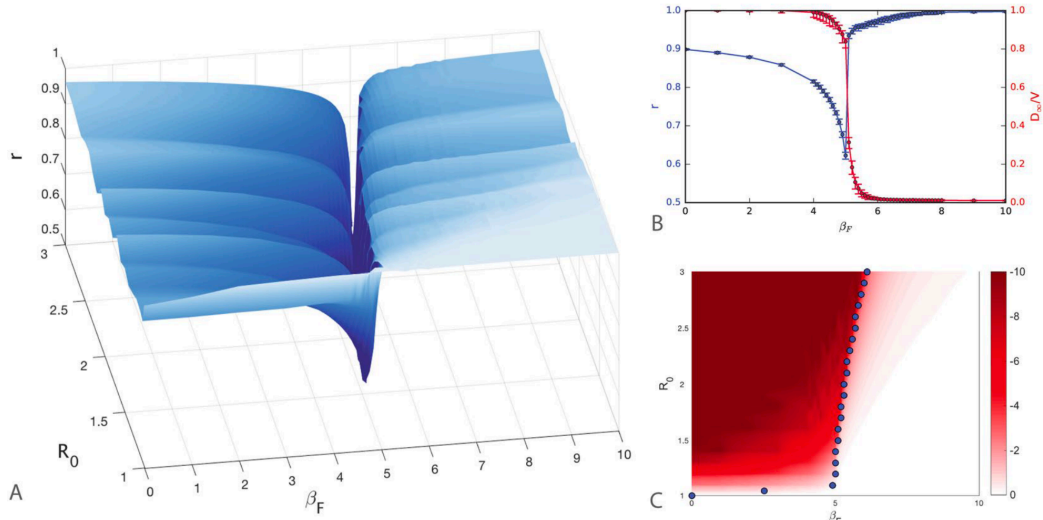
Figure 13. Resilience and diseased populations in a heterogeneous metapopulation system with individual self-dependent travel reduction. (A) 3D surface representing resilience in a heterogeneous metapopulation system as a function of local threshold R_0 and the fear parameter β_F: two areas of high values of resilience are separated with a narrow region of very low ones. (B) Comparison between resilience (blue) reported as cross-sections of the 3D plot for $R_0 = 1.3$ and the final fraction of diseased populations D_∞/V (red): while the increase of the fear transmissibility parameter β_F brings to a constant the fraction of the diseased populations it also causes an initial decrease of resilience to a minimum value after which the system bounces back to optimal conditions. (C) Even in this case the minimum values of resilience (blue points) correspond to the transition region from high to low final diseased populations. The colormap of the logarithmic of the healthy populations $(log(1 - D_\infty/V))$ is shown as a function of the local epidemic threshold R_0 and the fear parameter β_F.

parameter β_F, after which the fraction of diseased populations D_∞/V starts to decrease (i.e. $D_\infty/V < 1$). The minimum value of resilience, in this case, corresponds to the value of the fear transmissibility, after which a reduction of the fraction of diseased populations is observed. Although the approach to this critical boundary corresponds to a reduction of the infection risk, similarly to the case of travel restrictions, the measured resilience of the system decreases and attains its minimum value right at the transition point.

Effects of system-wide travel restrictions in data-driven simulations. As a further confirmation of the validity of the theoretical construct above described, we tested our results in a data-driven modelling setting. We considered the Global Epidemic and Mobility model (GLEAM) (Balcan et al. 2009b; Tizzoni et al. 2012) which integrates high resolution demographic and mobility data by using a high-definition, geographically structured metapopulation approach. The model's technical details and the algorithms underpinning the computational implementation have been extensively reported in the literature. GLEAM is a spatial, stochastic and individual-based epidemic model that divides the world population into geographic census areas, defined around transportation hubs and connected by mobility fluxes. The population of each census area is obtained by integrating data from the high-resolution population database of the 'Gridded Population of the World' project of the Socioeconomic Data and Application Center at Columbia University (SEDAC). The mobility among subpopulations is comprised of global air travel and the small-scale movement between adjacent subpopulations; i.e., the daily commuting patterns of individuals. Commuting and short-range mobility considers data from 80,000 administrative regions in 5 different continents. Here, we considered the Continental United States and simulated an SEIR contagion process, in which infected individuals do not travel. The number of infected subpopulations at the end of an outbreak and resilience as a function of the global mobility restrictions that result are shown in Figure 14. The initial conditions of the epidemic were set with 5 infected individuals in the city of New York, assuming $\beta = 0.48$, $\lambda = 0.66$ and $\mu = 0.45$. Mobility restrictions are implemented by reducing all the mobility flows connecting diseased subpopulations by a factor p, thus considering the heterogeneities of the subpopulations due to their different local mobility patterns (see SI). The control time T_C used in the calculation of r corresponds to the epidemic extinction time for the different values of the diffusion rate. As with the theory-driven model here we observe that a reduction of the travel diffusion p brings a constant reduction of diseased populations, but also reduces resilience until a critical value $p_c = 1.2x10^{-4}$. In Figure 4B we illustrate the geographical spreading of the contagion process and the reduction of traveling of each subpopulation tracked by the model in the Continental USA for values of p corresponding to three different regions of the diagram of Figure 14A. The figure clearly illustrates three regimes: i) for low travel reduction, a very severe epidemic hits all the subpopulations, but the short duration allows the system to go back to normal in a short time (high values of resilience); ii) for travel reduction close to the global invasion threshold, a small number of subpopulations are

hit but the system critical functionality is compromised for a very long time, thus, resulting in a low values of resilience; iii) travel reduction above the critical threshold allows the system to contain the epidemic at the origin with low risk and high values of resilience. It is worth remarking that in the data-driven model, the minimum value of resilience is reached for travel restrictions that correspond to a reduction of mobility of three to four orders of magnitude. This is because in modern transportation networks the global invasion threshold, as already pointed out in other studies (Cooper et al. 2006; Epstein et al. 2007; Gomes et al. 2014; Colizza and Vespignani 2008), is reached only for very severe travel restrictions that are virtually impossible to achieve. In other words, in realistic settings the feasible increase of travel restrictions appears always to decrease resilience. This calls for a careful scrutiny of the trade-off between individual's risk and resilience, as the region where both are achieved is virtually not accessible.
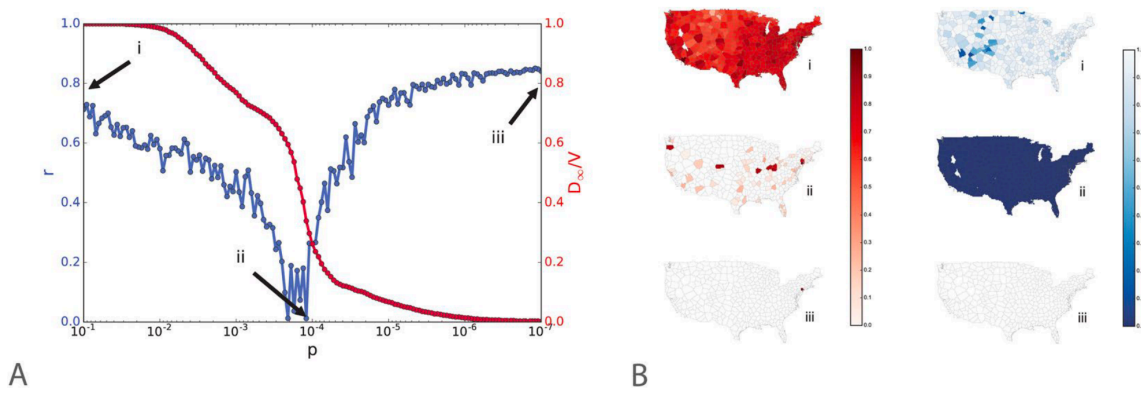
Figure 14. Resilience and epidemic size in the data-driven scenario. (A) The plot shows the difference between resilience (blue) and the final fraction of diseased populations (red) for different values of the diffusion rate p. Here, we can identify three critical regions of the system. i) diffusion rate $p = 0.1$ above the critical invasion threshold. Even if the system is characterized by sub-optimal resilience, the disease spreads all over the system. ii) the reduction of the diffusion parameter p results in a significant decrease of the number of diseased populations but also in a dramatic decrease of resilience; iii) below the critical invasion threshold resilience goes back to high values as fraction of diseased populations approaches zero. Python 2.7 (https://www.python.org/) and the Basemap library (https://pypi.python.org/pypi/basemap/1.0.7) were used to create these maps. (B) Epidemic size (red) and resilience (blue) for the different values of the diffusion parameter p corresponding to the three aforementioned regions.

2.2.3 Discussion

The realistic threat quantification is difficult and evaluation of vulnerabilities and consequences of new disease epidemics is certainly a challenge. We analyzed the impact of an infectious disease epidemic in structured populations by considering a definition of system resilience that takes into consideration not only the number of infected individuals but also

society's need for maintaining certain critical functions in space and time (Linkov et al. 2014b). In particular, we assume that the limitations and disruptions of human mobility deteriorate the system's functionality. We observe that containment measures that limit individuals' mobility are advantageous in reducing risk but may deteriorate the system's functionality for a very long time and thus correspond to low resilience. Although we have considered only two of the many dimensions encompassing the functionality of socio-technical systems (Holling 1973; Linkov et al. 2014c), we show that study of resilience allows stakeholders to measure the impact of epidemic threats and differentiate between different management alternatives. It is straightforward to envision more realistic definition of the critical functionality. The components of critical functionality could be weighted according to objective/subjective evaluation of their relevance to stakeholders. Finally, cost-benefit analyses and ethical considerations should be included in the analysis of the societal impacts of disease that could lead to long lasting effects or even death of the affected individuals. This study highlights the importance of resilience-focused analysis for selecting intervention strategies. The natural tendency to be conservative in managing potentially inflated risks associated with new and emerging epidemics can result in unnecessary burdensome and possibly ineffective actions like quarantines as well as travel bans49. The emerging field of resilience assessment and management29 and its implementation (Ganin et al. 2016; Gao et al. 2016) could thus evaluate cross-domain alternatives to identify a policy design that enhances the system's ability to (i) plan for such adverse events, (ii) absorb stress, (iii) recover, and (iv) predict and prepare for future stressors through necessary adaptation. To this end, the framework we presented can be of potential use for optimizing the policy response to a disease outbreak by balancing risk reduction with the disruption to critical functions that is associated with public health interventions.

2.2.4 Methods

Disease propagation and self-initiated behavioral changes. The metapopulation system is described by a scale-free network (SF) with a power-law degree distribution $P(k) \sim k^{(-\gamma)}$, which is generated by the configuration model (Catanzaro et al. 2005) with the minimum degree $m = 2, \gamma = 2.1$. (For the travel restriction scenario, in the SI, we report a comparison of the results between the heterogeneous networked system described above and a metapopulation system formed by a random network with Poisson degree distribution, which is generated by the Erdos-Rényi (ER) model (Erdős and Rényi 1960).) The networks

have $V = 5000$ nodes and average degree $\langle k \rangle \sim 6$, while the total number of individuals is $N = V^2 = 25.10^6$ which are distributed among the subpopulations nodes proportional to their degree distribution. At the beginning, 10 populations are selected at random and 50 individuals are set as exposed. All other individuals across the system are initially susceptible. We study a compartmental scheme that extends the basic SEIR (Anderson et al. 1992) model by considering separate behavioral classes within populations (see SI for the detailed description of the model). For this reason, we assume that individuals can spontaneously change their behavior because of the fear of the disease entering in a specific compartment S^F of susceptible individuals. In the case of travel restrictions, we set the transition rate from exposed to infected $\lambda = 0.67days^{-1}$ and recovery rate $= 0.33days^{-1}$. In the case of the behavioral model, we set the disease parameters $\lambda = 0.3days^{-1}$ and $\mu = 0.1days^{-1}$ while we consider an infection probability reduction $r_b = 0.15$, the self-reinforcement parameter $\alpha = 0.1$ and the relaxation parameter $\mu_F = 0.5$. All the presented results are averaged over 100 simulations.

Mobility process. We adopt a simplified mechanistic approach that uses a Markovian assumption for modeling migration among subpopulations; at each time step, the movement of individuals is given according to a matrix d_{ij} that expresses the probability that an individual in the subpopulation i is traveling to the subpopulation j. We assume that the diffusion rate on any given edge from a subpopulation node of degree k_i to a subpopulation node of degree k_j is proportional to k_j (Colizza and Vespignani 2008) and it is given by $d_{ij} = \omega_0(k_i k_j)^\theta / T_i$, where $T_i = \sum_j w_{ij} = \sum_j \omega_0(k_i k_j)^\theta$ represents the total flow in i, while θ and the exponent ω_0 are system specific (e.g., and $\theta = 0.5$ and in the world-wide air transportation network (Barrat et al. 2004)). In this scenario, we consider $\theta = 0.5$ and $\omega_0 = 10^{-3}$.

Global invasion threshold. For the SEIR model it is possible to explicitly calculate the average number of infected subpopulations for each infected subpopulation in a fully susceptible metapopulation system as $R_* = \hat{N}p\frac{2(R_0-1)^2}{(\mu+\lambda)R_0^2}\frac{\langle k^2 \rangle - \langle k \rangle}{\langle k \rangle^2}$ (Balcan et al. 2009b) where $\hat{N}$ represents the average number of individuals in a subpopulation. The condition $R_* = 1$ defines the invasion threshold for the system. Only for $R_* > 1$ can the epidemic spread to a large number of subpopulations. The invasion threshold readily provides an explicit condition for the critical mobility p_c, below which an epidemic cannot invade the metapopulation system, yielding the equation $p_c = \frac{1}{\hat{N}}\frac{\langle k \rangle^2}{\langle k^2 \rangle - \langle k \rangle}\frac{(\mu+\lambda)R_0^2}{2(R_0-1)^2}$ (Colizza and

Vespignani 2008).

2.3 Resilience and efficiency in transportation networks

Urban transportation systems are vulnerable to congestion, accidents, weather, special event, and other costly delays. Whereas typical policy responses prioritize reduction of delays under normal conditions to improve the efficiency of urban road systems, analytic support for investments that improve resilience (defined as system recovery from additional disruptions) are still scarce. In this effort, we represent paved roads as a transportation network by mapping intersections to nodes and road segments between the intersections to links. We built road networks for 40 of the urban areas defined by the U.S. Census Bureau. We developed and calibrated a model to evaluate traffic delays using link loads. The loads may be regarded as traffic-based centrality measures, estimating the number of individuals using corresponding road segments. Efficiency was estimated as the average annual delay per peak period auto commuter and modeled results were found to be close to observed data, with the notable exception of New York City. Resilience was estimated as the change in efficiency resulting from roadway disruptions, and was found to vary between cities, with increased delays due to a 5% random loss of road linkages ranging from 9.5% in Los Angeles to 56.0% in San Francisco. The results demonstrate that many urban road systems that operate inefficiently under normal conditions are nevertheless resilient to disruption, whereas some more efficient cities are more fragile. The implication is that resilience, not just efficiency, should be considered explicitly in roadway project selection and justify investment opportunities related to disaster and other disruptions.

2.3.1 Introduction

Existing roadway design standards emphasize the efficient movement of vehicles through a transportation network (Samuelsson and Tilanus 1997; Beverly 2010; Hoogendoorn et al. 2014; Sami et al. 2013). Efficiency in this context may include identification of the shortest or fastest route (Samuelsson and Tilanus 1997; Rodrigues et al. 2008; Chang and Nojima 2001; D'Este et al. 1999), or the route that minimizes congestion (Yan et al. 2006). It is the primary criterion on which road networks are modeled and design alternatives considered (Chang and Nojima 2001; D'Este et al. 1999; Allen et al. 1993; Yamashita et al. 2004). The Texas A&M Transportation Institute defines and reports traffic delay

in urban areas as the annual delay per auto commuter (Schrank et al. 2015). Other studies define efficiency as delay for the individual driver in terms of time spent moving or stopped (D'Este et al. 1999), or mean travel time between all origin-destination pairs in the network (Allen et al. 1993). However, as the experience of any motorist in large American cities can attest, conditions beyond the scope of the roadway design, including congestion, accidents, bad weather, construction, and special events (e.g., a marathon race), can cause costly delays and frustrating inefficiencies that result in fuel waste, infrastructure deterioration, and increased pollution (Çolak et al. 2016; Turnbull 2016). Evaluating road networks based only on efficiency under normal operating conditions results in little to no information about how the system performs under suboptimal or disrupted conditions. Infrastructure systems that exhibit adaptive response to stress are typically characterized as resilient (Holling 1996, 1973; Linkov et al. 2014a; Ganin et al. 2016; Park et al. 2011; Flynn and Burke 2012; Hollnagel et al. 2006b; Alderson et al. 2015). Given the essential role of transportation in emergency response, provision of essential services, and economic well-being, the resilience of roadway networks has received increasing policy attention. Nonetheless, scholars have yet to converge on a shared understanding of resilience suitable to guide design, operation, and reconstruction of roadway networks. Although resilience in infrastructure systems is characterized as a multidimensional concept (Woods 2015; Seager et al. 2017b), in many engineering and civil infrastructure implementations, resilience is defined as the ability of a system to prepare for, absorb, recover from, and adapt to disturbances (Linkov et al. 2014a). Specific to transportation, resilience has been defined as "the ability of the system to maintain its demonstrated level of service or to restore itself to that level of service in a specified timeframe" (Freckleton et al. 2012). Others describe transportation resilience as simply the ability of a system to minimize operational loss (Pant 2012), or will use the term synonymously with robustness, redundancy, reliability, or vulnerability (King and Shalaby 2016; Li et al. 2016; Kermanshah and Derrible 2016). Current efforts in transportation resilience research have focused on framework development and quantification methods. These efforts include the specification of resilience indicators, such as total traffic delay (Freckleton et al. 2012; Cox et al. 2011), economic loss (Cox et al. 2011), post-disaster maximum flow (Carey and Hendrickson 1984), and autonomous system components (Murray-Tuite 2006). Practical concerns with this type of resilience evaluation are that it relies on uncertain performance data, and often omits indicators that are unquantifiable (Flynn and Burke 2012). Other resilience approaches apply traffic

51

network modeling to identify locations for critical buildings (e.g. hospitals and fire stations) (Melkote and Daskin 2001), minimize trip distance for individual passengers (Thiagarajan et al. 2009), and minimize travel time across the system (Çolak et al. 2016). One drawback of existing network resilience methods is that they are data intensive, often requiring limited information about resources for unusual road system repair (King and Shalaby 2016; Kermanshah and Derrible 2016) or network behavior following a disruptive event (Simonovic 2016). Moreover, existing resilience quantification approaches lack calibration and testing across a range of transportation systems. Because many disruptive events, and their associated consequences, are difficult to predict, resilient road systems must be characterized and evaluated by the capacity to adapt to a variety of different stress scenarios. Partly because of these obstacles, joint consideration of efficiency and resilience has yet to be implemented for transportation networks. In this paper, we study the interconnections between resilience and efficiency among road transportation networks in 40 major U.S. cities. We develop an urban roadway efficiency model, calibrate it based on the observed data (Schrank et al. 2015) of annual delay per a peak period auto commuter and apply the model to calculate efficiency in 40 cities. Then, we model traffic response to random roadway disruptions and recalculate expected delays to determine the sensitivity of each city to loss of roadway linkages. The results may reveal important considerations for assessing proposals for improvement of roadway infrastructure that maintain efficiency under stress conditions.

2.3.2 Methods

To develop the urban roadway efficiency model, we first define the urban areas boundaries and construct the urban road networks before we evaluate the population density within cities utilizing the Census Bureau datasets (U.S. Census Bureau 2012a, 2010) and Open Street Map datasets (OpenStreetMap contributors 2017). We rely on this data to assess commuter patterns, which we use to measure efficiency and resilience of road networks. Alternative approaches to transportation have been offered and include those based on percolation theory and cascading failures (Berezin et al. 2015; Majdandzic et al. 2016; Zhao et al. 2016), human mobility patterns studies (González et al. 2008; Cho et al. 2011; Calabrese et al. 2013), queueing (Gross et al. 2008; Jiang and Adeli 2003), and the use of historical data to predict traffic. We review these approaches in Supplementary Materials and note that the main benefit of our model is that it relies solely on readily available public

data, rather than on particular data sets which may or may not be practical to obtain for any particular region. The model's algorithmic simplicity allows us to consider spatial topologies of cities in high resolution including tens of thousands of nodes and links. We do not create a more accurate transportation model than the existing ones, but we are able to obtain measurable characteristics of transportation systems (average delays) using our model.

Geospatial Boundaries and Population Density. To define geospatial boundaries for the transportation infrastructure networks we use the U.S. Census Bureau geospatial dataset (U.S. Census Bureau 2012a) for urban areas – densely developed residential, commercial and other nonresidential areas (U.S. Census Bureau 2012b). We approximate the exact urban area polygon with a simplified manually drawn one (Figure 15A) and include all roadways within 40 km (25 miles) of it in the network. For each of the links we calculate its length based on the polyline defining the link, and assign a number of lanes m and the free-flow speed (FFS) (see Supplementary Materials). We next estimate population in vicinity of each intersection i using the Census Tract data (U.S. Census Bureau 2010). To this end, we split the map into Voronoi cells centered at intersections and then estimate the population of each cell N_i as

$$N_i = \sum_t N_t \frac{Area(P_t \cap P_i)}{Area(P_t)} \tag{18}$$

Above N_t is the population of Census Tract t, and P_i, P_t are the polygons of the cell and the tract respectively (Figure 15B).
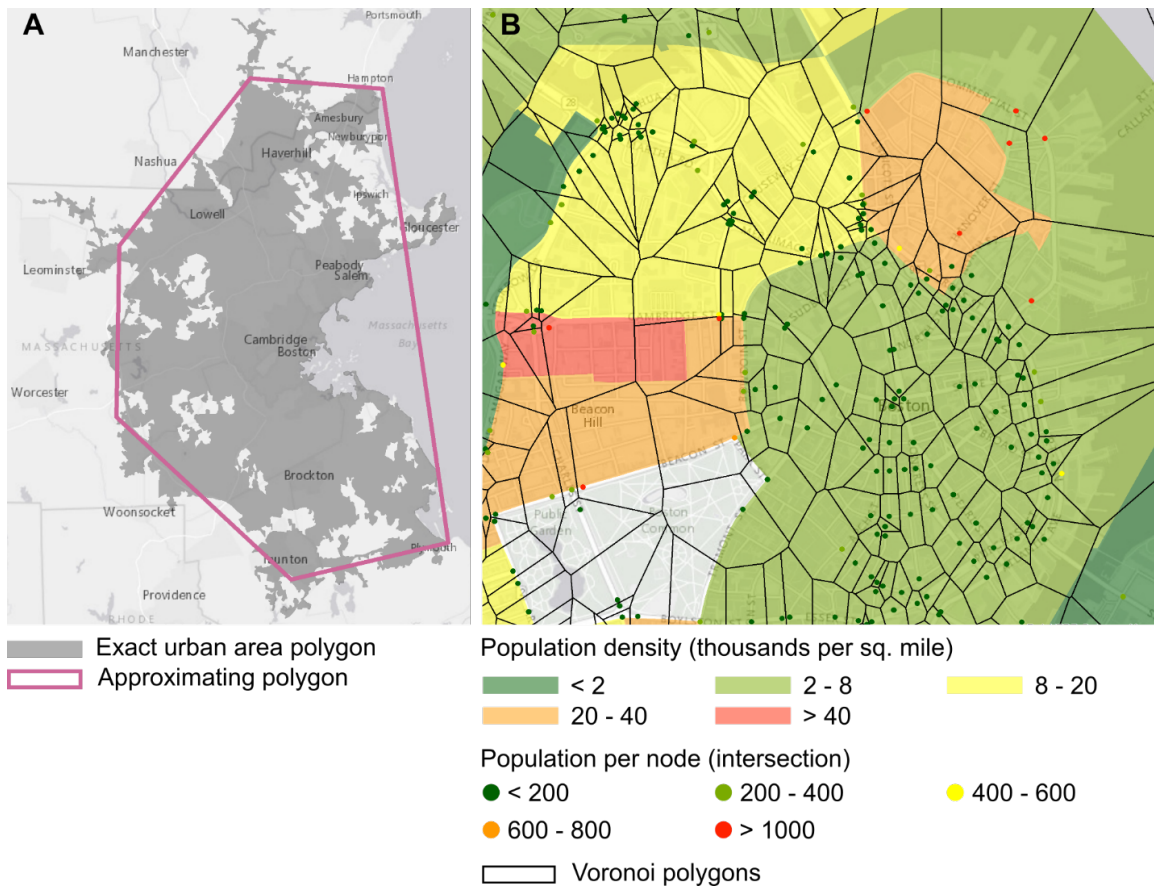
Figure 15. Definition of urban areas and assignment of nodes' population. A. Boston, MA-NH-RI urban area as defined by the U.S. Census Bureau shapefiles (gray background). To simplify the model and the algorithms calculating the distance from network nodes to the city boundary, we approximate each of the urban areas shapefiles with a coarse manually drawn polygon (pink outline). B. Assignment of the number of people to departing from each of the network nodes. Population distribution (color polygons, red corresponds to higher population density), Voronoi polygons (black outline) and network nodes (dots) in downtown Boston.

Transportation Model. We build on the gravity model for trip distribution assignment. The gravity model (Casey 1955) is a classical model for trip distribution assignment and is extensively adopted in most Metropolitan Planning and statewide travel demand models in the U.S. (Sabyasachee et al. 2013; Cascetta et al. 2007; J. de D. Ortúzar and Willumsen 2011; Transportation Research Board 2007). Other trip distribution models include, for example, destination choice models (Fotheringham 1983; Woodside et al. 1989). However,

these models are not as widely used in large-scale, as the detailed data required by these models is frequently unavailable (Sabyasachee et al. 2013). We assume that (i) the flow of commuters from origin region o to destination region d is proportional to population at the destination Nd, and that (ii) the flow of commuters depends on the distance x_{od} between the origin and destination and is given by a distance factor $P(x_{od})$. Using these assumptions, we can estimate the fraction of individuals commuting from region o to destination region d, f_{od}, as

$$f_{od} = \frac{N_d P(x_{od})}{\sum_k N_k P(x_{ok})} \tag{19}$$

Then, the commuter flow from origin region o to destination region d is

$$F_{od} = N_o f_{od} \tag{20}$$

While individual driving habits may vary (Lima et al. 2016), we assume that all drivers tend to optimize their commute paths such that their travel time is minimized. This assumption allows us to calculate commute paths for every origin destination-pair using inferred free-flow speeds. To calculate commuter flows between all pairs of intersections we estimate distances x_{od} as the distance of the shortest time path from o to d. Further, in place of the distance factor $P(x_{od})$ we use the distribution of trip lengths from the U.S. Federal Highway Administration National Household Travel Survey (NHTS) (U.S. Department of Transportation 2009; Haaren 2012), which we approximate with the exponential function (Figure 16A).
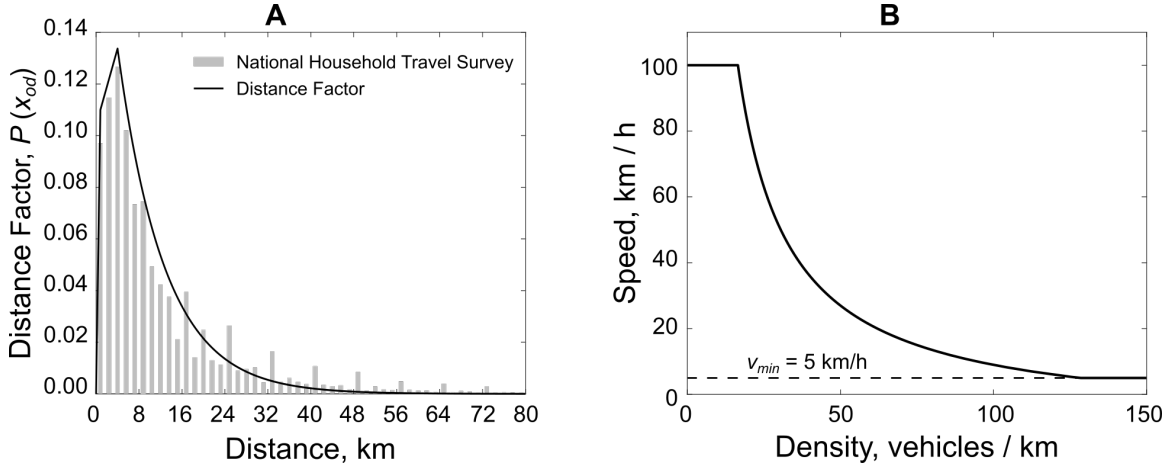
Figure 16. Model details. A. Distance factor $P(x_{od})$ (Eq. 19) of trips given the distance between nodes (solid line) and the statistical data (bars). B. Dependency of speed on density for $V = 100$ km/h.

Next, we define the commuter load on each road segment as

$$L_{ij} = \sum_{o,d} F_{od}\theta_{od}(ij) \qquad (21)$$

where $\theta_{od}(ij)$ is a binary variable equal 0 when link ij is not on the shortest path connecting nodes o and d, and 1 otherwise. Note that in Eq. 21 we only consider origins that are not farther than 30 km from the urban area boundary polygon. The nodes farther than 30 km from the boundary are only used as destinations to evaluate the fraction of commuters not going towards the urban area (Eq. 19). Because most commuters travel during peak periods, commuter loads L_{ij} can be regarded as traffic-based centrality measures estimating the number of individuals using corresponding road segments. Then, the cumulative time lost by all commuters

$$\Delta T = \beta \sum_{ij \in E} L_{ij} \left(\frac{(l_{ij} + l_0)}{v_{ij}} - \frac{(l_{ij} + l_0)}{V_{ij}} \right) \qquad (22)$$

where V_{ij} and v_{ij} are respectively the free-flow and the actual traffic speeds along the ij road segment, l_{ij} is its length, l_0 is the length correction due to traffic signals, and β is the proportionality coefficient same for all urban areas. The summation in Eq. 22 includes only links, whose origins and destinations are within the boundary polygon. A similar

56

equation was obtained for the moving delay in (Jiang and Adeli 2003) where the authors looked at the delay induced from road repairs. The actual traffic speed v_{ij} depends on many factors including the speed limit, the number of drivers on the road, and road conditions. While there exist a number of approaches to estimate actual traffic speed (Greenshields et al. 1934; Smulders 1990), we have chosen to use the Daganzo model (Daganzo 1994) to derive the traffic speed:

$$v_i j = \alpha \frac{l_{ij} m_{ij}}{L_i j} - v_{veh}, \; subject \, to \, v_{ij} \in [v_{min}, V_{ij}] \tag{23}$$

where v_{min} is the minimum speed in the traffic, v_{veh} is the correction for the finite size of the car and α is the proportionality coefficient (Figure 16B).

Efficiency and Resilience Metrics. In order to summarize results of the model, we provide working definitions for metrics of resilience and efficiency as follows. We measure efficiency as the average annual delay per peak period auto commuter. In practice, lower delay mean higher efficiency. There are multiple ways to map from delays to efficiency, such as taking the inverse values of delays, taking negative values of delays, etc. To avoid ambiguity and facilitate the interpretation of results we keep using the delays themselves to quantify the transportation efficiency of urban areas. We operationalize resilience through the change in traffic delays relative to stress, modeled as loss or impairment of roadway linkages. Looking at resilience from the network science perspective, we focus on topological features of cities, rather than on recovery resources available. Sterbenz et al. (2010) evaluate a network's resilience as a range of operational conditions for which it stays in the acceptable service region and highlight that remediation mechanisms drive the operational state towards improvement. We are studying how availability of alternate routes helps remediate the consequences of the initial disruption to the network. In the traffic context, the immediate impact of a given physical disruption (and the time for it to unfold) in terms of closing lanes or reducing speed limits on affected roads will not vary much from network to network, although the number and type of such disruptions will. Likewise, the speed of restoring full functionality (through action in the physical domain) is not so much dependent on the road network as it is on the nature of the disruption (snow vs. earthquake vs. flood) and the resources the city allocates to such repair, and the level of functionality such repairs achieve ought to be the full pre-disruption functionality, i.e., eventually all roads can be fully

cleared or restored. However, the immediate loss of function for a given traffic flow can very quickly be partially recovered after a disruption by action in the information domain, namely rerouting of traffic. From the new steady state at that level of functionality, full functionality is gradually restored. Thus, our model proxies for resilience and is calibrated against the data that proxies for efficiency. At the same time, we note that to fully capture resilience characteristics of a transportation system, it is required to analyze recovery resources available and the effectiveness of coordination between the relevant authorities. Lower additional delay corresponds to higher resilience, but using the same reasoning we had for efficiency, we quantify resilience through additional delays.

2.3.3 Results

Efficiency. Taken together, our traffic model has three parameters, proportionality coefficient α, minimum speed v_{min}, and finite vehicle size correction v_{veh} and is summarized in Eqs. 22 and 23. Given parameter values of the model, one can estimate the total delay incurred by all commuters in any given suburban area or, equivalently, the average delay per commuter. We take v_{veh} = 9 km/h and v_{min} = 5 km/h and calibrate the model to determine the value of α to match the real data on the annual average delay per a peak period auto commuter provided by the Urban Mobility Scorecard (Schrank et al. 2015).

We divide the 40 urban areas into two equally-sized groups for model calibration and validation respectively. We have found that for the 20 urban areas used for calibration, the R-squared coefficient took values in the range [-0.01; 0.83] (Figure 17). This allows us to set model parameters α and β (see Methods) as follows: $\alpha = 4.30x10^4 hour^{-1}, \beta = 10.59$. These values correspond to the Pearson coefficient of 0.91 ($p = 2.17x10^{-8}$). To validate the model, we estimate travel delays in 20 different urban areas. As seen from Figure 17, the estimated travel delays are significantly correlated ($R = 0.63, p = 3.00x10^{-3}$) with actual delay times (Schrank et al. 2015), validating the transportation model. Figure 18 is a Google Maps representation of real and modeled results for Los Angeles (LA) and San Francisco (SF). Road conditions under real, average traffic patterns at 8 am provided by Google Maps are in Figure 18A, 18D. Modeled conditions are given for comparison in Figure 18B and 18E. Finally, Figure 18C and 18F show the new, modeled traffic patterns that result from redistribution of travel in response to a disruption of 5% of the links.

58
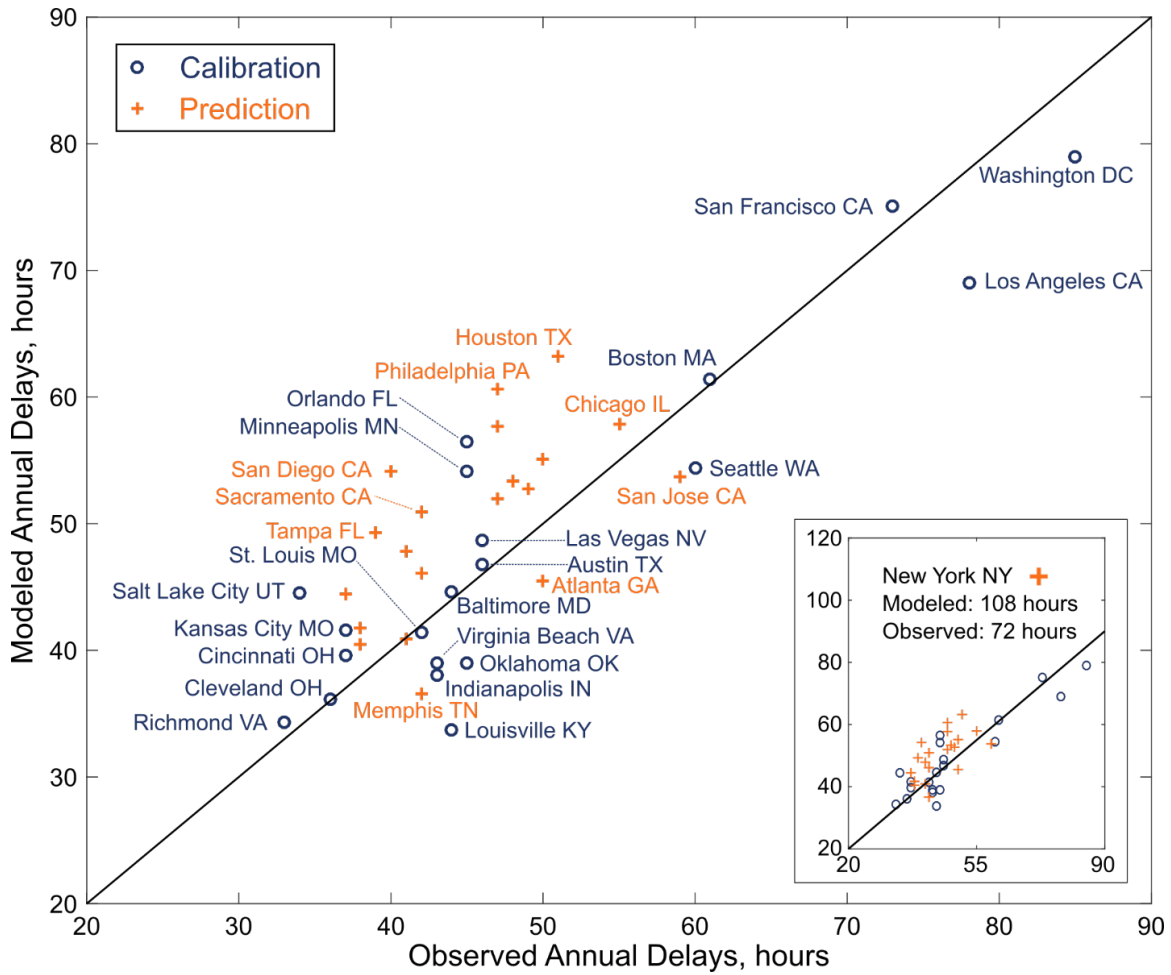
Figure 17. Modeled and observed delays in 40 urban areas. Pearson correlation coefficient and p-value between observed and modeled delays are $(0.91, 2.17x10^{-8})$ for the 20 cities used to calibrate the model and $(0.63, 3.00x10^{-3})$ for the 20 cities used to validate the model. Observed delays were taken from the Texas A&M Transportation Institute Urban Mobility Scorecard (Schrank et al. 2015).
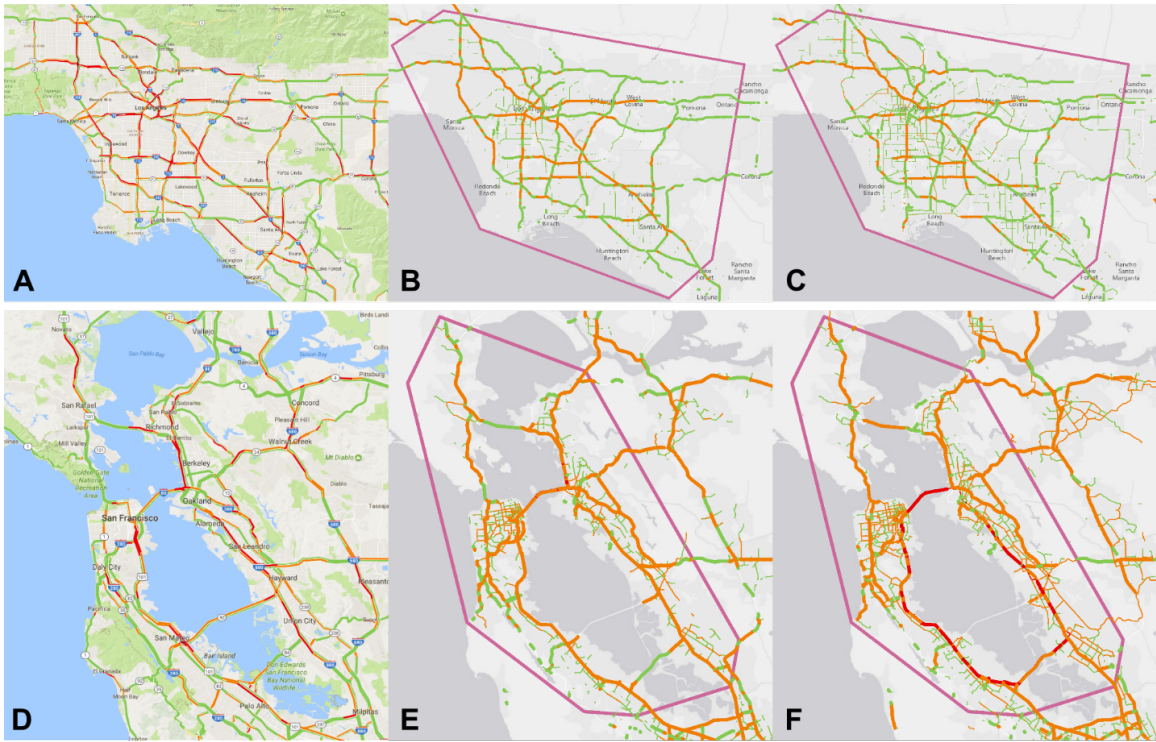
A, D Google Map typical traffic at 8am

B, C, E, F Modeled delay per km (min): ━━ < 1.2   ━━ 1.2 - 12   ━━ 12 - 24   ━━ > 24
                                            ━━ Highways   ─── Other roads

▭ Approximating urban area boundary polygon

Figure 18. Traffic distributions. Typical congestion at 8 am for Los Angeles (top) and San Francisco (bottom) as given by Google Maps (A and D), modeled with no disruptions (B and E), and modeled with a 5% link disruption (C and F). Notably, in Los Angeles the disruption results in traffic redistribution to smaller roads, while in San Francisco it results in increased congestion along the major highways.

Resilience. Our approach to model stress is inspired by the percolation theory. For every independent simulation of stress, we select a finite fraction of affected road segments r at random, with the probability of failure proportional to segment length. We collect statistics for 20 realizations of the percolation. On failed segments, free-flow speeds are reduced to 1 km/h (representing near total loss), and loads L and traffic delays are then recalculated using the updated free-flow speeds. Low stress scenarios ($r < 0.1$) might be caused by accidents or construction. Larger disruptions might occur during power failures that disrupt traffic signals, or severe flooding that makes many roadways nearly impassable. Finally,

widespread stress might be caused by snow, ice, or dust storms that impact nearly the entire roadway system. Figure 19 displays the analysis of delay times in six representative urban areas for the full spectrum of adverse event severities, $r \in [0; 1]$. Some routes within a single urban area experience longer delays than others. The inset to Figure 19 shows the delay distribution for both LA, which is narrowly clustered, and Boston, where greater variability between roadways is evident. Traffic delay times grow rapidly as r increases and reach saturation (all routes moving at 1 km/h) as r approaches 1. We determine the most resilient urban transportation network to be Salt Lake City UT, while the least resilient among the 40 metropolitans is shown to be Washington DC. Figure 20 shows both the efficiency (in blue) and resilience response (additional delays due to 5% link disruption, in orange) for of the 40 urban areas modeled. Some cities with high efficiency under normal operating conditions (i.e., low delays) nevertheless exhibit low resilience (i.e., a sharp increase in traffic delays) under stress. Virginia Beach VA, Providence RI and Jacksonville FL all fall into this category of urban areas in which traffic operates well under ordinary circumstances, but rapidly become snarled under mild stress. On the other hand, Los Angeles (LA) is notorious for traffic delays under all conditions – yet minor stress levels result in little degradation of efficiency. By contrast, normal traffic delays in San Francisco (SF) are comparable to LA, but mild stress in SF results in large increases in additional delays. These examples indicate that resilience (i.e., additional delay response to stress) is independent of normal operating efficiency.
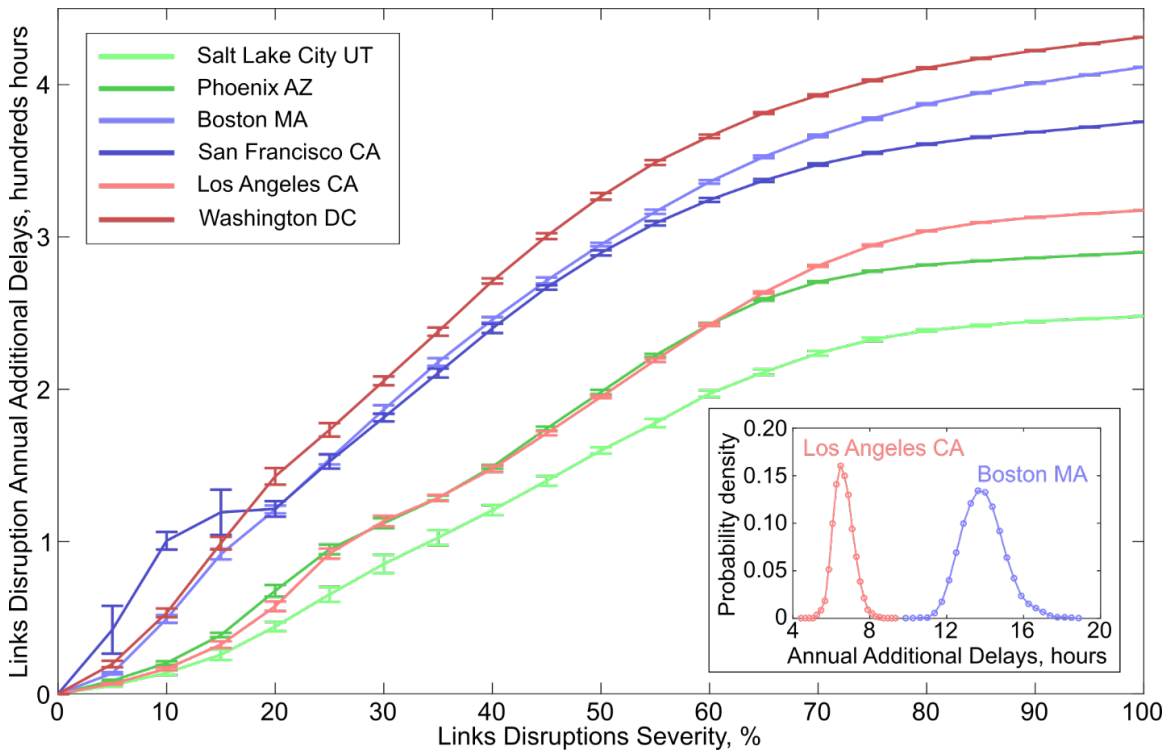
Figure 19. Dependency of the additional delay on the severity of the links disruption for six representative urban areas. Error bars show mean values ± standard deviation. The inset shows distribution densities for two selected urban areas for 1,000 realizations of 5% disruption. Note, San Francisco's unique topology makes it susceptible to failures of a small number discrete roadways, and this produces an anomalous impact at 5% - 15% disruption.
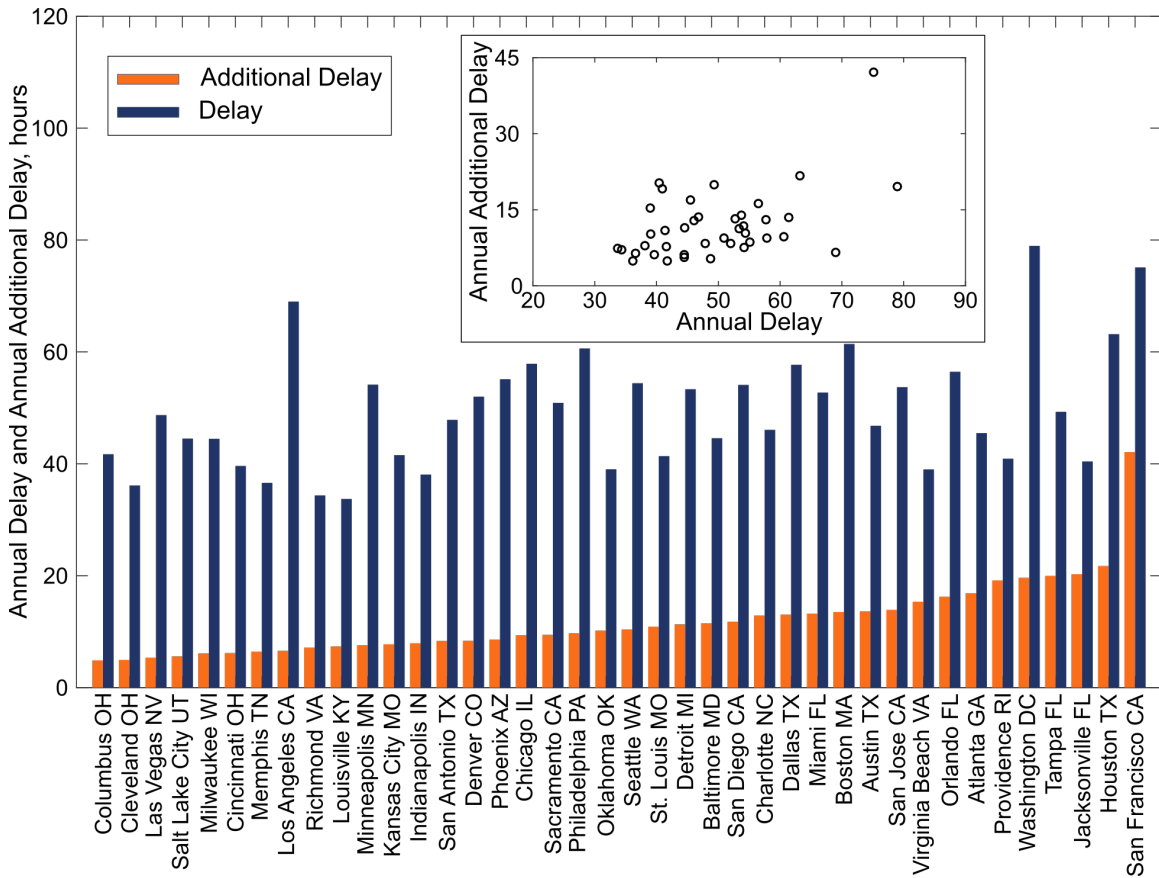
Figure 20. Comparison of resilience and efficiency metrics. Annual impact of 5% disruption (Additional Delay) has low correlation with normal annual delay per a peak period auto commuter (Delay): Pearson $R = 0.49, p = 1.18 x 10^{-3}$.

2.3.4 Discussion

The disturbances affecting the road infrastructure are often complex, and their impact on the structure and function of roadway systems may be unknown (Kermanshah and Derrible 2016; Murray-Tuite 2006). These disturbances might be natural and irregular, such as distributed road closures caused by an earthquake or homogenous vehicle slowing down due to a snowstorm. The disturbances might also be anthropogenic and intentional, such as a street fair or marathon race. Whatever the disturbance, the results of this analysis allow several meaningful inferences to be made that may have important implications for highway transportation policy. The first is that resilience and efficiency represent different aspects

63

related to the nature of the transportation systems, they are not correlated and should be considered jointly as complementary characteristics of roadway networks. Second, there are characteristic differences in the resilience of different urban areas, and these differences are persistent at mild, medium, or widespread levels of stress (Figure 19). Except for San Francisco CA, which is the most fragile of all cites represented in Figure 19 at stress levels $r < 20\%$ but then surpassed by Boston MA and Washington DC, the rank ordering of urban area resilience is insensitive to stress level. That is, cities that exhibit relatively low resilience under mild stress are the same cites that exhibit low levels of resilience (relative to peers) under widespread roadway impairment. This suggests that the characteristics that impart resilience (such as availability or alternate routes through redundancy of links) are protective against both the intermittent outages caused by occasional car crashes as they are to snow and ice storms. For cities without resilience, a widespread hazard such as snow may lead to a cascade of conditions (e.g., crashes) that rapidly deteriorate into gridlock. This was exactly the case for Washington DC 20 Jan 2016 under only 1" of snow (Richardson 2016), and for Atlanta GA two years earlier, which experienced 2" of snow in the middle of the day that resulted in traffic jams that took days to disentangle (Sen 2014). While popular explanations of these traffic catastrophes focus on the failure of roadway managers to prepare plows and emergency response equipment, Figure 19 suggests that cities with similar climates (Memphis TN, Richmond VA) are less likely to be impacted, regardless of the availability of plow or sand trucks. The third inference follows from Figure 20, which suggests that urban areas that make capital investments to reduce traffic delays under normal operating conditions may nevertheless be vulnerable to traffic delays under mild stress conditions. Because such stressors are inevitable, whether from crashes, construction, special events, extreme weather, equipment malfunctions, or even deliberate attack, investment strategies that prioritize reduction of normal operating delays may have the unintended consequence of exacerbating tail risks – i.e., the risk of worse catastrophe under unlikely, but possible conditions. Finally, the exceptional positon of New York City in Figure 17 calls attention to the fact that substitutes for roadway transportation are available in many cities, and have an important role to play in relieving traffic congestion. According to the Texas A&M Institute (Renn 2011; Lomax et al. 2011), public transit reduces delays per peak hours auto commuter in the New York urban area by 63 hours, in Chicago by 23 hours, and by less than 20 hours in other urban areas. Because our model considers only roadway transit, and New York City contains a myriad of nonroad-based options to

avoid roadway congestion, it is unlikely that our model can provide informative results for the New York urban area. Although interest has increased in policies that enhance roadway resilience, few analytic tools are available to guide new investments in achieving resilience goals. It is widely understood that roadway infrastructure is expensive, both in acquiring land for rights-of-way and in construction of improvements, and thus decisions regarding alignment, crossing, and access made over a period of decades may have long-lasting consequences that are observable in traffic data today. Consequently, urban areas exhibit different unintentional traffic characteristics, including delays under normal and random stress conditions. Investments motivated exclusively by expected efficiencies under normal operating conditions are unreliable safeguards against loss of efficiency under stress conditions. Therefore, new analytic tools are required that allow designers to assess the adaptive capacity of roadway infrastructure, and assess the potential of new investments to provide enhanced resilience. The adaptive network-based model described herein is one such approach.

2.4   Lack of Resilience in Transportation Networks: Economic Implications

Disruptions to transportation networks are inevitable. When road networks are not resilient, or in other words, do not recover rapidly from disruptions, unpredictable events can cause significant delays that may be disproportionately greater than the extent of the disruption. Enhancing transportation system resilience can help mitigate the consequences of disruptions; however, required investments are difficult to justify given the low probability of such events. This paper calculates economic implications of unmitigated random disruptions in urban road systems. We modeled delays in transportation networks and demonstrated how resilience can be integrated into macroeconomic modeling via the transportation planning model, REMI TranSight. The model was applied to 10 cities in the United States to forecast the impact of disruptions on gross domestic product (GDP). Different disruption scenarios were modeled and the magnitude of disruption was used to calculate additional delays in transportation networks, which were then integrated into the TranSight model. The results were compared to a baseline case, where economic impact was assumed to be proportional to the magnitude of disruptions. Results show that losses in GDP were far more pronounced in the case scenario as compared to the baseline. The losses tended to be higher in wealthier and more economically productive cities. The economic output tends to rebound one to

two years after a disruptive event. We conclude that different topology in transportation networks in different cities requires explicit consideration and quantification of resilience to support investment decisions designed to improve transportation networks in cities.

2.4.1 Introduction

Currently, most mandated development-related transportation planning is intended to prepare for frequently occurring and observable disruptions, while unpredictable events that have not yet occurred attract less attention. The current norm for improving transportation networks and remedying the economic impact of delays is undertaken through management of the network for specific threats to reduce the travel time and improve efficiency motivated by the goal of reducing delay (Belenky 2011). Current practice is to evaluate road performance with Level of Service (LoS), or a similar measure of efficiency, during the worst traffic of an average day and when the whole road network is running as expected (Transportation Research Board 2010). This emphasis on travel time and the monetary value of its duration allows prospective projects to enter the realm of cost-benefit analysis but carries the shortcoming of representing only observable and predictable events. Recent experiences clearly show that a focus on resilience is necessary. The economic impacts of resilience (or lack thereof) in transportation networks have not been studied, but will ultimately be essential to justify future investment. The National Academy of Sciences (2012) defines system resilience as the capacity to be prepared for an event, absorb its shock, and then recover and subsequently adapt following the event. Lounis and McAllister (2016) quantify resilience as the change in a system's performative function and quality over the course of either or both an extreme event or a gradual accumulation of internal or external stress. Here, and in previous publications (Ganin et al. 2017, 2019), we consider transportation resilience to be a property of the network that allows for fast recovery of travel to expected speeds, despite disrupted roads and intersections in the network. Cumulative delays associated with disruptions can be considered a metric of network resilience.

The transportation field recognizes that variances in expected travel times have a cost, even when not incurred, as they need to be planned for by travelers. This is known as the Value of Reliability (VOR) (U.S. Department of Transportation 2016). While VOR is a well-researched field and methodologies to estimate it do exist (Fosgerau and Karlstr"om 2010; Lam and Small 2001), no standardized method has yet been adopted in the U.S.

(U.S. Department of Transportation 2016). The cost of unreliable transportation may not be consistent across sectors as some workers have more flexibility in where and when they work than others. These costs may also fluctuate with the amplitude of the delay, with small delays potentially being negligible, though more research is needed (Fosgerau et al. 2007; Mackie et al. 2003; U.S. Department of Transportation 2016). A more standard metric used in planning is Value of Time (VOT), which is generally calculated on the assumption that variance in travel time from one scenario to another is certain. VOT is normally linearly calculated based on wage rates (U.S. Department of Transportation 2016). Both of these methods fail to capture the costs of unpredictable delays, which cannot be accounted for in schedules and may have different costs than VOR or VOT.

The realm of possibilities between average traffic conditions and unpredictable disruptions, such as natural disasters, is comprised of events that can have significant disruptions to mobility, and yet are outside the normal realm of planning and analysis. These events may be unpredictable in both space and time and unknown in nature, to the point of being random. Therefore, these unpredictable events should be the focus of resilience inquiries here and elsewhere (Ganin et al. 2019, 2017). While mitigating the risk of disruption (i.e., strengthening key nodes and links in transportation networks) is appropriate for specific hazards, events that are highly uncertain in space and time challenge our ability to characterize vulnerability to them and implement effective risk mitigation measures. Cost concerns ensure that completely minimizing physical risk at one location may inherently limit our ability to reduce risk elsewhere. Similarly, hardening a transportation system against the risk of new types of disruptions, such as cyber-attacks on Intelligent Transportation Systems (ITS), is difficult because the potential range of risk is too vast, or poorly understood, to be effectively predicted (Ganin et al. 2019). Network-wide management is therefore more appropriate than location-specific solutions, where the objective is to keep people and goods flowing through the network in spite of disrupted parts of the network. An apt measure of a network's performance with respect to that objective is resilience. The capabilities and characteristics that should be fostered to deliver system resilience in the face of disruptive events are currently studied in many disciplines, including transportation (e.g. Cox et al. 2011).

Across various contexts, there is a growing recognition that lack of resilience can have grave socioeconomic consequences, especially in the context of damaged interconnected

infrastructure (Florin and Linkov 2016). Such is the finding of a recent World Bank report on infrastructure, as a key enabler of economies and the macroeconomic impacts, not being resilient (Hallegatte et al. 2019). Presidential Policy Directive 21 - Critical Infrastructure Security and Resilience (The White House 2013) formalized the call to enhance the nation's critical infrastructure functioning and resilience by recognizing the importance of operable critical infrastructure, including transportation systems. This resilience-related Policy Directive focused on 16 sectors that were considered vital to national economic security, public health, and safety ("Critical Infrastructure Sectors," 2019). The current state of practice in transportation planning lends itself to the conclusion that advances in resilience research need to be integrated into planning norms to help account for uncertain events and emerging risks. Quantitative resilience modeling results can be used in tandem with efficiency-driven modeling efforts to conduct tradeoffs among multiple objectives in transportation network investment.

The extent of socioeconomic impacts due to disrupted infrastructure is a currently researched field, yet mostly limited to forecasting or assessing impacts of specific disruptive events. For example, Ham et al. (2005) assessed the anticipated economic implications if an earthquake were to occur in the New Madrid Seismic Zone in the U.S. Midwest. They concluded that the ensuing disruption to U.S. commodity flow could pose a significant threat to economic stability and recovery at the regional, national, and international scales. Ham et al. (2005) define transportation network resilience as referring to the adaptability of commodity flow such that goods can be transported via multiple modes. Similarly, Tatano and Tsuchiya (2008) developed a spatial computable general equilibrium (CGE) model to estimate economic losses (e.g., changes in the cost of travel time) attributed to earthquake disruption of freight and passenger transportation flow. They used the 2004 Niigata-Chuetsu Japan earthquake as a case study, and regional economic losses were measured as a function of inter- and intra-regional trade (Tatano and Tsuchiya 2008). Pelling et al. (2002) discussed how the 1995 Kobe, Japan earthquake increased transportation costs in the region by over 50% and increased the cost of goods in the region by 10%. Internationally, the disablement of the Kobe Port halted the import and export of goods. Pelling et al. (2002) suggest that disasters possess an "inflationary potential" due to their capacity to affect the "production, distribution, marketing, and consumption" functions of markets. Cho et al. (2015) show how disruptions to critical highway infrastructure, such as highway bridge and

tunnel damage, can lead to economic losses in the U.S. on a state-by-state and industry basis. They conclude that the states and industries that can adapt to disruption suffer the least economic loss. Therefore, redundancy in transportation networks can mitigate the consequences of a disruptive event. This conclusion is mirrored by Worton (2012), who suggested that resilience engineering should focus less on efficiency than on the capacity for preparedness, recovery, and adaptation (Mattsson and Jenelius 2015). The shortcomings of existing studies on the economic implications of transportation disruptions are twofold: 1) resilience is often conflated with potential economic damages and 2) economic modeling is not underpinned by any analysis of transportation network topology.

This paper aims to demonstrate that joining resilience analysis with regional economic modeling can advance the methodological approach necessary for planning. This integration will allow for transportation costs to be associated with travel time delays, as delays are a primary contributor to economic impact. The current paper uses REMI TranSight (Regional Economic Models Inc. 2018), an established economic modeling process that uses input-output, computable general equilibrium, econometric, and economic geography methods. Disruption-induced delay results generated by Ganin et al. (2017) were translated to economic outcomes by formulating them as inputs to REMI TranSight. This expanded the ways to use travel delay information, which is more traditionally used in scenario and policy analysis. We introduce a process for quantifying the economic implications of resilience, or lack thereof, and demonstrate this process for 10 U.S. cities. This can be used to develop a planning approach that explicitly considers resilience. The methodology used in this study can be instrumental in the transition from current risk-based planning to true resilience planning, supported by economic analysis and subsequent selection of management alternatives.

2.4.2 Methods: Integration of Transportation Network Resilience and Economic Models

This paper joins two independently developed and documented models to assess the economic impacts of resilience in transportation networks: 1) Ganin et al. (2017) urban traffic network simulation, which estimates travel times given either a predictable disruption (e.g., peak commute hours) or random disruptions and accidents (e.g., natural disasters and major accidents), and 2) Regional Economic Models Inc. (2018), a regional economic forecast-

ing model oriented specifically for simulating the outcomes of changes in transportation systems.

Network Model for Simulating Delays Associated with Disruptions. Effects of disruptions on an urban transportation infrastructure are taken from Ganin et al. (2017). Specifically, Ganin et al. (2017) modeled travel delays that result from road network disruptions during peak-hours for private vehicle commuters in 40 urban areas in the continental United States.

In order to study travel delays under both normally functioning and disruption scenarios, Ganin et al. (2017) first built transportation graphs (as in Figure 21) for each of the urban areas, comprised of intersections connected by roadways, and then simulated commuter trips on the graphs. Specifically, urban areas were spatially divided based on the Voronoi tessellation method [1] and were centered at each road intersection. The expected population of each cell was calculated using data from the U.S. Census Bureau. Trip distribution was accomplished with a modified gravity model and the privately-owned vehicle was the only mode of travel studied. Route assignment was done assuming free-flow speeds on all roadways. Further, commuters were assumed to travel along the routes which minimized their trip durations. The output of the model is the annualized travel time per commuter. Free parameters of the model are calibrated to match measured data on the annualized delay per a peak-hours auto commuter, as given by the Urban Mobility Scorecard (Schrank et al. 2015).

---

[1]Voronoi tessellation is a technique for partitioning a plane into cells around points such that the distance from a given point to the boundaries of the cell is less than that of the other points (Skiena 2008).
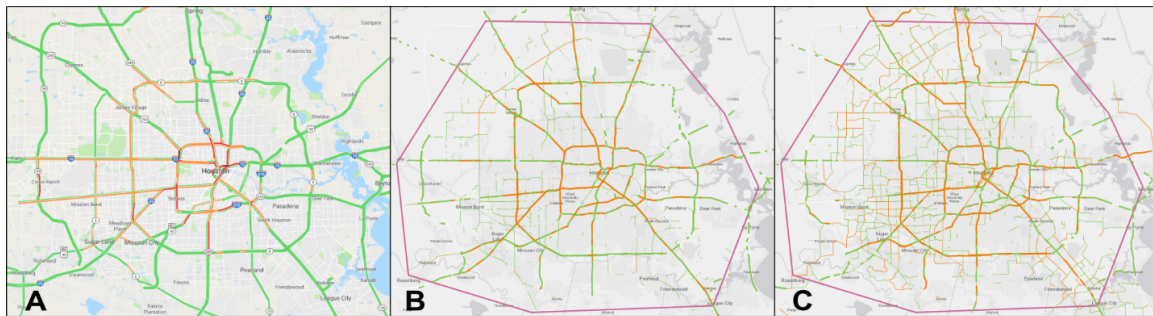
A   Google Map typical traffic at 8am

B, C   Modeled delay per km (min):   —— < 1.2   —— 1.2 - 12   —— 12 - 24   —— > 24

     —— Highways   —— Other roads

▭ Approximating urban area boundary polygon

Figure 21. Roadway networks of Houston, TX: (A) Congestion patterns at 8 am per Google Maps, (B) Modeled delays per kilometer of travel in normal conditions, and (C) Modeled delays per kilometer of travel under 5% disruption. The purple line shows an approximation of the urban area boundary in panels (B) and (C).

To characterize the resilience of modeled cities, Ganin et al. (2017) generated disruptive events on the transportation networks by disabling links. Resilience was quantified as the additional delay resulting from these disruptions. More specifically, Ganin et al. (2017) randomly selected a fixed fraction ρ of network links that was made non-functional by reducing the free-flow speeds of the network links to 1 km/h. Links were selected at random, with probabilities proportional to their lengths. Then, traffic was redistributed per the updated link free-flow travel times, assuming the same origin-destination demand matrix and assuming travelers were still using routes that minimized trip duration. Additional delays that resulted from these events were evaluated as the difference between annualized travel time with and without a disruption. The details of each step of the model development can be found in Ganin et al. (2017).

Economic Modeling to Forecast the Impacts of Transportation Network Changes. TranSight (Regional Economic Models Inc. 2018) is designed to be used with transportation forecasting models to translate the outcomes of improvement measures into regional economic implications. Regional Economic Models Inc. (REMI), the developer of TranSight and related products, maintains models for a wide variety of U.S. regions and states in order to support research. For a particular city or region, the effects of transportation

projects are forecasted in economic terms that include gross domestic product (GDP), employment, delivered price, commodity access, labor access, and relative cost of production. REMI mixes techniques from Input-Output (I-O) and Computable General Equilibrium (CGE) modeling, as well as economic geography and econometric techniques (Regional Economic Models Inc 2017). In this research, TranSight is paired with forecasted delays that result from transportation system disruptions instead of changes in travel times that are associated with road network improvements.

The inputs to TranSight from a transportation model include: (a) changes to Vehicle Miles Traveled (VMT), (b) changes to Vehicle Hours Traveled (VHT), and (c) vehicle trips attributable to improvement measures. Changes in velocity (measured by VMT divided by VHT) from the baseline scenario to the improvement (test) scenario are formulated as having a proportional effect on transportation cost and changes in trips that can be made in a given amount of time (measured by the number of trips divided by VHT) are formulated as having a proportional effect on accessibility cost. Changes from the baseline transportation network are presumed to affect various economic variables via changes in "effective distance," which functions to change travel time or commuting time and expenses. Cost-savings due to reduced travel times accrue to industry firms in the model due to reduced commuting and transportation costs and increased access to markets.

Cost-savings, capital investment, and other financial and economic concerns associated with prospective infrastructure projects are related to the regional economy. Changes to economic variables are modeled via policy variables that represent the effect of travel time on individual spending on fuel and subsequently disposable income and consumer spending. Additionally, costs to industries are related to transportation via the extent to which labor demand can be met and the composite price of goods sent to market. All of the linkages between travel times and associated costs to the policy variables that are used in TranSight are detailed in Regional Economic Models Inc (2017) and the TranSight 4.2 User Guide (Regional Economic Models Inc. 2018).

Connecting Transportation Network Resilience and Regional Economic Modeling. For the specified road network disruption severity ρ, quantifying the fraction of affected roadways, the network model (Ganin et al. 2017) calculates the resulting average annualized travel time $T(\rho)$ per a peak-hours commuter, which can be compared to travel time in the

absence of disruption $T(0)$. The topological attributes that yield higher or lower $T(\rho)$ values are not called into question in this research. We assumed a linear relation between the transportation costs and the relative increase in travel time due to network disruption, estimating the corresponding percent increase in transportation costs $c(\rho)$ as

$$c(\rho) = \Delta T(\rho)/T(0) \qquad (24)$$

where $\Delta T(\rho) = T(\rho) - T(0)$.

For the purposes of this demonstration, we relied on changes in gross domestic product (GDP) as an indication of economic impact for individual cities assessed. We did not probe the specific mechanisms of the economies of each city. To quantify the effects of the transportation cost increase on GDP, we utilized the TranSight model, and then generated the relative change in GDP as a function of disruption severity for the 10 cities of interest. Two scenarios are modeled:

1) Baseline scenario. Assumes that transportation cost increases are directly proportional to road disruption severity:

$$c^0(\rho) \equiv \rho \qquad (25)$$

2) Test scenario. Assumes that transportation cost increases are proportional to the additional travel time induced by disruption, not to the fraction of roads affected (see Eq. 24).

2.4.3 Results

Table 5 displays the percent increase in transportation cost computed for different cities at specific values of road network disruption severity ρ and serves as the basis for subsequent analysis of economic implications.

We observed that the same disruption results in different cost increases for different cities (note that cost is proportional to additional delay). The cost escalated quickly – even at 3% disruption, San Francisco exhibits a cost increase of 34%. The economic impacts of transportation disruption were most pronounced in San Francisco, where 5% disruption resulted in a 51% increase in transportation costs. Similar disruption in Los Angeles

73

Table 5. Transportation cost $c(\rho)$ increases by city that result from road network disruptions of severities varying from 1 to 5 percent. Shown in red are the transportation cost increase values exceeding 25%.

| | | Fraction of Affected Roadways (Network Links), ρ | | | | |
		1%	2%	3%	4%	5%
	Atlanta	4%	10%	16%	23%	**33%**
	Detroit	3%	6%	9%	14%	19%
	Houston	5%	11%	16%	24%	**32%**
	Jacksonville	7%	13%	22%	**33%**	**44%**
Transportation Cost Increase, $c(\rho)$	Los Angeles	1%	3%	5%	7%	9%
	Miami	4%	9%	13%	18%	23%
	Orlando	4%	9%	14%	20%	**26%**
	San Francisco	9%	20%	**34%**	**43%**	**51%**
	Seattle	3%	6%	9%	13%	17%
	Tampa	6%	12%	20%	**26%**	**37%**

resulted in only a 9% increase in transportation costs. Jacksonville is overall the second most severely impacted city considered, with a 44% increase in transportation costs.

TranSight forecasts future economic performance in order to understand the short- and long-term implications of disruptions and management policies. Figure 22 shows the impact of a 5% increase in transportation costs on the GDPs for 2019 and the forecasted subsequent five years for the baseline (22a) and test case (22b) scenarios. All simulated cities showed significant impact on GDP in 2019, the year in which the disruption occurred; the residual effect of that shock over the five subsequent years, however, was relatively small. No remediation or recovery activity was simulated and transportation costs returned to prior levels in 2020-2024. All of the cities recover to within 0.2% of their expected GDP (simulated GDP in the absence of a shock) by 2020 and have exceeded it by 2023. The profile of the initial impact and recovery was very similar for the test case, yet the magnitude was substantially greater, with disruption roughly an order of magnitude worse

than the baseline case. Notice the difference in the y-axis of the Figure 22 graphs.
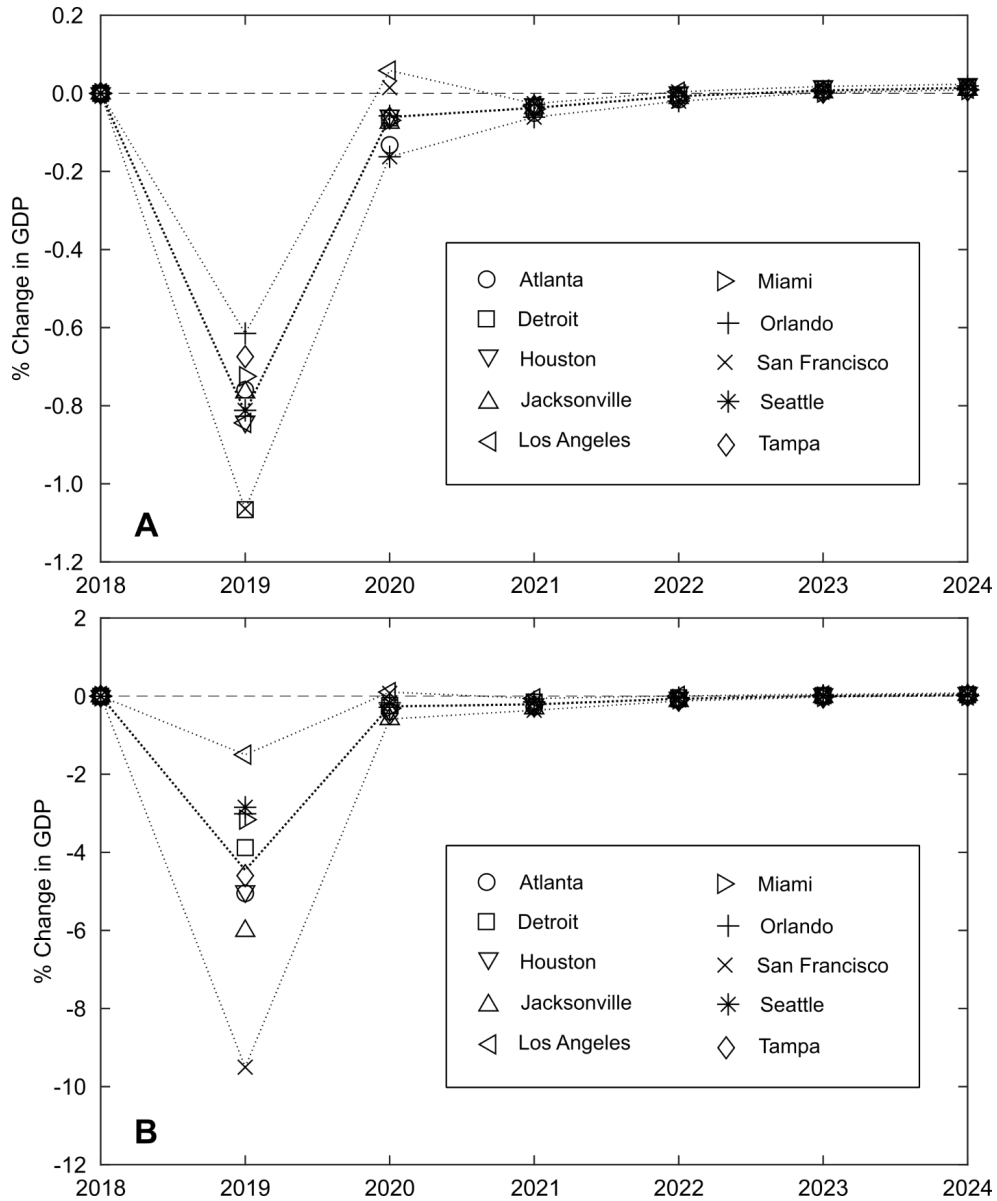
Figure 22. Temporal performance of regional economies (as measured by GDP) in response to a 5% increase in travel costs that lasted one year (A) and to a 5% transportation network disruption that lasted one year (B). Dotted lines show the lowest, mean, and highest values for each year. The dashed line corresponds to zero change in GDP from expected GDP in the absence of disruption.

In Figure 23, GDP changes from the baseline scenario are displayed by the colored bars and those for the test scenario are shown with the transparent bars. For both scenarios, GDP progressively declined as a function of disruption severity. At the same time, we note that changes in GDP due to travel time delays were significantly more consequential than for their baseline counterparts. For example, in this model, a random disruption of 3% of road segments in the San Francisco urban area results in a $c(\rho) = 34\%$ transportation cost increase. This leads to a 6.64% GDP decrease, significantly more than the baseline result of 0.64%. Not all cities show such disparate results between the two scenarios. For example, for Los Angeles (the 5th bar), a 1% roadway disruption increases the travel time by approximately 1% and therefore the GDP effects are the same (the transparent and the colored bars completely overlap).
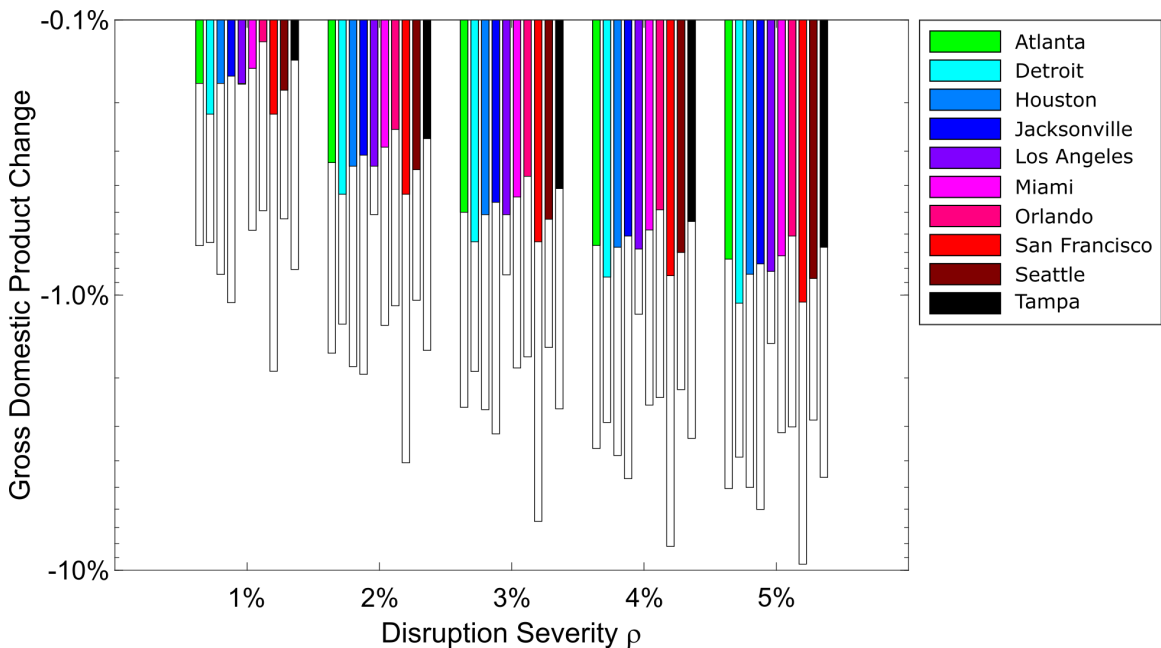
Figure 23. Impact of disruption on GDP (vertical axis scale is logarithmic). The colored bars correspond to the baseline scenario, with transportation cost increase being proportional to the disruption severity ρ, while the transparent bars present the case when transportation costs change per travel time increase in response to a disruption.

Even though higher percent additional delay due to road network disruption predictably results in greater economic impact (Figure 24), the lack of complete linearity signifies that

there may be differences between the city economies in question that lead to the different impacts of regional disruption on economic outputs. For 5% disruption, the proportion of GDP loss in Orlando, Tampa, and Jacksonville are lower than the trend while the rate of GDP loss in San Francisco and Detroit are higher than the trend.
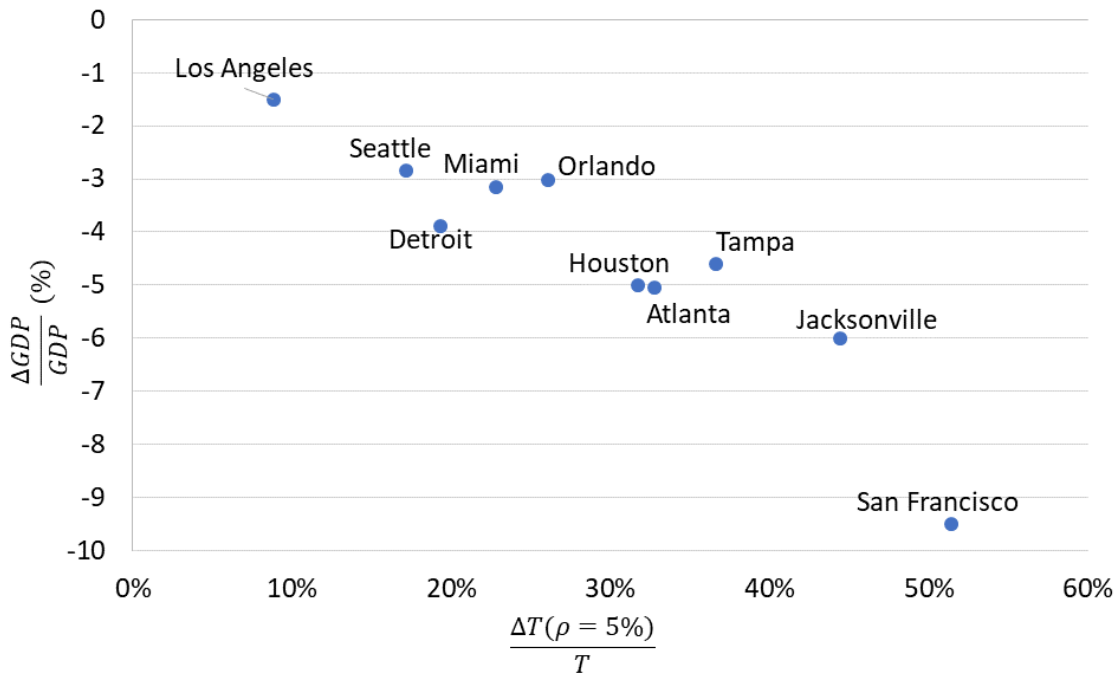
Figure 24. Percent additional travel time that results from the simulated disruption of 5% of the road network is plotted against the percent loss of GDP that results from that delay. Loss of GDP is related to the delays caused by disruption.

A more detailed study of the relationship of total regional GDP and GDP per capita loss to travel disruption may be important for revealing fundamental differences across studied metropolitan areas. The rate of GDP loss associated with each additional 1% network disruption is not constant within or across cities. Some cities lose a similar amount of GDP with each additional increment of disruption while others lose a more variable amount (see Figure 25; note that the losses are not displayed cumulatively). For example, Orlando has a similar amount of loss each time another 1% of the roads are disrupted (visualized through tight clusters of graphical points). Houston suffers less loss from 2 to 3% disruption than from 1 to 2% but then much more loss from 3 to 4%, which is visualized as more disperse

points. This demonstrates that the impact of additional disruptions is difficult to predict and that there may be thresholds and unknown factors in play that affect whether additional road network disruptions will translate to disproportionate economic impacts.
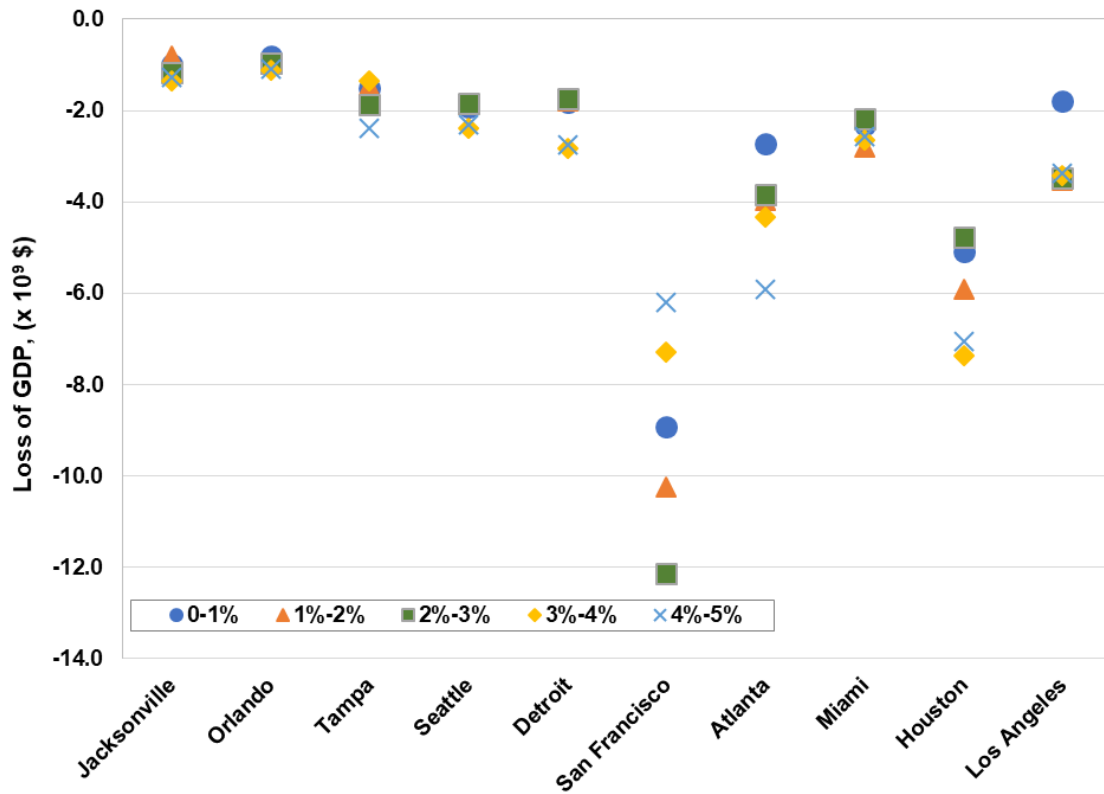
Figure 25. GDP loss attributable to each additional 1% road network disruption (not cumulative). Losses attributed to an increment of disruption are variable and do not have a pattern that follows the addition of increments.

To investigate whether the tolerance of city economies to road network disruptions is related to their size (as reflected by GDP) or wealth (GDP per capita), we look at how much the forecasted GDP or GDP per capita changes over a 1% disruption increment. The results here are for the disruption window of 4% to 5%. Loss of GDP over that increment is negatively correlated with GDP (Figure 26A). That is, larger economies, as measured in GDP, tend to suffer more loss in productivity than smaller economies when road network disruption is increased from 4% to 5%. All cities lose an additional 1% (+/- 0.4%; mean = 0.98%, standard deviation = 0.8%) of GDP with the exception of Los Angeles (0.33% loss) (loss as

78

a percent of GDP is displayed as inset graph). In comparing the main and inset graphs, we can see that the largest economy (Los Angeles) is the most tolerant to network disruption but the next two largest economies (San Francisco and Atlanta) are comparably intolerant, as they lose a similar percent of their GDP as some of the smallest economies in the sample (Tampa and Jacksonville). Modeled economic impact is also negatively correlated with wealth, as measured in GDP per capita (see Figure 26B), although to a lesser extent than GDP. Cities with higher productivity per capita have a higher rate of loss when percent road network disruption is increased from 4% to 5% than cities that have lower per capita GDPs. The cities that are more tolerant, in that they lose less GDP per capita than their counterparts, have a wide range of wealth; Seattle, Miami, and Orlando lose a similar percent of their GDP per capita, as evident from the inset.
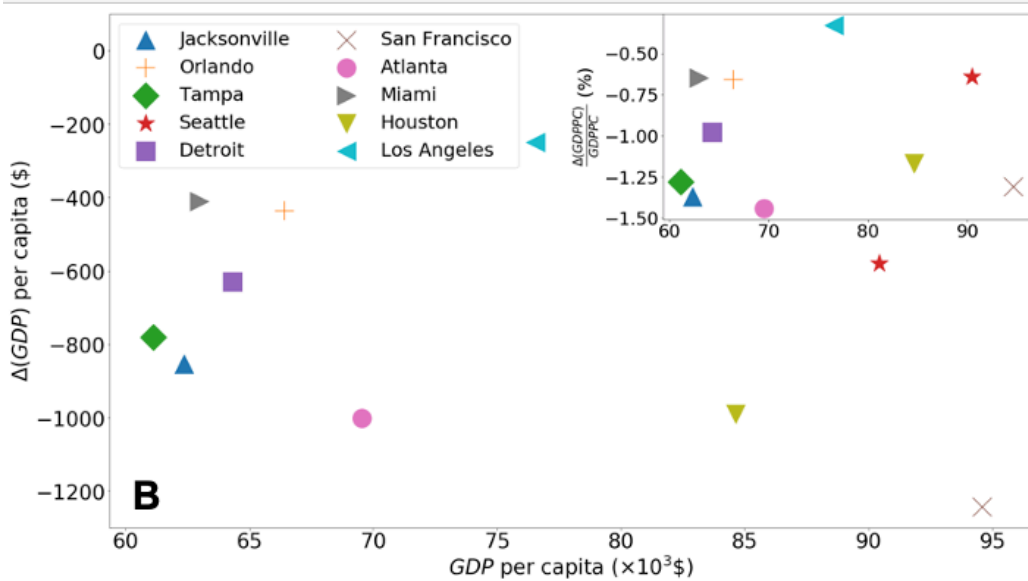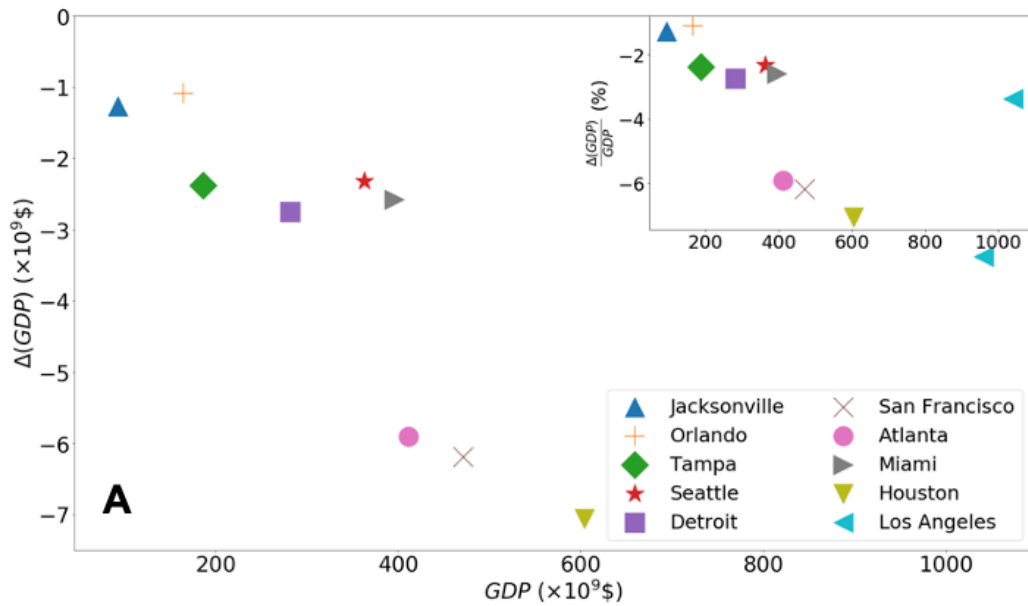
Figure 26. Comparison of city economy size (as measured by GDP) (A) and wealth (as measured by GDP per capita) (B) as explanatory variables of the tolerance of city economies to additional disruption. Results are for additional GDP losses attributable to increasing percent network disruption from 4% to 5%. Insets show losses as a percent of GDP and GDP per capita, respectively.

2.4.4 Discussion and Conclusions

This study is the first to integrate network resilience models for transportation networks with economic models. We found that travel delay resulting from disruptions in transportation networks result in non-trivial economic impacts. The economic outcome of disruptions cannot be predicted by a simple incremental change of velocity-related parameters (VMT, VHT). This research demonstrates that economic models must also be paired with structural network models and analysis in order to best reflect the impact of disruptions on economically important processes, such as the access of commuters to their workplaces and firms moving their commodities to market. This work also shows that the degree of disruption (i.e., percent of roadways disrupted) cannot be assumed to be proportional to transportation cost increase. Rather, network disruptions can result in a wide variety of transportation cost increases depending on a number of different factors. As such, one important conclusion of this work is that losses associated with network disruptions may be an order of magnitude higher than the size of the disruption itself, as evidenced in Figure 23.

Improving the ability of infrastructure systems to continue to function after unexpected disruptions (i.e., resilience) is emerging as a high priority in infrastructure planning. Resilience is an objective that has to be balanced against other system performance objectives, including the efficiency of transportation networks. Ganin et al. (2017) used a previously developed methodology for network resilience quantification (Ganin et al. 2016) and demonstrated that efficiency and resilience are not correlated in the cases of 40 U.S. cities. Yet, Ganin et al. (2017) did not assess the economic implications of a lack of resilience in transportation networks. Rose (2017) gives a comprehensive overview of the challenges faced in accounting for the economic impacts of disruptions, including that of translating damaged public infrastructure into broader economic losses.

It is intuitive that the effects of disruptive events are not only direct physical impact. The extent to which these impacts can propagate through regional economies needs to be quantified in order to motivate remedial action and inform priorities. A combination of factors is likely contributing to the forecasted impacts of road network disruptions on economies. Potential factors are characteristics of both the road networks as well as the economies. In this proof-of-concept study, only exploratory analysis of selected economic characteristics was used. A more detailed analysis of transportation networks was not attempted.

The question of what makes some cities more sensitive to transportation disruptions than others is important to study further. Our findings show that the economic effects of road disruptions do not perfectly scale with the productivity of city economies (in terms of GDP) or wealth (in terms of GDP per capita) and that not all of the regional economies in the study are equally affected by disruptive events. A naïve expectation is that the more productive an economy, the more sensitive it will be to transportation infrastructure disruption due to its reliance on a mobile workforce. Yet, the results of this study imply that more complex processes are at work. Further evidence of the differing sensitivity of city economies to disruptions include that each additional increment of disruption does not show a pattern across cities and that recovery from simulated disruption is not the same across cities. We can speculate that the sensitivity of an economy is dependent on the reaction of the transportation network to a disruption as well as the dependence of the economy on its transportation network, but more research is needed to understand the factors at play. The implication of this result is that both region-specific economic and region-specific transportation models are necessary for resilience planning.

The model is limited, however, in that the calculated transportation cost increase was applied for a period of an entire year; a finer temporal resolution would be preferable and would yield more interesting results in terms of recovery of economies from disruptive events. Additionally, the modeled cost increases may be conservative as, in reality, unexpected delays can result in disproportionately higher cost increases as compared to expected delays. For example, traffic congestion does not disrupt businesses on a day-to-day basis because it can be planned; unforeseen delay cannot be planned around and may cause disruptions that hurt businesses. Because Value of Reliability is not taken into account in the economic model used here, the economic impacts of transportation disruption were conservative. Other limitations include excluding public transportation, which is an important consideration in coupled network models. Similarly, the traffic model is intentionally abstract and simple and only uses readily available data. Integration of explicit travel pattern (e.g. using cell phone data) could improve specificity of the model, but it will limit its universal appeal because these data may not be available for all cities and may be of varying quality. Finally, the current model does not differentiate between the expected travel times, the variance of everyday traffic, and unpredictable delays. Future work should investigate how to integrate potential methods of measuring the Value of Reliability, separate from expected travel times,

as well as how unexpected delays may best have their costs modeled and integrated. Similarly, whereas this research considers how different city economies recover from a simple and passive standpoint, future work should consider the economies themselves with a more in-depth consideration of the specific economic indicators and recovery patterns (Rose and Dormady 2018).

A key motivation of this research is to highlight that efficiency, risk reduction, and resilience are different and often competing objectives as it relates to making investment decisions on improving transportation networks. Although disruptive events result in delay, and sometimes much more pronounced delay than routine congestion, they cannot be prevented solely by improving efficiency of a transportation network or its hardening to reduce risks. Risk reduction can have limited success for uncertain events. For example, in flow networks, such as transportation systems, efficiency may be achieved by having sufficient roadway capacity while resilience can result from the availability of (potentially not very efficient) alternative routes, which would ensure graceful performance degradation under disruption as opposed to a complete collapse. Planning for resilience is functionally different than planning for efficiency and hardening (Ganin et al. 2017).

Future work should explore the driving factors of uneven outcomes across cities, namely the reasons for transportation network sensitivities to disruption as well as local economy sensitivities to transportation failures and ability to recover. Additionally, future work can advance areas in which this study were limited and move the method toward practical application in transportation planning. For example, this paper did not investigate the topological attributes of road networks that yield more or less delay. This paper also did not investigate the mechanisms by which network disruptions cause economic impact in the results. Efforts to enhance the resilience of road networks to disruption and/or of the regional economy to a lack of road network resilience will need to study the outcomes of the models in detail. This paper is the first to provide fusion of a well-established Regional Economic Model (Regional Economic Models Inc. 2018) with the transportation network resilience model developed by Ganin et al. (2017), and to apply it to multiple cities in the United States. This explicit integration of a temporal resilience model linked with a regional economic model allows for the comparative evaluation of multiple cities with respect to the impact of resilience. This paper demonstrate that neglecting to plan for disruptive events from a resilience perspective can result in a disproportionate impact on city economies.

2.5 Vulnerabilities of Intelligent Transportation Systems (ITS) and the Need for Resilience

Many cities are adopting Intelligent Transportation Systems (ITS). These are "smart" systems that combine connectivity, coordination, adaptivity, and/or automated response for transportation policy optimization. Integration of "smartness" in many systems, including critical transportation infrastructure, is expected to increase efficiency, even in response to disruptions. However, the control and sensing systems of ITS open new vulnerabilities to nonphysical attacks. Current approaches to manage vulnerability focus on risk characterization and analysis. Emerging technologies by their nature have risks that are not fully known. Resilience, defined as the system's flexibility and recoverability, may be used to address unknown threats and adverse events. We conducted a study of network resilience in response to cyber physical random and targeted disruptions in 10 urban areas. Disruptions were affecting either intersections or roadways under different impact scenarios. Modeled attacks disrupting 20% of intersections caused on average 14.6% higher additional delays than the same severity attacks on roadways. Targeting by length caused the most severe disruptions. Contrary to expectations, targeting nodes and links with the highest number of shortest paths (betweenness) was not noticeably worse than random attacks. Thus as cities adopt ITS and other smart systems resulting in potentially unknown vulnerabilities, it is important to consider resilience of transportation infrastructure affected by potential cyber attacks.

2.5.1 Introduction

Smart cities have been a rising new concept for urban planners, with new initiatives pushing for more efficient urban transportation, such as the US Department of Transportation's Smart City Challenge (U.S. Department of Transportation 2017). In smart cities, systems use connected technologies to collect and analyze data in real time, and to make predictive and adaptive decisions (Albino et al. 2015). Smart transportation systems are generally called Intelligent Transportation Systems (ITS). Many ITS applications are focused on improvements at intersections and freeway ramps/merges (Coogan and Arcak 2015; Ghena et al. 2014; Makarem and Gillet 2012; Onieva et al. 2012). This is often an efficient implementation as improving intersections allows for improvements at limited points to extend to the roads linking to the intersections themselves. While the high degree of connectivity in transportation systems allows smart cities to improve early detection and response, in-

crease efficiency and reduce risks, the connectivity also presents a new vulnerability itself (Lombardi et al. 2012). Systems that can either be controlled remotely, respond to traffic conditions, or communicate with other systems are vulnerable to having the pathways that they use to communicate with each other and sense the world being remotely hijacked (Amoozadeh et al. 2015; Petit and Shladover 2015; Reilly et al. 2015). The potential for remote disruption may make these intersections more attractive targets to malicious actors, as they may no longer need to physically expose themselves, to cause a disruption. Additionally, just as improvements in intersection performance spread to their linking roads, disruptions will also propagate.

While improvements in security can reduce risk, they cannot eliminate it. One framework to mitigate some of the possible impacts of failures is to increase system resilience through redundancy, modularity and/or autonomy. Resilience is defined by the National Academy of Sciences as the ability to prepare for, absorb, recover from, and adapt to disturbances (Linkov et al. 2014c; Council 2012). Resiliency is the subject of new policy initiatives, such as Presidential Policy Directive 21, which seeks to promote resilient infrastructure that is not just efficient but that is able to withstand changing conditions (Committee 2015). This effort has not yet yielded a clear and agreed upon method to measure resilience in transportation.

Transportation resilience has been defined as the ability to maintain a "demonstrated level of service or restore itself to that level of service in a specified timeframe" (Freckleton et al. 2012). Even so, efforts to quantify resilience have been hindered by the difficulty in complex transportation systems modeling. Existing resilience methods are either qualitative or data intensive, often requiring specific system knowledge (Hughes and Healy 2014). Further, many conventional traffic models now utilize private datasets including mobile phone GPS data (Calabrese et al. 2011; Woodard et al. 2017). The limited availability of this data makes verification and replication difficult. This leads to a question – As ITS systems become more prevalent in a city, what are the implications for transportation resilience? This paper argues that network science (Ganin et al. 2016) provides a foundation for the evaluation of tradeoffs in designing smart and resilient transportation systems. This paper contributes to the literature by providing a method and metric to assess ITS road network resiliency and a case study of this method for 10 cities.

This section is organized into 7 subsections: states goals, defined resilience and describes the paper's organization; goes over relevant background context and literature; describes the methodology used in this paper's analysis; summarizes the results extracted from the analysis; discusses the results and their implications for practice; concludes the section by summarizing the previous subsections and discussing methodology limitations and potential areas of future investigation.

2.5.2 Background

Resilience and Risk in the Context of ITS Networks. ITS networks and components have taken many forms and undergone multiple evolutions of organizational structure. Early examples of networks billed as ITS include synchronized traffic lights (Auer et al. 2016). These systems could be based only on timings, and therefore be entirely autonomous, or some level of connectivity between the intersections, to minimize commuter time in red light phases (Auer et al. 2016). Later systems included centralized controllers which could connect to and coordinate actions among many areas and respond to changing conditions (Auer et al. 2016). These hierarchal systems, however, are vulnerable to widespread disruption as failure at controlling nodes can affect performance of large percentages of the system, giving them low resilience. A newer strategy is to design autonomous components that communicate together and are aware of each other's' protocols (Auer et al. 2016; Makarem and Gillet 2012; Shladover et al. 2012). If each component can make decisions independently then failures can be contained. Among these developing autonomous components are autonomous and connected vehicles; capable of performing some or all driving tasks, while communicating and cooperating with other vehicles and infrastructure (Administration. 2016). The communications component can, however, also be used as an avenue of attack, or to propagate disruption by feeding in false information (Amoozadeh et al. 2015; Petit and Shladover 2015; Reilly et al. 2015). Even systems without communicative protocols can be vulnerable if they are designed to sense traffic conditions (Petit and Shladover 2015).

ITS vulnerabilities were showcased when channels used by Sensys Networks VDS240 wireless controllers were spoofed by security researchers into changing traffic patterns through tricking the control system into believing that different traffic patterns, than those actually present, were occurring (Cerrudo 2014; Zetter 2014). This major security issue stemmed from unencrypted communications between sensors and the signal control system,

leaving cities like Washington, DC, Seattle and New York City vulnerable to attack and prompted a major security patch of the system (U.S. Department of Homeland Security, Industrial Control Systems Cyber Emergency Response Team 2014). Furthermore a study by researchers at the University of Michigan not only found similar vulnerabilities to the aforementioned tests, but also noted that many devices were found online by the search engine SHODAN, with little to no authentication protection (Ghena et al. 2014). Other researchers have found access to traffic camera data and even traffic signals which could be put into test modes without credentials (Goldman 2013). SHODAN and other mass scanning research efforts have been used at larger and larger scales, such as Project Shine (Shodan Intelligence Extraction) where researchers worked with the Industrial Control Systems Cyber Emergency Response Team (ICS-CERT) of the Department of Homeland Security (DHS) to categorize 460,000 IP addresses with predicted system impact and linking them with location, sector type and organization name (Team 2012).

While cyber-attacks remain a key threat, physical access and insider knowledge remain key issues. In 2009, two Los Angeles traffic engineers pled guilty to illegally accessing a city computer to cause major disruptions by disabling controls at four intersections (Grad 2009; O'Leary 2009). Although this incident was performed as part of a lab dispute, and not a targeted event, it emphasizes the vulnerabilities of transportation systems to insider attacks of malicious intent. When insider knowledge and wireless access are both unavailable, physical access can compromise hardware (Byres and Lowe 2004). Although traffic light systems are designed to absorb physical failures, a coordinated attack with cyber physical components can amplify damage (Li 2016). For example, even simple attacks on traffic signals, such as copper theft near freeways, can temporarily cripple transportation systems (Reilly et al. 2015). Furthermore, while smart grids have traditionally been able to protect their hardware with physical access controls and protections, two way meters being deployed now are accessible by consumers and adversaries (Khurana et al. 2010). As ITS, vehicle-to-vehicle and vehicle-to-infrastructure technologies are developed and adopted, city vulnerability increases. A more recent example is spoofing mobile phone data GPS, thereby influencing applications such as Google Maps or Waze into rerouting drivers around non-existent traffic jams (Sinai et al. 2014).

Because it is nearly impossible for a system to mitigate all possible threats to smart infrastructure, transportation security must incorporate resilience thinking; to absorb and recover

from possible attacks when they do occur. As such, the Department of Transportation has developed the Cyber Security Evaluation Program and Cyber Resilience Review process to better access cyber security (U.S. Department of Homeland Security 2015). Recent work in modeling has been aimed at modeling possible physical impacts of cyber-security breaches, such as in regional transport (Omer et al. 2013), freeway traffic control (Reilly et al. 2015) and traffic intersections (Laszka et al. 2016). However, these models are often not detailed enough at the city level scale to characterize urban resilience (Omer et al. 2013; Reilly et al. 2015) or focus on directed attacks (Laszka et al. 2016). To bridge this gap in understanding of the impact of attacks on city scale transportation systems, we illustrate an application of a transportation resilience and efficiency model (Ganin et al. 2017) and use it to analyze directed and random attacks on both node and link disruptions within selected United States urban transportation networks.

Literature Review. Table 6 summarizes 17 recent, published on or after 2007, relevant and reviewed ITS papers. Column 2 notes if the paper discusses ITS effects on network performance. Column 3 notes if the paper discusses ITS effects on risk at all, to a network or node. Column 4 notes is the paper discusses ITS effects on network resiliency. For the purposes of this table resiliency is considered addressed if the paper measures system performance from the immediate disruption, until a steady state has been achieved, or full system recovery. Column 5 notes if the article addresses security pathways and vulnerabilities within ITS devices themselves. Column 6 notes if the paper used a metric to measure the impact on network performance, risk and/or resiliency, or device security. ITS effects on risk and resiliency are considered fulfilled if addressed from the perspective of additional vulnerabilities that these technologies might pose. Among the reviewed literature the majority of papers addressed either device security or ITS performance effects, in isolation. Three papers discuss the risks that security vulnerabilities might have on performance, but only on a limited basis. Two of these papers only considered the limited number of vehicles directly connected to each other in a Cooperative Adaptive Cruise Control (CACC) platoon. This limits direct applicability to a single link or piece of a link, rather than network performance. Petit and Shladover investigated more effects, but only qualitatively. The lack of metrics limits the direct applicability as comparisons among different disruptions are difficult. None of the reviewed papers investigated how ITS technology, its adaptation and vulnerabilities would affect transportation network resiliency. Resiliency is important as it

Table 6. Summary of ITS Literature 2012 through 2016

Paper	Addresses Performance	Addresses Risk	Addresses Resiliency	Addresses Device Security	Outputs Metric(s)
(Mahmassani, 2016)	✔				✔
(Verma et al., 2016)				✔	
(Vijayakumar et al., 2016)				✔	✔
(Amoozadeh et al., 2015)	✔	✔		✔	✔
(Coogan and Arcak, 2015)	✔				✔
(Khondaker and Kattan, 2015)	✔				✔
(Petit and Shladover, 2015)		✔		✔	
(Ghena et al., 2014)				✔	
(Sedjelmaci et al., 2014)				✔	
(Whaiduzzaman et al., 2014)				✔	
(Zhang et al., 2014)				✔	
(Daeinabi and Rahbar, 2013)				✔	
(Gerdes et al., 2013)	✔	✔		✔	✔
(Makarem and Gillet, 2012b)	✔				✔
(Onieva et al., 2012)	✔				✔
(Shladover et al., 2012)	✔				✔
(Abdel-Aty et al., 2007)	✔				✔

shows how well a system is able to perform during and recover from failures, which may be investable with ITS adoption. Resiliency can also be used to plan for unknowable events and vulnerabilities, which traditional risk frameworks cannot. This section contributes to the literature by developing a method to evaluate the resiliency effects of ITS systems on a metropolitan area, when applied to intersection improvements. This section then presents a case study of 10 different cities against random and targeted disruptions.

2.5.3 Modeling and Methods

This section looks at transportation modeling from the classical graph percolation theory perspective and abstracts from the inherently complex and probabilistic nature of causing a disruption. Thus, we assume that an attack succeeds and results in a disruption of a certain fraction of either intersections or roadways.

We apply the model developed in Ganin et al. (2017) to evaluate delays in large transportation networks. The framework's main advantage is that it provides a unified approach to all large urban areas and does not require extensive datasets. Resilience is assessed based on the additional delays resulting from adverse events. The framework's algorithmic simplicity results in acceptable computing requirements. Ganin et al. (2017) focused on natural disasters and modeled disruptions as being random with likelihood proportional to road length. Malicious attacks may be both random and targeted and not necessarily correlated simply with length. Here, we expand upon the previous model to account for malicious attacks. In the rest of this section, we briefly outline the most important assumptions and equations of the model and refer the interested reader to (Ganin et al. 2017).

Roadways Geospatial Topologies. We employ U.S. Census Bureau shapefiles (U.S. Census Bureau 2012a) to define urban area boundaries. To simplify calculation of distances in the network, we manually draw simplified polygons around the exact urban area shape. The drawn polygons are available in the supplementary data. Every point comprising a roads polyline was checked to be within the boundary of the boundary polygon, or within 40 km of the boundary.

We use datasets (OpenStreetMap contributors 2017) provided by the OpenStreetMap (OSM) foundation to build transportation networks. Open Street Maps is a knowledge database that seeks to provide free access to current geographical information (Haklay and Weber 2008). We processed it to exclude minor and residential streets and service roads and clipped it (ESRI 2012) to only include features within 50 km of the boundary.

Trips Generation and Distribution and Traffic Assignment. Our model utilizes 2010 census datasets (U.S. Census Bureau 2010) to assign populations values to nodes in the network. This dataset contains population values broken down at the census tract level, wherein each small subdivision of a county is updated by local participants prior to each decennial census but remain relatively consistent over time. First, we build Voronoi polygons

90

(Voronoi 1908) around the intersections in the graph. As the census tracts and the Voronoi cells do not directly overlap, we estimate the population of each Voronoi cell using density values from the census tracts and the spatial overlap of census tracts and the cell. We estimate the population of each cell N_i in Eq. 26 1, where N_t is the population of Census Tract t, and P_i, P_t are the polygons of the cell and the tract respectively.

$$N_i = \sum_t N_t \frac{Area(P_t \cap P_i)}{Area(P_t)} \tag{26}$$

To determine the total number of trips to and from activities in an analysis area, most models use specific factors that influence the number of trips in a region including vehicle ownership, income, household size, type and density of development, availability of transportation alternatives and quality of the transportation system (Martin and McGuckin 1998). However, because our model is already based on simplified models of city networks, with much larger tracts than most perform, we use density of population as a proxy for these values. Specifically, we estimate the flow of commuters from an origin region o to destination region d to be proportional to the population at the destination N_d and inversely proportional to the cost function of distance between the two nodes. In our model, we assume that all drivers tend to optimize their commute paths such that travel time is minimized. Given this assumption, we assigned paths for every origin destination pair using free-flow speeds. Next, we estimated the number of commuters using a particular link and used that and the road capacity assessed based on its OSM type and maximum speed (if known) to evaluate the density of vehicles on the link. Then, we evaluate the actual speed of vehicles along that link based on their density, the maximum speed and road type of the link.

Delays Evaluation. Delay is evaluated as the difference between the time it would take an average commuter to reach the destination in the network under load and in the empty network (moving with the maximum speed along each link). In Ganin et al. (2017), the model was calibrated for 20 urban areas and validated for 20 other urban areas in the U.S. The calibration fitted the modeled delays to the annual average delay per peak-hour auto commuter given in the Urban Mobility Scorecard (Schrank et al. 2015). The resulting Pearson correlation coefficient is 0.9 for the calibration group and 0.6 for the validation group. In this study, we choose the 10 cities with the best fit from those used for both

calibration and validation.

Adverse Events. Ganin et al. (2017) studied link disruptions with risk proportional to their lengths. Such a formulation was an abstraction of natural disasters, which were assumed to occur with equal likelihood within the network's spatial area. In contrast, ITS attacks typically target elements of the network, as discussed in Section Background. For simplicity and impartiality, we define the adverse events space to include random attacks hitting nodes or links with equal probability. A node's disruption is modeled as the corresponding disruption of its incoming links. We further compare targeted and random attacks. Targeted attacks affect nodes or links with the highest loads, lengths, or betweenness values. We assume that traffic speed along a disrupted link is 1 km/h. This allows the model to estimate an upper bound of disruption effects. Actual disruptions are likely to have variable and asymmetric effects (Petit and Shladover 2015). This complexity was ignored as it would risk ignoring unknown attack vectors, while an upper bound would remain useful, even if it is known to overestimate disruptions.

2.5.4 Results

Figure 27 presents an example of the changes in a transportation network loads distribution caused by a disruption of 20% of nodes in Washington DC. Real traffic distribution at 8 am as given by Google Maps is shown in Figure 27A and serves as a benchmark. Figure 27B and Figure 27C describe the modeled traffic distributions and highlight roads with the most congestion: links having congestion of less than 0.01 hour/km per commuter annually are not shown. We assume regular conditions in Figure 27B and a disruption of 20% of intersections in Figure 27C. We base our methodology to evaluate stress caused by links and nodes disruptions on the percolation theory. In percolation studies (Stauffer and Aharony 1994) it is typically assumed that a certain fraction r of nodes or links in the network is affected by an adverse event and is disrupted. We refer further to this fraction as severity of the adverse event. Ganin et al. (2017) considered random disruptions of links with risk proportional to their length. Here, we assume a uniform likelihood of damage and add node disruptions. Nodes and links disruptions are considered separately. We collect statistics for 20 independent realizations of an adverse event. Random nodes disruptions may be caused by accidents provoked by a hacker's attack on traffic lights while links disruptions may stem from natural disasters and accidents such disasters cause. We report

our results for the case where the disruption severity is 20% in Figure 28B and Figure 28C. For reference, in Figure 28A we start by providing the correlation plot for the 10 selected cities. As these 10 cities have the best fit, the correlation (Pearson $R = 0.99(p = 4.76 \cdot 10^{-8})$ is significantly higher than that for all 40 cities. Next, in Figure 28B we compare nodes and links disruptions. We observe a strong correlation between additional delays caused by these disruptions (Pearson $R = 0.98(p = 1.08 \cdot 10^{-6})$. In addition, it is evident that node disruptions cause higher additional delays than links disruptions of the same severity. Finally, Figure 28C correlates modeled regular delays to those caused by a nodes disruption. Contrastingly to the previous cases, here we observe a much lower correlation: Pearson $R = 0.75(p = 1.26 \cdot 10^{-2}$.
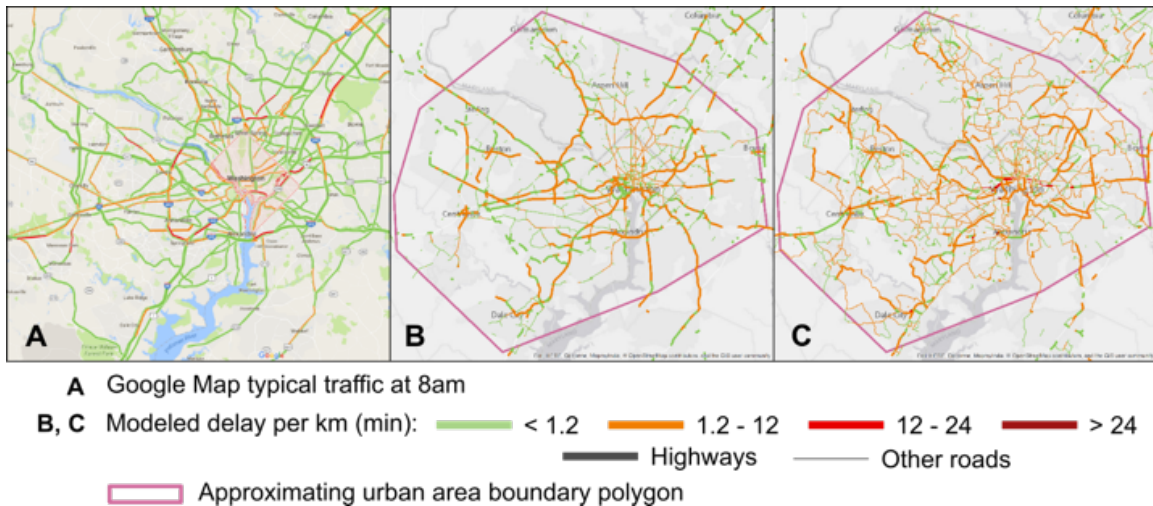
Figure 27. Traffic distribution in Washington DC. A. Typical congestion at 8 am as given by Google Maps. B. Predicted by our model for the case of no disruptions. C. Predicted by our model under the assumption of a 20% node disruption.

To characterize the difference between links and nodes disruptions we look at the case of Washington DC (20% disruption) in detail (Figure 29). Figure 29A shows a scatter plot comparing the two disruptions. In both disruption scenarios, we choose the same fraction of nodes or links (20%) to disrupt. However, as the incoming degree of nodes varies a disruption of 20% of nodes may result in a larger or smaller fraction of disrupted links. Notably, even for the cases when the fraction of damaged links is lower than 20% the corresponding delays are still higher than those in the links case.
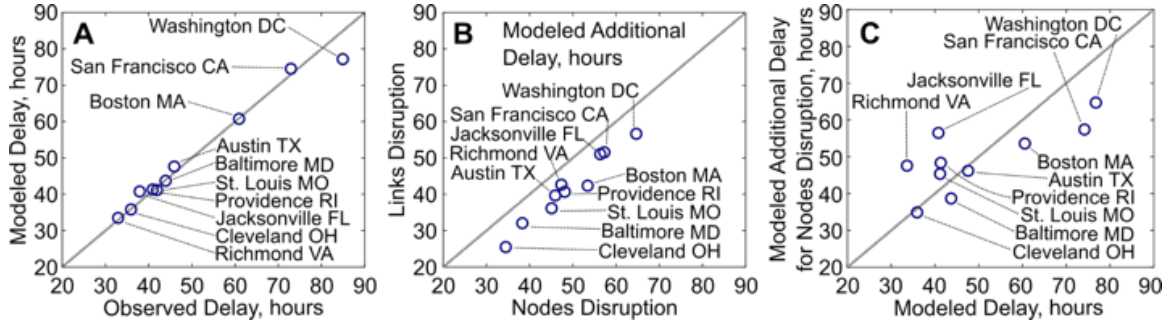
Figure 28. Correlation plots. A. Correlation between observed delays and modeled delays for 10 selected urban areas: Pearson coefficient $R = 0.99(p = 4.76 \cdot 10^{-8})$. B. Correlation between modeled additional delays for 10 urban areas for a 20% random links disruption and a 20% random nodes disruption: Pearson coefficient $R = 0.98(p = 1.08 \cdot 10^{-6})$. C. Correlation between the modeled delays and the additional delays due to a 20% random nodes disruption: Pearson coefficient $R = 0.75(p = 1.26 \cdot 10^{-2})$.

To calculate the distribution of the number of disrupted roads as a result of failure of n intersections, we apply the generating function formalism (Wilf 2006). First, we note that the number of links disrupted due to a successful attack on a single node is described by the network's incoming node degree distribution. This follows from our choice to model node blockages as disruptions of the corresponding incoming links. As each incoming link leads to a single node, numbers of links affected by the disruption of multiple nodes can be characterized as independent and identically distributed (i.i.d.) random variables. Denoting the incoming degree as kin and the fraction of nodes with that incoming degree as p(kin), we can define the said generating function, $G_{in}(x)$, as

$$G_{in}(x) = \sum_{k_{in}} p(k_{in}) x^{k_{in}} \tag{27}$$

The generating function of the sum of n i.i.d. random variables equals the generating function of a single variable raised to the power n. Then, the multiplier for x^k in the expression $[G_i n(x)]^n$ defines the probability that exactly k links are disrupted. In Figure 29B we assess the distribution of the disrupted links from the value $[G_i n(x)]^{5033}$, where 5033 is the number of disrupted intersections in the 20% disruption case, and compare the analytical and numerical results.

Finally, we also provide a distribution of additional delays resulting from links and nodes
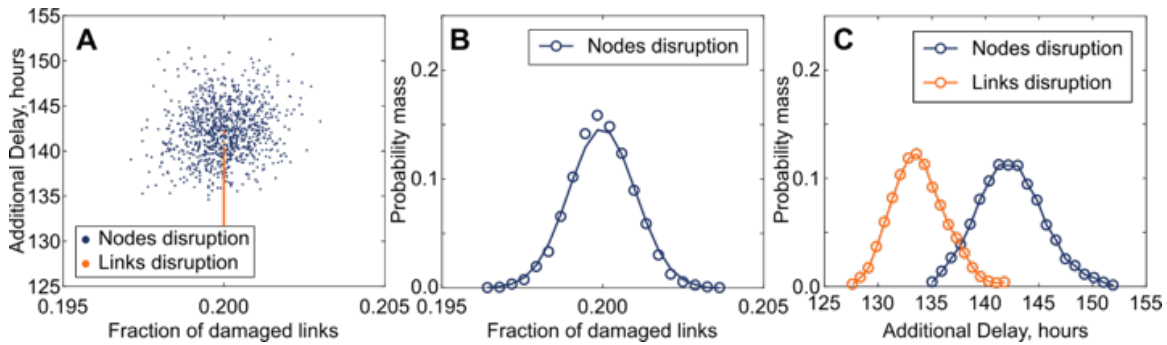
Figure 29. Detailed comparison of random disruptions intersections (nodes) and roads (links) in the Washington, DC urban area. A. Scatter plot showing the annual additional delay resulting from a disruption of 20% of intersections (orange dots) and 20% of roads (navy dots) for 1,000 random adverse scenarios. B. Distribution of the fraction of damaged links due to a 20% nodes disruption in 1,000 scenarios: circles show the results of the modeling while the solid line is estimated analytically. C. Distribution of annual additional delays in the cases of 20% nodes (navy) and links (orange) disruptions.

disruptions in Figure 29C and observe that these distributions allow us to use mean values to characterize resulting delays. We also note that the node delays expectantly have higher variability.

We further study how the severity of disruption affects the additional delay (Figure 30A). We focus on six representative urban areas and $r \in [0; 1]$. We observe that links disruptions result in lower additional delays and are less stressful for urban areas for all severities. Obviously, nodes and links curves cross at the point with disruption severity of 1. At that point, all roads are considered disrupted and the speed is reduced to 1 km/h. Finally, the dispersion values are low compared to the delays magnitude in all urban areas for all severities.
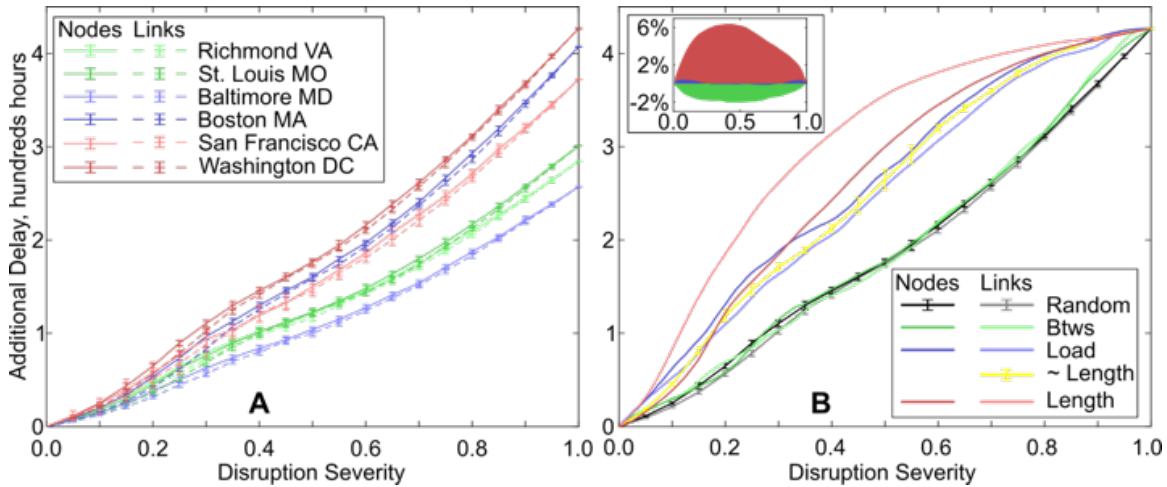
Figure 30. Delays and disruption severity, error bars show mean values ± standard deviation. A. Comparison between random disruptions of varied severity targeting intersections (nodes) (solid lines) and roads (links) (dashed lines) for 6 urban areas. B. Comparison between random (black and gray lines) and targeted attacks for Washington, DC. Nodes' disruptions are modeled as disruptions of their incoming links. Links to disrupt are chosen based on their betweenness centrality (green lines), load (blue lines), and length (red lines) values, nodes are chosen based on the sum of the incoming links' respective values. In addition, the yellow line shows a random disruption case where links are chosen proportionally to their length (Ganin et al. 2017). The inset shows how the fraction of disrupted links is different from the fraction of disrupted nodes in targeted nodes disruptions (by betweenness – green area, by load – blue area, and by length – red area).

Adversaries may be able to target nodes and links preferentially. To investigate how urban areas, respond to higher stress caused by intentional attacks, we focus on three strategies to target nodes and links deterministically. Attack Strategy 1 targets network elements of higher betweenness centrality value. The betweenness centrality value for a link was evaluated as the number of free-flow time optimal paths, between all possible pairs of nodes, passing through the link. The path length was assessed from the time to traverse links at free-flow speeds. Attack Strategy 2 targets network elements of the highest load (estimated from regular traffic patterns). Finally, Attack Strategy 3 targets network elements of highest length. All the above properties were evaluated for links. For nodes, they were taken as the sum of the respective values of the incoming links. We present results of the modeling for the Washington DC urban area in 30B. We observe several interesting patterns: i)

96

targeting nodes and links based on betweenness centrality is not different significantly from the baseline case of random disruption; ii) Links disruptions per Strategy 2 result in lower additional delays as compared to nodes disruptions following the same strategy; both cases result in significantly higher additional delays than the baseline case; iii) Strategy 3 causes the fastest increase of delays with disruption severity for small values of r. Additionally, links disruption, in this case, results in higher delays than nodes disruption. For comparison, the yellow line in 30B shows the scenario where the probability of a link's disruption is proportional to its length per earlier work (Ganin et al. 2017). We also characterize the fraction of links disrupted when disruption targets nodes per all three strategies in the inset to 30B. We observe that targeting nodes by betweenness damages a lower, than the disruption severity r, fraction of links. Strategy 3 (length) is opposite, that is, the fraction of damaged links is higher than r, even though it still results in lower delays.

2.5.5 Discussion

ITS systems rely on automated sensors and adjust their functions accordingly to changes in the environment. While the use of sensors and electronics brings benefits it also increases systems' complexity and vulnerability. Thus, infrastructure systems are becoming more vulnerable to unknown hazards or risks, whose likelihoods and mechanisms may be poorly understood. Resilience focuses on responding to unknown challenges by making the structure flexible and adaptable to changing conditions. Resilience focused planning prioritizes the general adaptability characteristics and unknown adverse events. In this paper, we conduct a resilience centered analysis by considering transportation infrastructure from a network science perspective abstracting from real events leading to disruptions. We make multiple important inferences from this analysis.

Our first result is that resilience and efficiency are not necessarily correlated. From Figure 29C, we observe that several cities that have similar values of modeled annual average delays per an auto commuter respond differently to random intersections disruptions. For example, Jacksonville FL and Richmond VA have disproportionally high increase in their delays in response to a moderate disruption compared to the rest of urban areas. Other such cities include Austin TX, Providence RI, and St. Louis MO.

Second, we observed that the predicted delays to intersections disruptions are slightly higher than those for roads disruptions. We determined that this pattern is applicable to all 10

97

urban areas and does not change with disruption severity. While this may be a consequence of the modeling assumptions, the finding may also be understood intuitively as follows. Typical traffic distribution according to the model is comprised of a set of highly occupied routes carrying the most load. Disruption of such a route results in traffic redistribution to an alternate route, assuming one is available. In the case of a random links disruption, there is a possibility that multiple links along the same route will be damaged and only single route will have to be altered. The likelihood of such a scenario for a nodes disruption is lower as nodes' incoming links point in different directions. On a more conceptual level, this may evidence on the roads disruption focus of system's designers. While the delays difference is not significant, it still highlights the importance of protecting intersections against adverse events. This recommendation is particularly important for smart systems dependent on other infrastructure networks (e.g. SCADA (Collier et al. 2016)).

Third, an important conclusion is that systems relative resilience is not sensitive to disruptions severity r. In Figure 30A, we observe that the ranking of urban areas does not significantly change with r. The most noticeable ordering change is Boston MA becoming second least resilient for $r > 30\%$ surpassing San Francisco. A similar observation was also made for link length proportional random disruptions in Ganin et al. (2017), in the case of uniform random disruptions. In contrast to Ganin et al. (2017), San Francisco CA performs significantly better and becomes only second least resilient (as opposed to being significantly worse than all other urban areas).

Finally, we also observe surprising response of transportation infrastructure to targeted disruptions (e.g. Washington DC in Figure 30B). Contrary to intuitive assumption, targeting nodes or links by their betweenness metrics is not significantly worse than a random selection case. This may intuitively be explained by the localized nature of betweenness focused attacks – such disruptions are clustered in a relatively small area corresponding to the "center" of the network. Attacks targeting nodes and links by their load do cause significantly higher values of additional delays than a random adverse event of the same severity. Consistently with the random attacks, intersections damage in this case is worse than roads damage. Last, we observe a change of this pattern for length based attacks: intersections damage is lighter than links damage. Moreover, length-based attacks are the worst of all three cases. This is understandable as rerouting from a long road requires much more additional roads. The observation calls attention to necessity of having more connections

98

between roadways which facilitate localized rerouting in response to a disruption.

2.5.6 Conclusion

Integration of smartness in many systems, including critical transportation infrastructure, is advancing faster than the ability of developers to evaluate the response of these integrated systems to natural (e.g., droughts, earthquakes, and floods) or deliberate (e.g., cyberattack and terrorism) threats. Due to the inherent design conflicts, no system can be fully smart and resilient. However, by balancing resilience and smartness within systems, transportation systems can increase the likelihood that system maintain functionality without compromising significant efficiency and sustainability especially during disruptions. There is no textbook solution to this optimization, and tradeoffs between smartness and resilience vary on a system-by-system basis. Some systems may benefit more from smart development and less from resilience (e.g., minor urban streets). Other systems may place a greater value on resilience if the disruptions are expected to yield high intensity loss (e.g., cities with critical bridges and tunnels), or if the operational decisions involve conflicting outcomes (e.g., prevent a disruption at a school or a hospital). Through a better understanding of the tradeoffs between smartness and resilience, and application of network science, system engineers can work to maximize efficiency and quality of life, and also minimize loss due to disruptions.

Limitations and Future Work. It is important to note, that because we consider only urban areas classified by the U.S. Census Bureau as either large or very large, the model should be validated using traffic from smaller urban areas before its application to those cities, as the travel patterns in small urban areas may differ from those in larger ones. Moreover, it may be necessary to include residential and more minor roads into consideration as these roads' role increases with the decrease of the size of the area analyzed. We believe this verification to be an important step in the further development of the model, but leave it out in this section.

The model presented in this study abstracts many components. A more complex environment could improve accuracy, but such advanced models also require much more detailed data. Examples of these data include the exact numbers of lanes and speed limits for all links, traffic lights locations and schedules, the exact temporal distributions of commuters' travel, fine grained spatial distributions of commuters' origins and destinations, considering

job sites, and, most importantly, the variability in the accessibility and the public perception of other modes of travel (such as public transportation, cycling, or walking). Some of this data may be less costly to obtain and integrate into the model.

One potential improvement is in the trip generation model. As of now trip origin and destination pairs are generated based with a gravity model using population and distance. This simplification is useful in scenarios where more specific data may not be available or other methods too costly to justify. Some of the cities included in this study, however, may have more specific data that can be easily incorporated, in their Household Travel Surveys (HHTS). A method, such as used in Chen et al. (2013), that incorporates this data, when available, would improve the results for many areas.

The simplified disruption model could also be relaxed. Currently all disrupted nodes reduce connecting link performance uniformly. This simplification acts to model an upper bound of possible disruption. In reality, different attack options will affect different node types and connecting links differently. Setting all lights to red turns an intersection after initial confusion, to a stop sign. Another attack would disrupt light timing so that vehicles face an increase in total red light timing when not turning or only allowing one direction of traffic have a green light (Petit and Shladover 2015). Each of these attacks will have different disruption effects. Including this complexity may improve model accuracy, at the cost of potentially ignoring unknown attack vectors.

3   Small-Scale Models

This project considered several small-scale models of Command and Control (C2) systems and architecture, primarily in the context of masters theses conducted by students at the Naval Postgraduate School. We summarize the work of each, in the order they were conducted.

3.1   A First-Principles Perspective

In the thesis of Cohick (2018), we consider the design of C2 architecture from the perspective of operations research. We investigate how one measures the performance of an organization in conducting a mission, and we contrast various C2 architectures in their ability to accomplish complicated tasks.

This research is inspired by retired Army General Stanley S. McChrystal's book titled, *Team of Teams: New Rules of Engagement for a Complex World* (McChrystal et al. 2015). McChrystal explains the challenges he faced after taking command of the Joint Special Operations Task Force (JSOTF) in Iraq beginning in September 2003. Despite being superiorly trained and equipped, his forces were unsuccessful in combating Al-Qaeda in Iraq (AQI), a very aggressive and unpredictable foe. The book details the transition of the JSOTF from a traditional military organization, reliant on planning and rehearsal, into a flatter, more integrated, and team oriented architecture that was able to aggregate information and adapt to external events faster than the enemy.

The story told in *Team of Teams* is relevant to both military and civilian organizations. Military units and civilian corporations alike are comprised of people, internal processes, and reporting relationships that were designed, or have evolved, to accomplish a mission.

A key lesson in *Team of Teams* is that the C2 organization and methods that work well in one environment for a given mission might not work well in a different environment and/or for a different mission. The implication is that an organization that is not adapted to its environment is bound to perform poorly. The story of the JSOTF's organizational transformation to overcome the challenges posed by AQI is presented as an exemplar for how modern organizations ought to adapt in the face of growing operational complexity.

Table 7 highlights characteristic differences between JSOTF and AQI.

Table 7. Contrast of JSOTF and AQI. (Adapted from Daft 2012)

JSOTF (Mechanistic)	**AQI (Organic)**
Centralized Structure	Decentralized Structure
Strict Hierarchy of Authority	Collaborative Teamwork
Specialized Tasks	Empowered Roles
Vertical Communication	Horizontal Communication
Many Rules, Formalized	Few Rules, Informal

In response to the JSOTF's lack of effectiveness combating AQI, McChrystal implemented major changes to its architecture and culture to increase adaptability in its new complex and rapidly changing environment. Large civilian organizations alike are changing their organizational architectures in order to remain relevant and profitable. But how does one know if an organization is appropriately adapted for a given mission and/or environment? Moreover, is it possible to design an optimal C2 architecture? And how can the field of operations research help to understand these issues?

3.1.1 Model

We develop a mathematical framework to support analysis of simple word matching game, designed as a highly stylized version of an intelligence gathering operation. Using the lexicon defined by Fullerton (2014), we present the game's characteristics:

Components. The components of the game are simply the players and resources available to each player. Each player represents one member in an organization.

Space. The resources come from the game's space (an organization's environment) which we implement using a distribution of letters A through Z. During every turn, each player receives one letter from the distribution which we can vary, thus changing the organizational environment in which we examine the performance of a specific architecture.

Goals. The overall goal of the game is for an organization to collect and assemble the required letters to match a specific objective. The objective can be partitioned into small tasks and assigned to specific members in the organization. Each player is assigned a primary (and possibly a secondary) task which they strive to accomplish using the letters

they receive. Like the environment, we can vary the organization's objective in order to measure an architecture's performance under different stresses.

Mechanics. The game's mechanics are defined by the specific organizational architecture being examined. Each architecture will have different connections between members, reporting relationships, and collaborative processes. We can change the game's mechanics in order to analyze different architectures.

Rules. Game rules direct player actions. They dictate what each player will do each turn according to the architecture being examined. The number of discrete simulation time steps it takes the organization to match its objective is the output (score) of the game.

Using a Discrete-Time Markov Chain to represent play in any specific game, we model a series of random transitions from one *state* to another. For our word matching game, we define the state space in terms of the number of letters that have been obtained. Thus, the game starts with an initial state (no matches) and concludes when the system reaches a terminal state (matching complete). At every point in the game, we assume the probability of the next transition depends only on the current state (i.e., the process is memoryless). Time advances in discrete steps and the *state space* for the game is the set of all possible states. The transitions between states depend in general on the number of players and connections between them as determined by the C2 architecture being examined.

Players in a game can be working independently or in coordination. We define the organizational architectures we seek to explore via mathematical analysis and simulation in terms of the relationship between two types of players.

Worker. A worker is a player in the game that receives an input of letters from the environment (distribution of letters). We illustrate workers using a red rectangle. Each is assigned a task (sub-component of the organization's overall objective), annotated in the rectangle. Workers can only communicate or collaborate with other players if connected by a linkage in the architecture. Workers have no subordinates.
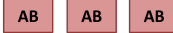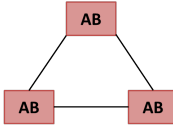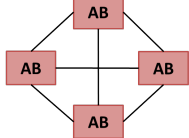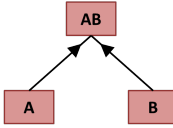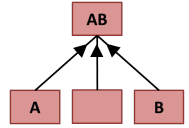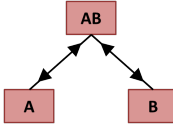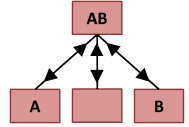
Manager. A manager is a player in the game connected to one or more subordinate workers. Managers are illustrated with blue rectangles and do not receive an input of letters from the

environment. The manager's task is to collect completed tasks from workers and match the organization's overall objective, depicted in the rectangle.

We consider different hierarchical structures for coordination, and evaluate the performance in terms of the expected number of discrete time steps to complete a complete match. We assume the frequency with which letters are found by workers follows the letter distribution found in 40,000 words from the English dictionary (Cornell Department of Mathematics 2004).

We present Table 8 to summarize the architectures we have examined. The bracketed letter following each result annotates the method used to calculate the expected value (A = Analytic equations; S = Simulation).

Table 8. Architecture Performance Summary

	1 Agent	2 Agents	3 Agents	4 Agents
Independent Agents	69.0536 [A] 68.7626 ± 1.2761 [S] We begin with the simplest architecture consisting of one agent.	36.4587 [A] 36.6198 ± 0.6240 [S] Doubling the number of agents reduces the expected value nearly in half.	26.3074 [A] 25.8209 ± 0.4025 [S] More agents yield better performance, but benefit begins to taper off.	20.9992 [A] 20.9587 ± 0.3104 [S] Additional agents continue to improve performance, but it does so less and less.
Fully Connected Agents		34.7854 [A] 34.8019 ± 0.6404 [S] Connecting two agents allows them to share letters, thus improving performance compared to two independent agents.	23.2637 ± 0.4259 [S] Connecting three agents continues to decrease the expected value. Realistically, now sharing rules are needed to avoid creating copies of letters. Time deciding how to share, is time an agent is not matching new letters.	17.7431 ± 0.3175 [S] Using simple sharing rules, additional agents result in improved performance but the rate of improvement slows.
Hierarchy One-way Comms			69.3888 ± 1.2861 [S] With one-way communication, performance is nearly identical as a single agent architecture.	A = 67.1308 ± 1.3182 [S] B = 36.6723 ± 0.6139 [S] Assigning "A" or "B" to the added worker makes a huge difference. When the environment is known, tasks can be assigned to optimize performance.
Hierarchy Two-way Comms			34.9463 ± 0.6484 [S] The expected value is nearly same as two connected agents since the manager can share letters with other agents.	A = 23.5521 ± 0.4289 [S] B = 23.6020 ± 0.4365 [S] With two-way communication, the architecture's sensitivity to task assignments is much less, thus making the architecture more robust.

3.2   A Project Planning Perspective

The thesis of Beavers (2019) considers the C2 organization and function of the staff of dedicated men and women serving within any of the Navy's maritime headquarters.

As the theater combatant commander's maritime warfighter, the Navy fleet commander manages theater maritime assets to prepare and provide forces, and employs those naval forces to support and perform operational missions. To execute these duties, each of their maritime headquarters (MHQ)s is designed to be appropriately manned, trained, and equipped to support both functions; even as the commander is assigned various responsibilities to conduct enduring and contingency missions as a Naval Force (NAVFOR) Commander, Joint Force Maritime Commander (JFMCC), Coalition Force Maritime Component Commander (CFMCC), or even support for joint task force (JTF) operations.

Within the MHQ staff, a separate maritime operations center (MOC) bridges the command and control (C2) gap between strategic level guidance and tactical employment of forces. These MOC staffs perform tasks that provide the commander with the organizational structure, procedures, and expertise support needed to make and communicate informed decisions, while setting the conditions for subordinate success; separate from the MHQ's fleet management mission. However, even with well-defined requirements in place, Manpower, Personnel, Training, and Education (MPT&E) impacts still pose potential risks to a MOC's ability to effectively support the commander's C2 role. On a practical level, it may not be possible for all staffs to be sufficiently equipped or manned to designed levels at all times, and en route training timelines could either cause gaps in billet fills, or delays in training for newly reporting personnel. The potential cumulative effects of these impacts could result in a reduced, but highly dedicated staff, capable of accomplishing its day-to-day missions, or even contingency operations for a limited time. However, this level of performance would require the consolidation of multiple MOC position responsibilities, and the leveraging of in-house experience as a replacement for proper training and education. Ultimately, this model could prove unsustainable, as the increased operations tempo of contingency operations stress the MOC's current manpower beyond the limits of what their efforts are able to mitigate, and leads to either delays in task completion and/or impacts to the quality of support they are able to provide.

This thesis develops a logic for objectively measuring the required effort a staff member

must exert to accomplish tasks within a time period, and provides a way to gain insights into the relative quality of work and risk assumed by them to do so, given the current manning and experience levels of their staff members. We focus on a method for considering staff activities, conducted under the constraints of limited resources by a versatile workforce, as project scheduling problems for task-on-node networks. To create those networks that represent the functions of a MOC, we describe a procedure for converting the supporting tasks that make up the mission essential tasks (METs) a MOC is assigned, into a collection of network nodes with precedence relationships, expected durations, and resource requirements in the form of the functional team-leads needed to work tasks.

3.2.1 A Project Scheduling Perspective

We explore current techniques for representing and solving resource-constrained project scheduling problems, in search of a model that best describes the "real world" aspects of a staff, their work, and the MPT&E factors that impact them. In reviewing the merits of the Critical Path Method (CPM) and the Project Evaluation and Review Technique (PERT) for identifying the tasks who most affect the completion time of a network, we ultimately dismiss them as necessary, but not sufficient because they do not account for resource requirements or constraints. To better suit the level of complexity a model must account for to represent staff activities, the family of Resource Constrained Project Scheduling Problems (RCPSPs) is explored. Many variations to the RCPSP exist, each adding a level of variability with the aim of modeling a more general scenario; including multiple combinations of resource types to complete the same task, and the simultaneous consideration of multiple projects. Each variation has its own set of solution methods as well, with each technique leaning farther from finding optimal answers and more towards "good enough" as complexity increases.

The Project Management Institute's Project Management Body of Knowledge defines a *project* as "a temporary endeavor of tasks undertaken to create a unique product, service, or result which may be subject to precedence and require resources" (Snyder 2013). Examples of projects exist in all aspects of the economy, including manufacturing, construction, repair and maintenance, design and development, and government programs. In general, all projects consist of *tasks* that must be accomplished according to a schedule constrained by precedence relations among those tasks, associated resource requirements for their execution, and an overall goal of minimizing total project cost; cost being in relation

to either monetary, temporal or any other consideration deemed important to the project manager.

According to Vanhoucke (2012), project managers have utilized formalized methods for organizing, scheduling, and controlling projects to best achieve their stated goals, since at least the late 1950s. He further explains that the earliest of these approaches are powerful but simple techniques for analyzing, planning, and scheduling large, complex projects. Known as the *Critical Path Method (CPM)* and *Project Evaluation and Review Technique (PERT)*, Vanhoucke (2012) describes how these tools provide a means for determining which tasks, of the many that comprise a project, are "critical" because of their impact on total project time and/or cost.

Gen et al. (2008) illustrates that, while CPM and PERT are exceedingly popular in industry due to their simplicity in construction and relative ease in determining exact solutions, they do not directly take resource requirements into account. They explain that this is both unrealistic and potentially costly to the project manager who may have to consider variability in resource availability and/or even executing multiple competing projects. Gen et al. (2008) goes on to define the problem of scheduling tasks under resource and precedence constraints, with the objective of minimizing the project duration, as a *Resource-Constrained Project Scheduling Problem (RCPSP)*. These resource constraints, as referred to by Gen et al. (2008), are limited renewable supplies, such as manpower, materials and machines, which are necessary to completing project tasks. They go on to state that many variations of the RCPSP model have been developed and examined, all with the aim of more accurately modeling "real world" situations. As explained by Gen et al. (2008), these variants incorporate concepts such as multiple task duration-resource requirement options for a single task (known as *modes*), resources that are multi-skilled and interchangeable, or multiple projects that compete for the same resources but are conducted simultaneously. Kolisch (1995) goes on to elaborate that each variant of the RCPSP serves to create a more accurate model than that of the CPM, PERT or basic RCPSP; however, their utilities are diminished given that with each added level of complexity, the computational effort to find an optimal or even feasible solution for the project schedule grows exponentially.

The following sections provide necessary details of CPM, PERT, and RCPSP models.

3.2.2 CPM and PERT

Both the CPM and its extension PERT were first developed in the 1950s and made popular in industry by their use in the U.S. Navy's Polaris Submarine-Launched Fleet Ballistic Missile program (Kolisch 1995). These techniques assume a project possesses unconstrained resources and represents it as a directed graph network without loop cycles. Figures 31 and 32 show a network of six tasks for analysis by the CPM and PERT methods, respectively. They use a task-on-node diagram, with nodes describing project tasks and arrows displaying the precedence relationships among them. Additionally, each network has a start (ST) and an end (EN) node, with the values above each numbered node representing task duration.

Task duration is where PERT analysis diverges from the CPM. While task duration for the CPM is deterministic, PERT attempts to address time uncertainty in projects by assuming that each task duration is a random variable between two extreme values which follows a beta distribution (Vanhoucke 2012). The overall estimate is then calculated as a weighted average of the optimistic, most likely, and pessimistic estimates, with more weight given to the most likely value.
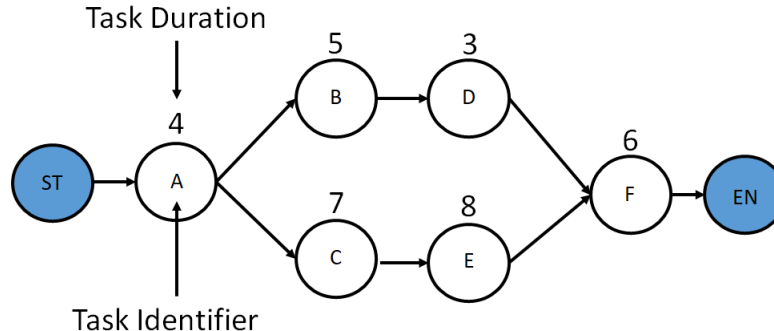
Figure 31. Critical Path Method. As appears in Beavers (2019), adapted from Pinha (2015).

The goal of CPM/PERT, and for that matter all project scheduling efforts, is to determine the tasks within a network of precedence relationships that cannot be delayed without delaying the duration of the entire project. Each of those tasks are referred to as a *critical task* and their relations within the project's network as the *critical path*. Vanhoucke (2012) gives a detailed explanation and practical example for how CPM/PERT problems can be solved to identify these critical tasks and paths, which he summarizes into three generalized steps:
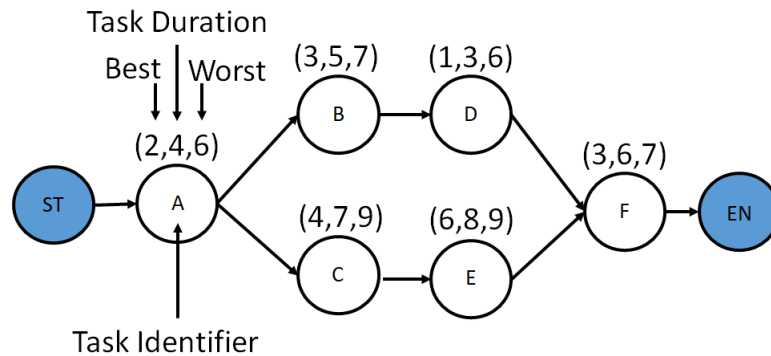
Figure 32. Project Evaluation and Review Technique. As appears in Beavers (2019), adapted from Pinha (2015).

1. Calculate the earliest start schedule
2. Calculate the latest start schedule
3. Calculate the slack for each task

Based on the work of Vanhoucke (2012), we introduce some basic notation to facilitate the calculation of each of these below. Let $j \in A$ denote the tasks in a given project (alias k). Let d_j denote the *duration* of task j. Let P_j and S_j denote the sets of *predecessor tasks* and *successor tasks*, respectively, for task j.

According to Vanhoucke (2012), a tasks's *earliest start time* is either equal to or larger than the earliest finish time from among all of its predecessors. The *latest finish time* is determined to be either equal to or less than the latest start time from among all of its successors. Finally, the amount of *slack* associated with each task represents the amount of time each task j can be delayed without affecting the duration of the entire project. A task with a slack of zero cannot be delayed without affecting the entire project's duration. Vanhoucke (2012) refers to such a task as a *critical task*, where the *critical path* is the path along a network that consists of these critical tasks. For Figure 31, nodes ST, A, C, E, F, and EN have a slack value of 0 and therefore, constitute the critical path for that network.

While simple forms of analysis, the CPM and PERT tools help to draw the focus of a project manager to the most consequential activities, providing an effective basis for the scheduling and monitoring of progress without the consideration of resources.

3.2.3 Resource Constrained Project Scheduling Problems (RCPSP)

The basic Resource Constrained Project Scheduling Problem (RCPSP) was first introduced by Pritsker, Watters, and Wolfe in the late 1960s (Pritsker et al. 1969). This simple version consist of interrelated tasks, each with a single associated time of duration and resource requirements pair. Figure 33 depicts the task-on-node network representation of a single six-task RCPSP project. Each task, nodes A through F, shows their single duration-resource requirements pair, which is predetermined and fixed in a RCPSP.
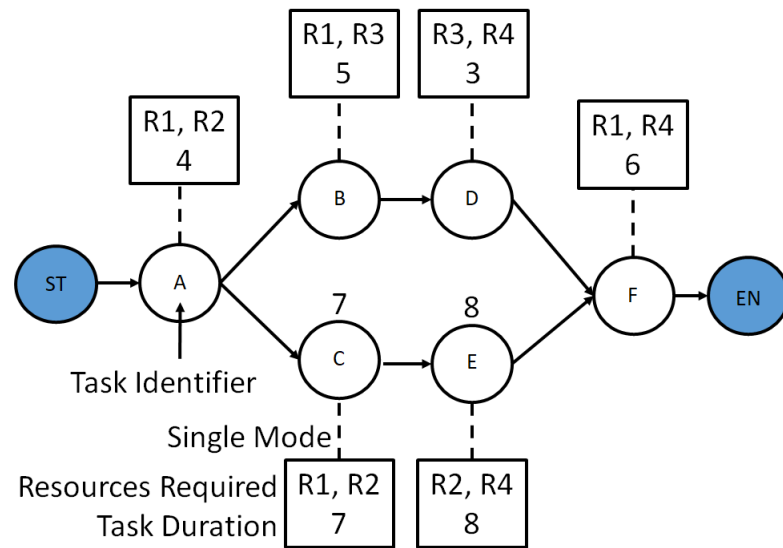
Figure 33. Resource Constrained Project Scheduling Problem (RCPSP). As appears in Beavers (2019), adapted from Pinha (2015).

Several variants of the RCPSP have been proposed since the initial work of Pritsker and his colleagues. As explained by Gen et al. (2008), these variants can be divided into categories based on their problem characteristics, i.e., the number of projects, the number of task-duration/resource requirement options available, resource consumption, and the ability to stop a task's execution before its completion (interruption) (see Figure 34). Gen et al. (2008) also states that RCPSP models can be labeled as either *single-project*, where there is only one project to be scheduled, or *multi-project*, where there is more than one project to be scheduled and executed.

According to Gen et al. (2008), RCPSP models can also be divided into two more groups based on the number of task duration-resource requirement pairs available to perform a task. These various alternative combinations are referred to by Gen et al. (2008) as

111

modes. Per Kolisch (1995), in the single-mode version, there exists only one task to resource paring capable of executing a task, whereas in the multiple-mode version, each task can be performed by selecting one out of many different combinations of resources that can complete a given task. Kolisch (1995) goes on to state that the resources in RCPSP models may also be considered as "limited-but-renewable" from period to period, "non-renewable," "of-limited-amount" for the duration of the project, or doubly constrained by both restrictions. Finally, Gen et al. (2008) explains that RCPSP models can again be divided into those that are *preemptive*, where a task can be interrupted after it has started, and those that are not (i.e., *non-preemptive*). However, the main body of work in the development of RCPSPs center on the non-preemptive case of its two main variants and their combination; the *Multi-mode Resource Constrained Project Scheduling Problem (MRCPSP)*, the *Resource Constrained Multi-Project Scheduling Problem (RCMPSP)* and their combination, the *Multi-mode Resource Constrained Multi-Project Scheduling Problem (MRCMPSP)* (Wauters et al. 2016).

Figure 34. Classification of Resource Constrained Project Scheduling Problems (RCPSPs). As appears in Beavers (2019), adapted from Gen et al. (2008).

The RCPSP, and its more complex extensions, are considered to be the standard approaches to project scheduling, but belong to the class of strongly NP-hard problems (Kolisch 1995). The majority of the solution methods to these types of problems tend to be either exact, or some form heuristic. Exact procedures strive to find the optimal solution, often through the enumeration and evaluation of all possible solutions. These procedures, however, are only useful for solving small projects that have undergone several simplifying assumptions.

Heuristic methods, on the other hand, search for a good enough but not necessarily optimal scheduling solution, using more realistic restrictions and assumptions (Vanhoucke 2012). Still, both exact and heuristic methods are limited in their efficiency for solving the complex models that more accurately represent "real world" problems. Therefore, many researches attempt to capture more realistic model restrictions through discrete event simulations that make use of priority rules (Pinha 2015). These priority-rule-based simulations are very popular among researchers for solving the more complex RCPSP extensions due to several reasons: they achieve a satisfying quality of solutions, they are computationally fast, and they proceed similar to the way decisions are made by project managers (Kolisch 1995).

3.2.4 Combinatorial Multi-Mode Resource Constrained Multi-Project Scheduling Problem (CMRCMPSP)

As best stated by Pinha (2015), the chosen representation for a problem determines its complexity and the size of the search space in which a usable solution exist. For instance, a highly specific model may significantly reduce the breadth of the solution search, and could even be relatively simple to solve, but may only apply to a single or limited range of problem instances.

To better represent the complex nature of "real world" situations and provide meaningful insight to problem solutions, models must possess the flexibility to consider multiple projects, with several tasks competing for limited resources that are themselves dynamic. Until recently, the MRCMPSP was the most general of resource-constrained scheduling model extensions cited in literature (Pinha et al. 2016). It allows for the simultaneous consideration of tasks from several projects, all of which are under precedence and resource constraints, with multiple modes available for each. For illustrative purposes, Figure 35 shows the six-task project discussed in previous examples, now with multiple modes available for each task and the inclusion of four resources (denoted R1 through R4). For example, task C has three modes; mode 1 calls for resources R1 and R2 for a duration of 7 units of time, mode 2 requires R2 and R3 for a duration of 8 units of time, and mode 3 needs R1 and R4 for a duration of 10 units of time.

As illustrated by Pinha (2015), multiple modes provide added flexibility to develop project schedules. In the example given in Figure 35, the duration of task C increases with the order of the mode option. However, this is not a strict limitation. Depending upon the

requirement, the mode order may be analogous to preferences for some value other than duration—such as quality, materials, and cost—that the project manager uses to determine the number and order of modes for each task. As Pinha (2015) goes on to explain, these preferences are also independent of resource requirements by other modes. For instance, if mode 1 requires more resources than any of the other modes, that does not necessarily equate to a shorter duration for mode 1. On the contrary, its duration will vary depending on its process, not it resource requirements. Conversely, less equipment with fewer laborers may be more efficient than the resources mode 1 calls for, but they may also be more costly, produce products of a lesser quality, or otherwise result in an outcome that is less desirable than the application of the resource requirements for mode 1.
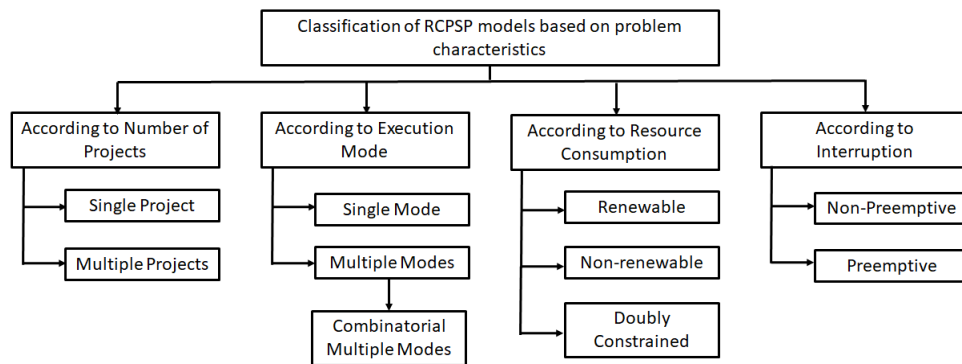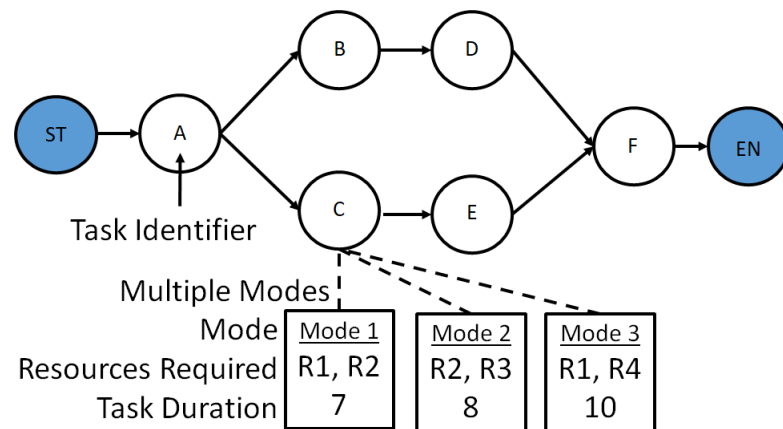

Figure 35. Multi-mode Resource Constrained Project Scheduling Problem (MRCPSP). As appears in Beavers (2019), adapted from Pinha (2015).

At its heart, the consideration of a MOC's operations and the risk associated with any under-resourcing of staff levels, training, etc., is a human resources problem. In the context of a project, human resources have multiple skills, with a certain level of proficiency in that skill, or skill level, which may or may not affect the overall performance of a task. While the MRCMPSP model can accommodate flexibility in the application of multiple resources, very few approaches are capable of handling multi-skilled workers. The Combinatorial Multi-mode Resource Constrained Multi-Project Scheduling Problem (CMRCMPSP) developed by Pinha (2015) is an extension of the RCPSP that expands its execution mode category (see Figure 34). It allows for several project tasks to be handled simultaneously, under precedence and resource constraints, with multiple modes available for each task. It also enables the consideration for the skill level of the resources at hand, which may be made

analogous to the level of risk assumed with assigning a resource to perform a given skill for a task.
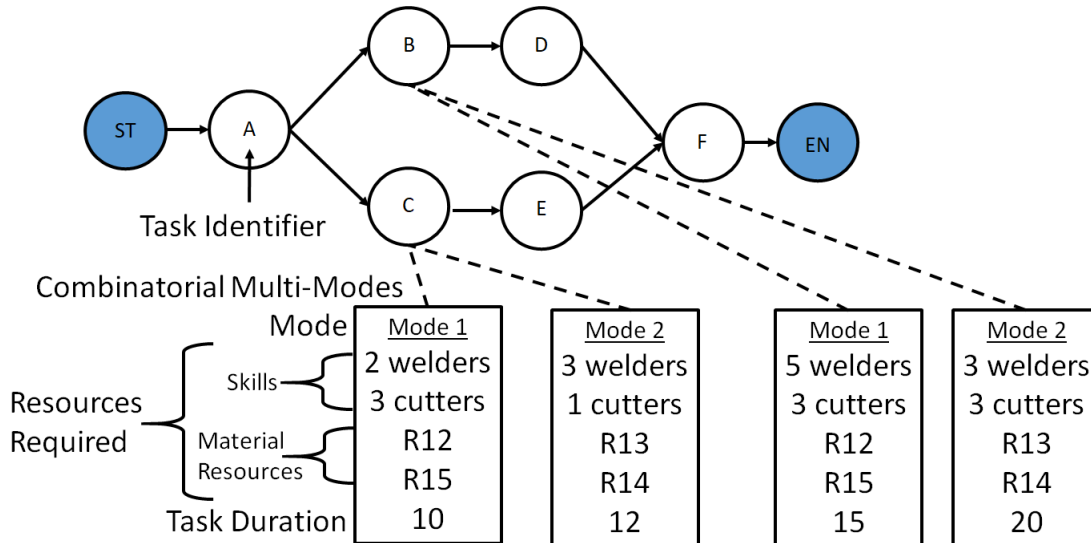
Figure 36. Combinatorial Multi-mode Resource Constrained Multi-Project Scheduling Problem (CMRCMPSP). As appears in Beavers (2019), adapted from Pinha (2015).

The CMRCMPSP differs from the MRCMPSP in the way it considers resources. In a CMRCMPSP, the quantity of modes does not solely define the multitude of options available to perform a task, and task performance may or may not depend on the resources selected. Per Pinha (2015), a mode in a CMRCMPSP is no longer a set of resources, but a set of combinational subsets of unknown resources who possess the skills needed to execute a given task but who also possess a certain level of proficiency in those skills. Figure 36 shows the previously discussed six-task project, now with both skill and resource requirements for each task and multiple modes to illustrate the CMRCMPSP's differences from the MRCMPSP.

3.2.5 Our Contribution

We seek a representation of MOC staff operations that is specific enough to generate realistic solutions and reduce that solution search space to a manageable size. At the same time, that representation must also be sufficiently general enough in nature to allow for the flexibility needed to model the "real world" characteristics of human resources performing

115

administrative tasks. We settle on an adaptation of the Combinatorial Multi-mode Resource Constrained Multiple Project Scheduling Problem (CMRCMPSP) model, which allows us to capture the flexibility of a MOC staff in many ways. It accounts for a supervisor's ability to consider staff members as sets of combinations according to the skills they can perform, rather than by job titles. This mirrors the situations where a staff must tap personnel for the skills they possess without regard for the positions to which they are assigned; in the name of "getting the job done." The CMRCMPSP model also incorporates the various options in the way resources can be applied to complete a task, the need to conduct multiple parallel projects that must share a common pool of resources, the abilities of talented staff members whom are able to wear many positional "hats" while contributing to multiple concurrent efforts with varying degrees of proficiency (based on level of training, education, and/or experience), and a supervisor's preferences or organizational policies for the employment of resources which serve to resolve conflicts between competing task for those limited resources.

Using the Python programming language, we create a custom discrete event simulation algorithm to solve our adaptation of the CMRCMPSP model, which develops a schedule for the given projects by using the available resource pool. The Dynamic Resourcing Allocation Analysis Simulation Tool Kit, or DRAASTK (see Figure 37), works to start tasks as early as they are able to be executed, subject to the restrictions of user-defined preferences.
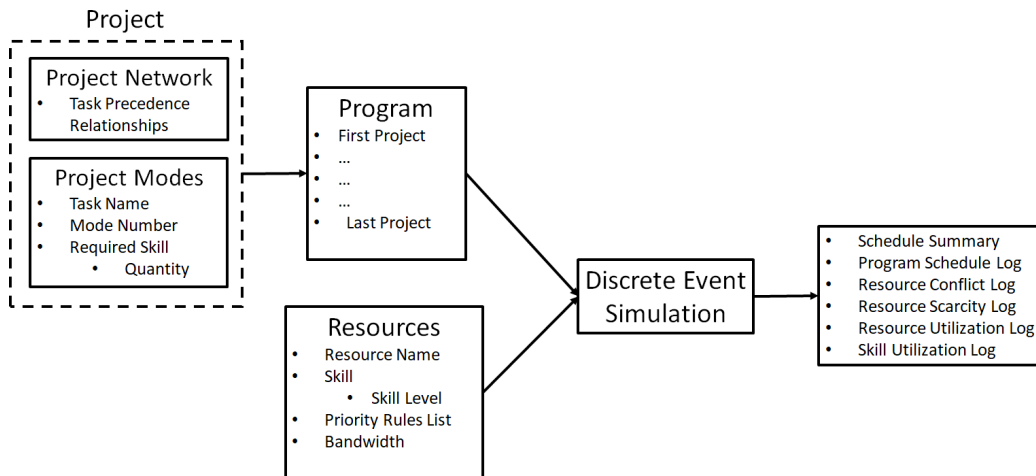
Figure 37. Dynamic Resourcing Allocation Analysis Simulation Tool Kit (DRAASTK) Overview Diagram

The Dynamic Resourcing Allocation Analysis Simulation Tool Kit (DRAASTK) accepts inputs for the project networks and modes that compose a program, as well as the pool of resources available to service the program. It then uses those inputs in a discrete event simulation to provide outputs that that give insights into project schedules and resource allocation and utilization.

Simulation outputs include information for both the task-resource parings and resource usage necessary to achieve the provided schedule. Along with a recommended schedule for task completion, a record for each delayed task, and the lacking required skills that caused its delay at the time of its evaluation, is provided to give insight into the organization's skill deficiencies and potential points for where additional resources should be focused. As for the current resource utilization and risk, a record for the percentage of time each staff member is involved with a task is taken; broken down by how much of that time is spent engaged with either a single or multiple simultaneous tasks as a measure of that staff member's exerted effort. The level of proficiency for which a resource holds for the skill applied to a task is also noted and provided. A subjective value based on the training, education, and experience level a staff member holds for a particular skill, the collective average for the level of proficiency of all staff members whose skills were utilized to complete a task is analogous to the relative level of risk in quality and performance a staff assumes to achieve the provided timeline.

As a resourcing decision aide, this model provides many opportunities for additional work, in

both data and simulation refinement. In particular, in order to consider MOC staff operations as networks to analyze, significant subject matter expert (SME) input is required to apply precedence relationships and expected durations to all supporting task and procedures. A MOC organizational review and development of these networks is an important first step in evaluating MOC staff operations. Developing granularity in the resource pool is also a viable area for expansion. Currently DRAASTK considers skills as the positions of team leaders common to all MOCs, but this can be evolved to include the duties and skills of each team member and their individual availability schedules, allowing for a more robust examination of a specific MOC staff. Finally, future work can focus on the flexibility of the simulation itself. Incorporating stochastic elements for initial and/or resource-based task durations, as well as consideration for individual resource availability, all works towards developing a more "real world" model and method for gaining viable insights to inform and aide decision makers with complex resource allocation resolutions.

Our method for paring a limited pool of skilled resources with the tasks that require them and recording the points of resource-induced delays and relative quality of skill in task completion are the main contributions of this thesis. Compared to a staffing profile inline with a MHQ's intended manning, the schedules and outputs generated by DRAASTK give insights into the amount of effort and risk to quality and timeliness an actual staffing profile exhibits. This information provides a quantifiable basis to articulate the existing MPT&E impacts and risks to mission that a staff experiences, and may inform the conversation needed to justify the resourcing requirements of a staff that continues to succeed, despite "doing more with less."

3.3 An Assignment-Based Perspective

The thesis of Diaz (2019) frames C2 design as a type of assignment problem between *tasks* (i.e., units of work to be done) and *resources* (i.e., personnel who perform work). As the number of tasks and resources get large, the assumption is that tasks are organized into *missions* and that resources are organized into *teams*. We formulate and solve a sequence of four different optimization problems that assign resources to tasks, missions to teams, tasks to missions, and resources to teams. Our models are based on the assumption that any resource can execute any task; however, the cost for a resource to execute a task depends on the suitability of the resource for that task. We characterize the type of work required by a

task in terms of a functional requirement, and we characterize the capability of a resource in terms of its resource specialization (e.g., Military Occupational Specialty). We also include nonlinear effects based on the interactions within a team and across a mission. Our goal is to understand the performance effects of optimal teams on the organizational architecture of a C2 network.

We compare all four assignment models using large and realistic data sets that demonstrate their distinct effectiveness and recommendations. We show that assigning tasks to resources without considering military missions or teams can efficiently produce a globally optimal task-resource assignment. However, implementing this assignment in practice is difficult because it requires a flat organizational architecture with high management overhead. Instead, models with predefined missions and teams simplify the computational effort of finding a good assignment, but often at the expense of a solution that is globally sub optimal. Models that form adaptive teams and/or mission sets can achieve better assignment and lower execution costs than current C2 architectures, but at the expense of significant increases in computational time. Thus, each model provides a new perspective on what teaming means while setting baselines for computational requirements, optimality, robustness, adaptability, and flexibility. Our results demonstrate that optimal teaming is data-, context-, and mission-dependent.

We conclude that organizations with different C2 architectures should consider different models to guide task-resource assignment decisions. Extremely flat organizations can benefit from the globally optimal solution found with the simple Task-to-Resource assignment model. However, current military force organization is not flat enough to harness these benefits, and missions and teams must be considered in assignment. Models with predefined missions and teams should only be used when task-resource decisions are time constrained. Instead, when faced with changing and unforeseen mission environments, military organizations may want a team-of-teams approach and would benefit from the adaptability of an assignment model that includes team formation. Moreover, when organizations are rigid but tasks can be redistributed to synchronize missions, the optimal assignment is best provided by the mission assignment model.

4 Large-Scale Models

The material in this section is taken from Kitsak et al. (2018) and also includes some work that is still under development.

4.1 Stability of a Giant Connected Component in a Complex Network

Robustness and resilience of networked systems under the impact of adverse events have been extensively studied in network science for two decades, but the research has been primarily focused on computing the mean-field properties, such as the expected size of the system's giant connected component (GCC) (Stauffer and Aharony 1994; Dorogovtsev 2010; Cohen and Havlin 2010). The inherent assumption is that only connected subnetworks retain their functionality, with the largest of these being most relevant to the overall performance of the system.

While the mean-field analysis of a system's behavior is undoubtedly an important first step toward understanding its robustness, in most practical situations it is insufficient to know the expected size of the so-called 'functional component.' Rather, the location of the functional component within the network itself is important. It is especially true in the case of resilience where critical system function and its recovery is of prime importance (Ganin et al. 2016; Linkov and Palma-Oliveira 2017a). For example, in the case of a major natural disaster, such as flood or an earthquake, one needs to know infrastructure units and transportation routes that are likely to remain functional. The efficiency of immunization strategies depends on our knowledge of low and high-risk groups in social networks. Likewise, the success of a marketing campaign depends on the knowledge of the target audience.
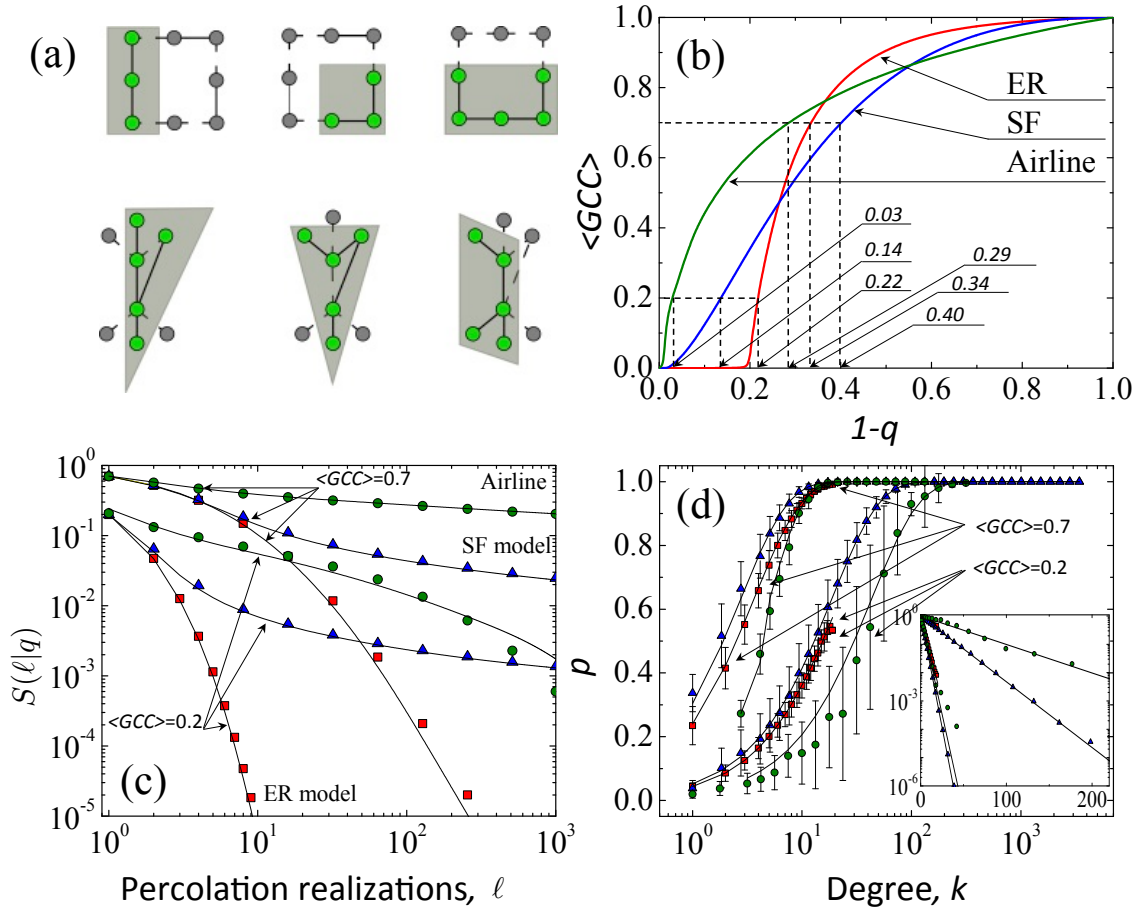
Figure 38. (a) GCCs shown for two toy networks at various links percolation realizations. Note that the GCC of the top network varies significantly depending on the set of removed links, while the GCC of the second network is more stable, and in all realizations includes the two central nodes. (b) The expected GCC size for an ER and a SF models, as well as the airline network as a function of the fraction of deactivated links q. Both model networks are of the same size, $N = 2 \times 10^5$ and the same average degree $\langle k \rangle = 5.1$. The SF network is characterized by $\gamma = 2.5$ and minimum degree $k_{\min} = 2$. Dashed lines show considered values of q that are chosen such that $\langle GCC \rangle = 0.2$ and $\langle GCC \rangle = 0.7$. (c) Stability $S(\ell|q)$ as a function of ℓ for the (red squares) ER model, (blue triangles) SF model, and (green circles) the Airline network described in Appendix 4.1.1. Solid lines correspond to analytical predictions by Eq. (31). (d) Probability for a node to appear within the GCC of a network as a function of its degree. The color scheme is that of panel (c). Data is binned logarithmically, and error bars display the standard deviation values. The inset shows the probability of the node *not* to appear within the GCC in the log-linear format with solid lines corresponding to analytical solutions given by Eqs. (29) and (30).

Yet, apart from a handful of recent works aiming to understand individual node properties in percolation and epidemic processes (Bianconi 2017; Hamilton and Pryadko 2014; Morone and Makse 2015; Rogers 2015; Kuehn and Rogers 2017), organization patterns of individual network's functional components are poorly studied. One reason is that the random nature of adverse events coupled with the complexity of relevant networked systems often makes the prediction of functional subgraphs impossible. Indeed, consider two toy networks of the same size and average degree, which we repeatedly subject to adverse events of equal magnitude, modeled by link percolation, i.e., by deactivating a fraction of links selected uniformly at random, Fig. 38(a). Even though the expected GCC sizes of both networks are similar, the GCC of the first network is unstable and strongly depends on the set of deactivated links. In contrast, the GCC of the second network seems to be centered at the two largest degree nodes and only weakly depends on the set of deactivated links. Clearly, nodes of the first network are topologically identical and, thus, are equally likely to enter the GCC. The two central nodes of the second network, on the other hand, have much higher probability to enter the GCC compared to the remaining nodes, serving as anchors for its GCC.

Our work is motivated by two recent results in percolation theory (Kuehn and Rogers 2017; Bianconi 2017) analyzing heterogeneous network responses to different percolation realizations. In particular, Kuehn and Rogers (2017) demonstrates a considerable variation of network's connected component sizes and individual node probabilities to appear within the GCC. Concurrently, Bianconi (2017) introduces a framework to quantify GCC fluctuations by analyzing network responses to two random but possibly correlated percolation realizations. In our work we ask a related question: How stable is network's GCC? In other words, we aim to quantify the extent to which the giant connected component of a network consists of the same nodes, regardless of the specific set of deactivated links during an adverse event. We analyze both single and multi-layer networks, finding that large GCC size does not necessarily imply GCC stability.

To simplify the exposition we model adverse events by link percolation (Stauffer and Aharony 1994; Bollobás and Riordan 2006) and limit our consideration to random network models with prescribed degree sequences. We also assume that there exists at most one GCC per network. At the same time we note that our approach is extendable to other types of percolation and more complex network topologies, including the situations with multiple

GCCs per network (Ben-Naim and Krapivsky 2008; Chen and D'Souza 2011). To quantify GCC stability we subject the network of interest to ℓ independent link percolation scenarios; each deactivating a random fraction q of network links. We then compute the sets of nodes $\Omega_t(q)$ for each percolation realization $t \in \{1, ..., \ell\}$ and use them to define the GCC stability function as the fraction of nodes inside all sets $\Omega_t(q)$, namely

$$ S(\ell|q) \equiv \frac{1}{N} \| \bigcap_{t=1}^{\ell} \Omega_t(q)\|, \tag{28} $$

where N is the network size and $\|\Omega\|$ is the cardinality of set Ω. In the case of GCCs consisting of random sets of nodes $S(\ell|q)$ decreases exponentially as a function of ℓ, while in the case of stable GCCs $S(\ell|q)$ is expected to decrease slowly or remain constant. We first focus on GCC stability of the Erdős Rényi (ER) (Erdős and Rényi 1960) and scale-free (SF) random network models. The former is fully prescribed by the number of nodes N and the constant connection probability p for every node pair. The latter is characterized by the scale-free (SF) degree distributions, $p(k) \sim k^{-\gamma}$ and is generated by the Molloy-Reed algorithm (Molloy and Reed 1995). Our results depicted in Fig. 38(b,c) indicate that GCC stability in ER is significantly smaller than that in SF networks. Indeed, $S(\ell|q)$ in the case of SF network models decays at significantly slower rates implying the existence of a stable sub-component, shared by all $\Omega_t(q)$ sets. At the same time $S(\ell|q)$ in ER network models seems to decrease exponentially, albeit at rates slower than expected for GCCs consisting of randomly selected nodes.

Our results for $S(\ell|q)$ are consistent with the observation that individual node probabilities to belong to the GCC are strongly correlated with node's degree. As seen from Fig. 38(d), the probability for a node to belong to the GCC increases as a function of its degree and approaches 1 for nodes of sufficiently large degree. These nodes belong to the GCC with high probability and serve as its anchors in a network, effectively stabilizing its GCCs. In contrast, the absence of such large degree anchor nodes in ER networks leads to unstable GCCs, Fig. 38(d).

To quantify the observed GCC stability phenomena we employ the generating function formalism (Wilf 2006). The starting point of our analysis is the mean-field solution for the

individual node probability to enter the GCC (Newman 2010):

$$p = 1 - [u(q)]^k,$$ (29)

where $u(q)$ can be regarded as the mean-field probability that a given end of randomly chosen link leads to a connected component of finite size (Callaway et al. 2000):

$$u = q + (1 - q)G_1(u),$$ (30)

and $G_1(x)$ is the generating function for the outgoing edge degree distribution, $G_1(x) \equiv \sum_{k=0}^{\infty} p_e(k)x^k$, where $p_e(k) = (k+1)p(k+1)/\langle k \rangle$, and $p(k)$ is the degree distribution (Newman et al. 2001).

Then, the probability that a node of degree k is present in the GCC independently in ℓ percolation realizations is $\left(1 - [u(q)]^k\right)^{\ell}$ and the expected stability of the GCC in ℓ independent percolation realizations is given by

$$S(\ell|q) = \sum_k p(k) \left(1 - [u(q)]^k\right)^{\ell}.$$ (31)

To validate our theoretical results we solved Eq. (30) numerically for each combination of $p(k)$ and percolation parameter q used in Fig. 38(b). We then used the obtained u values to superimpose the mean-field behaviors of individual node probabilities to enter the GCC and stability $S(\ell|q)$ with the numerical results, arriving at the excellent agreement between the two, Fig. 38(c,d).

The differences between $S(\ell|q)$ in SF and ER networks become more pronounced as expected GCC size decreases. Indeed, close to the criticality $S(\ell|q)$ in ER decays exponentially:

$$S(\ell|q) \approx \sum_{m=0}^{\ell} \binom{\ell}{m} (-1)^m \exp\left[-\langle k \rangle \left(1 - e^{-\epsilon\ell}\right)\right]$$
$$\sim \left[1 - u^{\langle k \rangle}\right]^{\ell},$$ (32)

where $\epsilon \equiv 1 - u \ll 1$ and $\langle k \rangle$ is the average degree in the ER network. In contrast, $S(\ell|q)$

in SF networks

$$S(\ell|q) \propto (\epsilon k_0)^{\gamma-1} \int_{\epsilon k_0}^{\infty} [1 - e^{-x}]^{\ell} \frac{dx}{x^{\gamma}} \sim \left(\frac{\epsilon k_0}{\ln \ell}\right)^{\gamma-1} \tag{33}$$

for $\ell \gg 1$. Here $\gamma > 2$ is the degree distribution exponent, $p(k) \sim k^{-\gamma}$, see Fig. 39 and Appendix 4.1.2 for derivation.
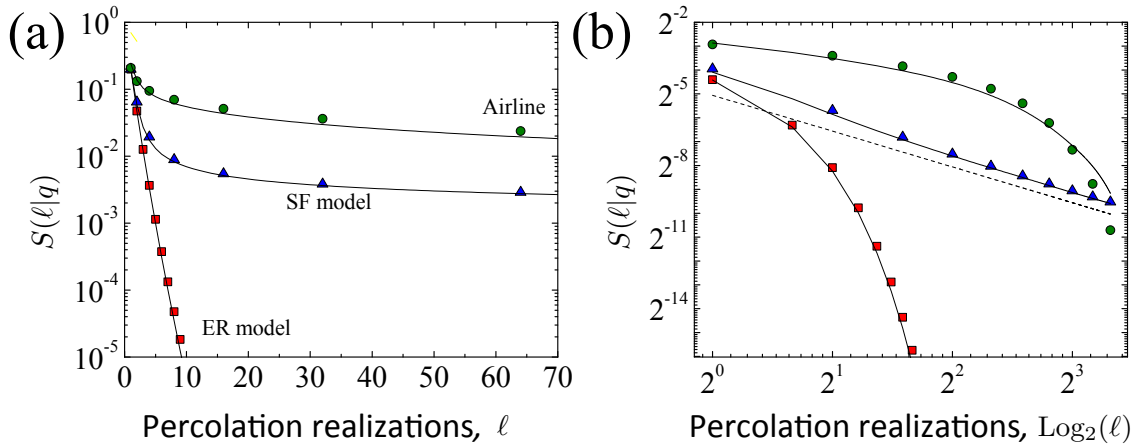
Figure 39. Stability $S(\ell|q)$ as a function of (a) ℓ and (b) $Log_2(\ell)$ shown at $\langle GCC \rangle = 0.2$ for the (red squares) ER model, (blue triangles) SF model, and (green circles) the Airline network. All parameters are those of Fig. 1(b,c) in the main text. The dashed line in panel (b) corresponds to the asymptotic regime of $S(\ell|q) \sim [\text{Ln } \ell]^{-1.5}$ prescribed by Eq. (6).

In the case of interdependent networks the deactivation of nodes or links in one layer leads to the deactivation of nodes in other layers that depend on it (Boccaletti et al. 2014b). In this case, the probability of a node to belong to the GCC depends not only on its own connections but also on the connections of supporter nodes in other layers. In the following, we consider a simple case of an interdependent network, consisting of equally sized layers A and B with one-to-one node interdependence; every layer representing a random network with prescribed degree distribution (Buldyrev et al. 2010). Assuming no correlations in interdependencies between the two layers, the probability of a node to belong to the GCC is

$$p = \left(1 - [u_A(\mathbf{q})]^{k_A}\right)\left(1 - [u_B(\mathbf{q})]^{k_B}\right), \tag{34}$$
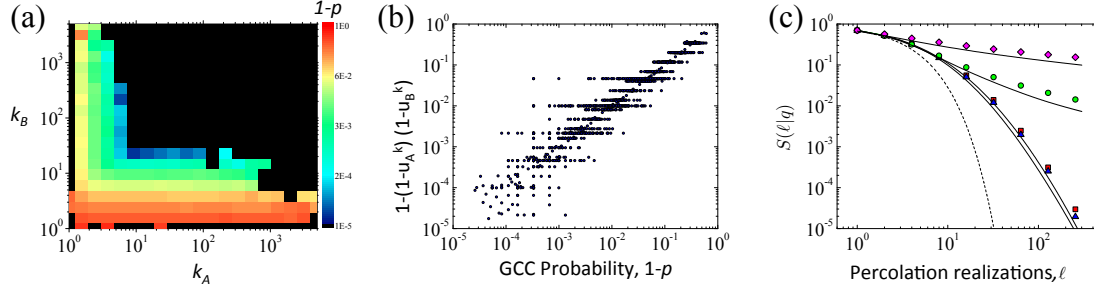
Figure 40. (a) Individual node probability to appear within the mutual GCC of the inter-connected network as a function of its degrees in the network layers. The interdependent network consists of two SF layers with random interdependence links. Both SF layers are generated using parameters reported in Fig. 38. Link percolation thresholds are chosen as $q_A = 0.49$ and $q_B = 0$, corresponding to $\langle GCC \rangle = 0.7$. (b) Individual node probabilities *not* to appear within the GCC compared to the analytical estimate of Eq. (34). (c) Stability of the mutual GCC as a function of number of independent realizations ℓ calculated for (green circles) SF-SF, (blue triangles) ER-SF and (red squares) ER-ER interdependent network models as well as (pink rhombi) the protein protein interaction network described in Appendix 4.1.1. Link percolation thresholds are, respectively, $q_A = 0.56$, $q_A = 0.63$, $q_A = 0.64$, and $q_A = 0.68$, $q_B = 0$, all corresponding to $\langle GCC \rangle = 0.7$. The dashed line corresponds to $S(\ell|q)$ of mutual GCCs consisting of random sets of nodes. Solid lines correspond to analytic solutions of Eqs (35-37)

where k_A and k_B are the degrees of the two interdependent nodes in layers A and B, respectively, $\mathbf{q} \equiv \{q_A, q_B\}$, q_A and q_B are the initial fractions of deactivated links in layers A and B respectively; while parameter u_A (u_B) is the effective probability that a given end of a randomly chosen link in layer A (B) leads to a mutual connected component of finite size. The mean-field parameters u_A and u_B can be found from the recursive approach of Buldyrev et al. (2010) and are given by the system of equations

$$1 - u_A = (1 - q_A)\left(1 - G_1^A(u_A)\right)\left(1 - G_0^B(u_B)\right), \tag{35}$$

$$1 - u_B = (1 - q_B)\left(1 - G_1^B(u_B)\right)\left(1 - G_0^A(u_A)\right), \tag{36}$$

where $G_{\{0,1\}}^{\{A,B\}}(x)$ are the generation functions for the regular and the outgoing degree distributions of layers A and B, see Fig. 40(a,b) and Appendix 4.1.4 for derivation. We note that this result is in good agreement with numerical results, which we obtain for ER-ER, ER-SF and SF-SF models as well as the network of protein-protein interactions (PPI), and depict for SF-SF network in Fig. 40(a).

126

Since interdependence links between the layers are established randomly, the joint degree distribution of the network is $p(k_A, k_B) = p_A(k_A)p_B(k_B)$, where $p_{A,B}(k)$ are degree distributions of layers A and B respectively. Then, the stability of the mutual connected component is simply the product of the contributions from the two layers:

$$S_{AB}(\ell|\mathbf{q}) = S_A(\ell|\mathbf{q})S_B(\ell|\mathbf{q}), \tag{37}$$

where S_A and S_B are the stability contributions of domains A and B respectively: $S_A \equiv \sum_k p_A(k)\left(1 - [u_A(\mathbf{q})]^k\right)^\ell$, and S_B is defined similarly. This result together with our results in Fig. 38 for single layered networks explains stable mutual GCCs in SF-SF networks and unstable GCCs in ER-ER and ER-SF networks that we observe in Fig. 40(c).

Taken together, our results indicate that networks with broader degree distributions are characterized by stable GCCs. Large degree nodes in these networks are part of the GCC with probabilities close to 1, and can be regarded as anchors keeping the GCC in place. At the same time, it is important to note that network robustness and GCC stability do not in general imply each other. We find that interdependent networks with broad degree distribution of their layers (e.g., SF-SF networks) have stable mutual GCCs. At the same time, however, SF-SF networks are known to be vulnerable to cascading failures (Buldyrev et al. 2010). Indeed, large degree nodes in interdependent networks are likely to depend on small degree nodes in other layers. As a result, failures of these small degree nodes immediately lead to failures of the large degree nodes that depend on them. In contrast, interdependent networks with narrow degree distributions, e.g., ER-ER networks, are less vulnerable due to a more uniform mixing of node degrees in their layers. The stability of connected components in interdependent networks, on the other hand, exclusively relies on the existence of anchor nodes, which according to Eq. (34) are large degree nodes dependent on other large degree nodes. While these anchor nodes are not frequent in SF-SF networks, they are non-existent in ER-ER networks, explaining why mutual GCCs are not stable in the latter.

Our findings open new avenues toward the design of efficient network reinforcement strategies. Indeed, building upon our results one can formulate the subgraph reinforcement problem as an optimization. Similar to the GCC stability, one can define the stability of any

subgraph Ω of interest:

$$S_\Omega(\ell|q) \equiv \|\bigcap_{t=1}^{\ell} \Omega_t(q)\| = \sum_{i \in \Omega} \left(1 - [u(q)]^{k_i}\right)^{\ell}, \tag{38}$$

where u is the mean-field probability that a given end of a randomly chosen link leads to a cluster of finite size and is given by Eq. (30). Then the reinforcement strategy is defined in using extra resources, e.g. additional nodes or links to maximize the $S_\Omega(\ell|q)$ objective function. It is also immediately clear that the optimal reinforcement strategies depend on the number of expected adverse events ℓ: if the system is designed to withstand a single adverse event, the reinforcement strategy should be aimed at maximizing the expected GCC size. On the other hand, if the system is designed to withstand multiple adverse events, the investments should be made to further reinforce or create the anchor nodes.

The mean-field analysis offered here has important limitations. The generating function approach works well for locally tree-like networks, while real systems contain significant amount of short loops. Likewise, it is now understood that heterogeneous distributions in particular can have considerable diversity in their topologies (Alderson and Li 2007) and that the hubs in these networks are not always centrally located. As a result, we do notice significant deviations for individual node probabilities in real networks, Fig. 38(d), that can't be explained by the mean-field approach.

Finally, within the engineering sciences there is a general understanding that the organization of many real-world networks is governed by processes that are not sufficiently characterized by stylized models of random networks and that these differences can have important implications for understanding robustness and vulnerability (Doyle et al. 2005; Florin and Linkov 2016). There is a growing body of work in large-scale optimization and game theory that leverages the specialized structure of these networks to discover specific vulnerabilities and to identify limited defensive investments that can maximally increase robustness and resilience (e.g., Alderson et al. (2015); Ganin et al. (2016)). Bridging the gap between the specific recommendations of these highly detailed models and the insights from more general models described here is an important goal for ongoing research.

4.1.1 Real networks

We test our GCC stability results on two real networks. The first real network is the complete US airport network in 2010, where nodes are US airports and an undirected link is established between two airports if there is a commercial flight between them. The network consists of $N = 1,574$ nodes and $E = 28,236$ links, and is publicly available from https://openflights.org/ and https://toreopsahl.com/datasets/.

The second is the network of human protein-protein interactions (PPI) that we derived from the BioGRID database (Stark et al. 2006). We represent the PPI dataset as a two-layer interdependent network, where layer A is formed by direct protein interactions, and links in layer B correspond to physical associations between proteins. In our analysis we focus on the mutual connected component of the PPI network that contains $N = 11,365$ nodes in each layer with the average degrees of $\langle k_A \rangle = 11.00$ and $\langle k_B \rangle = 25.38$ respectively.

4.1.2 GCC stability of single layer networks close to the criticality

Here we consider the asymptotic behavior of $S(\ell|q)$ for three types of network models: regular, ER, and SF network models.

4.1.3 Regular network models

All nodes of a regular network model have the same degree k_0, resulting in $p(k) = \delta(k, k_0)$, where $\delta(m, n)$ is the Kronecker delta. Then, GCC stability of a regular network is trivially $S(\ell|q) = \left(1 - u(q)^{k_0}\right)^{\ell}$ for all $u(q)$ values.

ER network models ER network models are characterized by the Poisson degree distribution, $p(k) = e^{-\langle k \rangle} \langle k \rangle^k / k!$, where $\langle k \rangle = p(N - 1)$ is the average degree. In order to obtain the asymptotic expression for $S(\ell|q)$ we first expand $\left(1 - u(q)^k\right)^{\ell}$ into the binomial sum. Then by reversing the summation order in Eq. (31) we obtain

$$S(\ell|q) = \sum_{m=0}^{\ell} \binom{\ell}{m} (-1)^m \exp\left(-\langle k \rangle \left(1 - u(q)^m\right)\right).$$

(39)

Close to the criticality $\epsilon(q) \equiv 1 - u(q) \ll 1$ and therefore, $1 - u(q)^m \simeq \epsilon(q)m$. Plugging this into Eq. (39) results in

$$S(\ell|q) \approx \left[1 - e^{-\epsilon(q)\langle k \rangle}\right]^{\ell} \approx \left[1 - u(q)^{\langle k \rangle}\right]^{\ell}. \tag{40}$$

SF network models SF network models are characterized by power-law degree distributions $p(k) \approx (\gamma - 1)k_0^{\gamma-1}k^{-\gamma}$, where k_0 is the smallest degree and $\gamma > 2$.

To deduce the asymptotic behavior of $S(\ell|q)$ we first approximate $u(q)^k$ in Eq. (31) as $e^{-k\epsilon(q)}$, where $\epsilon(q) \equiv 1 - u(q) \ll 1$. Then, by approximating the summation in Eq. (31) with an integral we obtain:

$$S(\ell|q) \approx (\gamma - 1)\left[\epsilon(q)k_0\right]^{\gamma-1} I_{\ell}(\gamma), \tag{41}$$

$$I_{\ell}(\gamma) \equiv \int_{\epsilon(q)k_0}^{\infty} x^{-\gamma}\left[1 - e^{-x}\right]^{\ell} dx. \tag{42}$$

Further, to deduce the asymptotic behavior of (42) we note that the function $[1 - e^{-x}]^{\ell}$ is monotonically increasing from 0 to 1 on the interval $0 < x < \infty$. It is useful to think about $[1 - e^{-x}]^{\ell}$ as a smoothed version of the step function $\theta(x - \ln \ell)$. To appreciate the shift we center the function $[1 - e^{-x}]^{\ell}$ at the position x_0 where it equals e^{-1}. From $[1 - e^{-x_0}]^{\ell} = e^{-1}$ we obtain $e^{-x_0} = 1 - e^{-1/\ell}$ leading, when $\ell \gg 1$, to

$$x_0 = \ln \ell \tag{43}$$

and thereby explaining the $\ln \ell$ shift. To appreciate the width of the quasi-step function $[1 - e^{-x}]^{\ell}$ we define the left boundary through $[1 - e^{-x_-}]^{\ell} = e^{-M}$ and the right boundary through $[1 - e^{-x_+}]^{\ell} = 1 - e^{-M}$. These definitions are reasonable if we choose $M \gg 1$. From these definitions

$$x_- = \ln \ell - \ln M, \qquad x_+ = \ln \ell + M \tag{44}$$

Thus the quasi-step function $[1 - e^{-x}]^{\ell}$ has the following properties:

1. It is essentially vanishes when $0 < x < x_-$.
2. It monotonically increases from 0 to 1 in the interval $x_- < x < x_+$.

130

3. It is essentially equal to 1 in the interval $x > x_+$.

We next establish the upper and lower bounds for the integral in (42). The upper bound is obtained by replacing $[1 - e^{-x}]^\ell$ by unity when $x > x_-$ and e^{-M} when $k_0\epsilon(q) < x < x_-$; the lower bound is obtained by replacing $[1 - e^{-x}]^\ell$ by $1 - e^{-M}$ when $x > x_+$ and zero when $0 < x < x_+$. Since $\int_{x_*}^{\infty} \frac{dx}{x^\gamma} = \frac{1}{\gamma-1} \frac{1}{x_*^{\gamma-1}}$, the bounds are

$$\frac{1 - e^{-M}}{x_+^{\gamma-1}} \le (\gamma - 1)I_\ell(\gamma) \le \left(\frac{1}{x_-^{\gamma-1}} + e^{-M} \frac{1}{[\epsilon(q)k_0]^{\gamma-1}} \right) \tag{45}$$

Since both $x_\pm \to \ln \ell$ in the $\{\ell, M\} \to \infty$ limit, such that $1 \ll M \ll \ell$, $I_\ell(\gamma) \sim \ln \ell^{1-\gamma}$ and

$$P(\ell|q) \sim \left[\frac{k_0\epsilon(q)}{\ln \ell} \right]^{1-\gamma}. \tag{46}$$

4.1.4   GCC stability of interdependent networks

Link percolation results for interdependent networks can be recovered from Buldyrev et al. (2010) by setting the fraction of active nodes to 1 and replacing degree generating functions with those for link percolated networks. Indeed, let us introduce the generating functions for degree distributions of layers A and B as

$$G_0^{\{A,B\}}(x) \equiv \sum_k p_{\{A,B\}}(k)x^k, \tag{47}$$

$$G_1^{\{A,B\}}(x) \equiv \sum_k \frac{(k+1)}{\langle k_{\{A,B\}} \rangle} p_{\{A,B\}}(k+1)x^k, \tag{48}$$

where $p_A(k)$ and $p_B(k)$ are the degree distributions of domains A and B respectively, while $\langle k_A \rangle$ and $\langle k_B \rangle$ are the corresponding average degree values.

Then the expected size of the mutual connected component μ_∞ is given by

$$\mu_\infty = xy, \tag{49}$$

where the mean-field parameters x and y can be regarded as the effective fractions of nodes belonging to the GCC of layers A and B respectively, and are given by the system of four

equations:

$$x = 1 - G_0^A (1 - y(1 - f_A)), \tag{50}$$

$$y = 1 - G_0^B (1 - x(1 - f_B)), \tag{51}$$

$$f_A = G_1^A (1 - y(1 - f_A)), \tag{52}$$

$$f_B = G_1^B (1 - x(1 - f_B)). \tag{53}$$

Here parameters f_A and f_B in analogy with the single layer case can be regarded as the effective probabilities that a given end of a randomly chosen link in layer A or B, leads to a connected component of a finite size.

By making use of Eqs. (50) and (51) one can rewrite the expression for the mutual connected component as

$$\mu_\infty = \left(1 - G_0^A (1 - y(1 - f_A))\right) \left(1 - G_0^B (1 - x(1 - f_B))\right), \tag{54}$$

from which one can immediately extract the individual node probability to appear within the mutual GCC:

$$p = \left(1 - u_A^{k_A}\right) \left(1 - u_B^{k_B}\right), \tag{55}$$

where k_A and k_B are the degrees of the two interdependent nodes in layers A and B, and the mean-field parameters u_A and u_B are given by

$$u_A = 1 - y(1 - f_A), \tag{56}$$

$$u_B = 1 - x(1 - f_B). \tag{57}$$

The same equations hold in the link percolation case, except the original generating functions $G_{\{0,1\}}^{\{A,B\}}(x)$ need to be replaced with those for percolated layers:

$$\tilde{G}_{\{0,1\}}^A(x) = G_{\{0,1\}}^A (1 - (1 - q_A)(1 - x)), \tag{58}$$

$$\tilde{G}_{\{0,1\}}^B(x) = G_{\{0,1\}}^B (1 - (1 - q_B)(1 - x)), \tag{59}$$

where q_A and q_B are the fractions of initially deactivated links in layers A and B respectively.

132

After a series of straightforward substitutions and simplifications the final result reads

$$p = \left(1 - [u_A(\mathbf{q})]^{k_A}\right)\left(1 - [u_B(\mathbf{q})]^{k_B}\right), \tag{60}$$

where

$$1 - u_A = (1 - q_A)\left(1 - G_1^A(u_A)\right)\left(1 - G_0^B(u_B)\right),$$
$$1 - u_B = (1 - q_B)\left(1 - G_1^B(u_B)\right)\left(1 - G_0^A(u_A)\right),$$

reproducing Eqs. (35) and (36) in the main text.

4.2 Reconstruction of Communication Paths in Incomplete Networks

Imagine a tourist standing at the intersection of West 14th street and 8th avenue in Manhattan, NY and trying to find her way to the Empire State Building. – one of the the main attractions that Manhattan has to offer, which is located at the intersection of West 34th st and 5th avenue. The tourist easily finds her way to the Empire State Building without using a map or a GPS. Now, imagine the same tourist navigating the Peace maze in the Northern Ireland. Without a map one can easily take hours to find the way out. From the graph theory standpoint, both the road system of Manhattan and the Peace maze are graphs or networks, where intersections are represented as nodes and road segments are links connecting them, and both navigation tasks can be be solved by finding a path between a pair of nodes. What makes Manhattan road network different from a maze is its grid-like structure and consistent street naming, allowing for efficient navigation. Being aware of these properties, one can find a pedestrian path to the intersection of interest by taking streets and avenues in the increasing or decreasing order. Further, if marked on the map, all reasonable pedestrian paths are aligned closed to the geodesic line connecting the two intersections.

The main idea behind our approach is that shortest paths in geometric networks are closely aligned along the geodesic lines connecting origin-destination node pairs. While many real networks, e.g., road networks and wireless communication networks, are embedded into real physical space, of our interest in this work are networks with effective or latent geometries. Nodes in these networks can be mapped to points in latent spaces such that connections between them preferentially take place between nodes separated by shortest

distances (Boguna et al. 2020). Of special interest here are networks whose latent geometries are not Euclidean but hyperbolic. A collection of recent works indicate that latent hyperbolic geometry is consistent with topological properties of real networks, such as hierarchical organization, scale-free property, strong clustering coefficient and self-similarity (Boguna et al. 2020). Notable examples of real networks with effective hyperbolic geometries are the network of protein interactions (Alanis-Lobato et al. 2018), the Internet at the Autonomous System (AS) level (Boguñá et al. 2010), and the World Trade Web (Serrano and Boguñá 2003).

In our work we focus on hyperbolic maps of the AS Internet, Fig. 41**a**, and the network of human protein interactions. In the former network nodes are Autonomous Systems (ASes) and connections are contractual agreements governing data flows between ASes, while in the latter the nodes are proteins and connections correspond to physical interactions between them. To demonstrate the idea of geometric path reconstruction we marked ASes lying on the shortest path between two AS pairs and corresponding geodesics in the 2-dimensional hyperbolic disk. As seen from Fig. 41**a**, nodes comprising the shortest paths lie in the vicinity of corresponding geodesics, shown with solid lines connecting origin-destination node pairs.
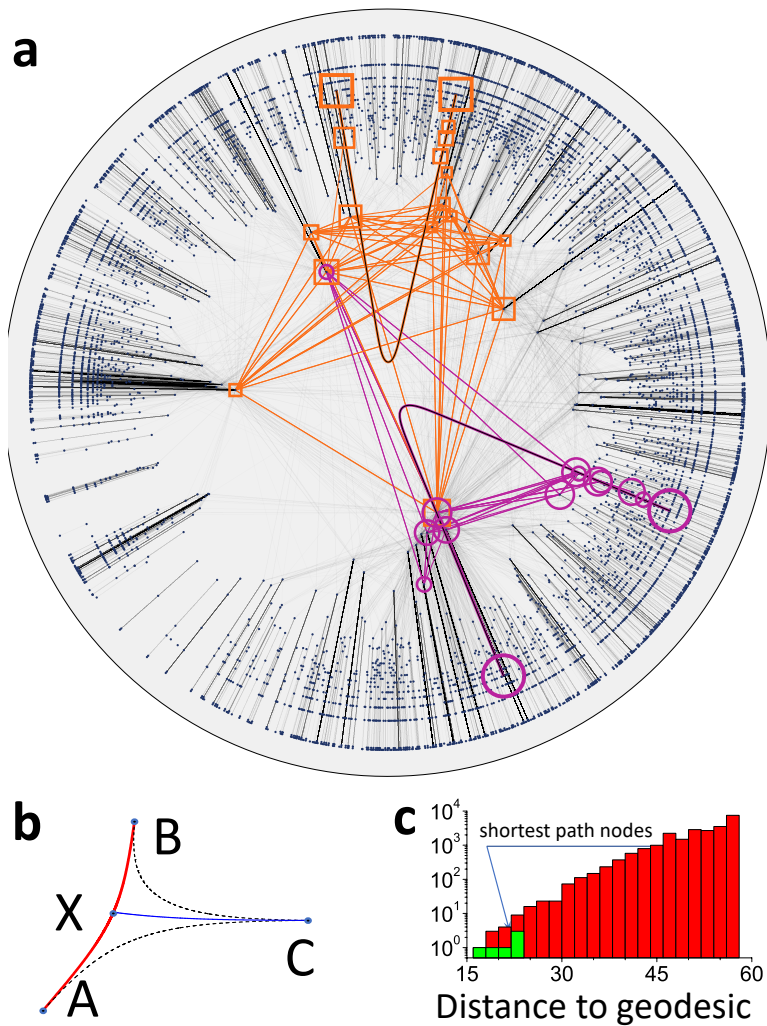
Figure 41. Latent geometry uncovers communication paths in the Internet at the Autonomous System level. **a**, Hyperbolic map of the Internet at the Autonomous System (AS) level. The latent space is the 2-dimensional hyperbolic disk and each point corresponds to an Autonomous System. Yellow squares and pink circles highlight ASes corresponding to communication paths between AS5392-AS8875 and AS1224-AS11650 pairs. **b**, Distance from point C to geodesic $\gamma(A, B)$ drawn between points A and B, as defined by Eq. (64). **c**, The distribution of distances to the $\gamma(AS5392, AS8875)$ geodesic from (green) shortest path nodes and (red) all Internet nodes. Our main observation, motivating this work, is that ASes comprising the shortest paths are localized in the geometric vicinity of hyperbolic geodesics.

To quantify the extent of the observed alignment we evaluated distances from all network nodes to the AS5392-AS8875 geodesic, finding that all 6 ASes comprising shortest paths between the AS5392-AS8875 pair are among 12 ASes characterized by the shortest distance

to the geodesic, see Fig. 41**c** and Methods. Our observation implies that the inverse is also true: nodes with shortest distances to the $\gamma(A, B)$ geodesic are likely to be the shortest path nodes connecting the $A - B$ node pair.

To test if this is indeed the case and whether the distance to the geodesic is robust with respect to the missing network data we removed 50% of the Internet links uniformly at random and then re-mapped the resulting incomplete network to the hyperbolic space. We then relied on the resulting inferred coordinates to identify ASes in the vicinity of the updated AS5392-AS8875 geodesic. To systematically test the accuracy of the distance to geodesic as a predictor of node likelihood to comprise the shortest path we used the area under the precision-recall characteristic (AUPR). We compared the accuracy of the distance to geodesic to that of degree-based ranking, brute-force network-based path ranking and reconstructed network-based path ranking, see Methods. Our results, computed for both the original and the incomplete AS Internet network indicate that the accuracy of the distance to the geodesic metric is comparable to that of the degree-based ranking and is substantially higher than network-based distances, Fig. 42**a**. This is the case since missing link hinder the identifiability of true shortest paths: even a single missing link may lead to a completely different set of shortest path nodes. Latent-geometric mapping of a network, on the other hand, provides an effective *mean field* image of the network, which is not sensitive to uniformly missing links.

We also observe that the removal of a small number of ASes closest to the AS5392-AS8875 geodesic fully disconnects the AS5392-AS8875 pair using both coordinates of the original and truncated network, Fig. 42**b** with minimal collateral damage to the rest of the network, Fig. 42**c**. This observation is in contrast with other network-based metrics that fail to disconnect the AS pair of interest upon the removal of 100 nodes. This result implies that not only shortest path nodes but also other nodes that lie on nearly shortest paths are localized in the vicinity of latent-geometric geodesics.

Our findings are not specific to the chosen AS5392-AS8875 pair and the AS Internet network. To establish whether latent-geometric geodesics can help identify shortest path nodes we conducted a series of systematic experiments on the AS Internet and the human interactome networks, finding that distance to geodesic measure given by Eq. (64) allows one to reliably identify shortest path nodes, Fig. 42**d**. Further, we found that targeted removal

of nodes with smallest distances to geodesic s allows one to disconnect corresponding node pairs by keeping the rest of the network minimally disturbed, see Fig. 42**e, f**.

Further, to test the limits of the distance to geodesic metric in the identification of shortest path nodes we repeated our experiments on incomplete random hyperbolic graph (RHG) models (Aldecoa et al. 2015). These experiments support our results on real networks and establish that the accuracy of the distance to the geodesic measure is highest at small RHG temperature values, when RHG links are preferentially established at shortest latent-geometric distances, and is nearly independent of the scale-free degree distribution exponent γ. As T increases and the role of geometry in link formation decreases the accuracy of the distance to geodesic measure decreases and becomes comparable to brute-force network-based methods.

Having established the accuracy of the distance to geodesic measure in the identification of shortest path nodes, we next discuss potential applications of shortest path identification to Internet routing and the analysis of biological pathways.

Anomaly detection for Internet routing

The Internet is a distributed system: there exists no central authority governing connections between ASes and routing of information between them. The Border Gateway Protocol (BGP) facilitates the interdomain routing process by discovering and distributing this reachability information to all ASes (Rekhter et al. 1994). Since the topology of the Internet is highly dynamic, ASes must constantly exchange and update this reachability information in small chunks, known as routing control packets or BGP updates. The BGP routing is based on trust, where an AS often accepts routes advertised by its neighbors without strict integrity tests (Rekhter et al. 1994). This exposes the entire system to a possibility of cyberattacks and malicious interventions. One class of cyberattacks is the BGP prefix hijacking. As a result of a BGP prefix hijack, an AS either claims ownership of prefixes that are owned by other ASes or announces that it can provide transit to a prefix or a set of prefixes. This attack can compromise the affected data flows by either exposing them to malicious actors or simply misroute them (Sermpezis et al. 2018).

To suppress prefix hijack attacks ASes need to be able to assess advertised BGP paths for their optimality and reject problematic ones. Since shortest and nearly shortest paths tend

to align along hyperbolic geodesics, we propose the hyperbolic stretch as the measure of the conformity of a BGP path to the corresponding geodesic, Fig. 43**a, b**. We define the hyperbolic stretch for a path connecting $A - B$ node pairs as the average normalized distance to the $\gamma(A, B)$:

$$D_{\Omega(A,B)} = \max_{C \in \Omega(A,B)} \frac{d(C, \gamma(A, B))}{d(A, B)}, \tag{61}$$

where $d(A, B)$ is latent geometric distance between ASes A and B and $d(C_i, \gamma(A, B))$ is given by Eq. (64).

We analyzed three notable prefix hijacks events involving China Telecom, Malaysia Telekom (MT), and Rostelecom (PJSC). For each event we analyzed BGP paths involving the hijacked ASes advertised both before and during the hijack event, finding that hyperbolic stretches of BGP paths advertised during the hijack events are significantly larger than those before the event, Fig. 43**c**.

Our preliminary results suggest that hyperbolic stretch can be used by ASes to rule out problematic paths. The hyperbolic map of the AS Internet can be regarded as certain mean-field map that is not sensitive to short-time perturbations of Internet topology. AS coordinates need to be recomputed once every several years, while coordinates of new ASes may be computed locally (Boguñá et al. 2010). Thus, each ASes may be endowed with the list of coordinates of all ASes, and these coordinates can be used to compute hyperbolic stretches of advertised paths. Alternatively, node coordinates could be carried alongside path information by BGP updates. Developing a latent-geometric classifier of BGP path and evaluating its efficiency is the subject for our future work.

Taken together, we established that latent-geometric geodesics serve as fairways for shortest and nearly shortest paths nodes in geometric networks Nodes in the vicinity of these geodesics are most likely to lie on shortest paths or may become shortest path nodes if network topology is perturbed. Our finding can be both be either a curse or a blessing depending on the circumstances. One could exploit the geometric localization of shortest paths to disrupt or eavesdrop communication paths of interest. On the other hand, the knowledge of geodesic fairways may help identify alternative optimal paths and rule out inefficient or fraudulent paths in the case of the incomplete networks. Two promising applications of our approach are the anomaly detection for Internet routing and the analysis of biological pathways.

At the same time, we emphasize that there is no one-size-fits all solution to the shortest path identification problem. In order to identify shortest path nodes in a partially known network one needs to know both the mechanisms of network formation and the character of missing data. Distance to geodesic, in this respect, assumes that link formation in the network is captured by its latent or effective geometry, and unknown links are missing uniformly at random. The first condition is a must: one cannot expect to identify shortest paths in non-geometric networks using geometric methods. The second assumption could probably be relaxed. The cornerstone of any latent-geometric inference task is the accurate latent-geometric map of the network of interest. In our work we used the HyperLink (HL) embedder (Kitsak et al. 2019), which has been shown to be accurate on uniformly incomplete networks (Kitsak et al. 2019). While the HL embedder is not expected to be accurate on non-uniformly incomplete networks, it should be straightforward to generalize the embedder to the special non-uniform case when the probability of a missing link is also a function of a latent distance between the nodes.

4.2.1 Methods

Distance to geodesic. Distance between two points $\{r_i, \theta_i\}$ and $\{r_j, \theta_j\}$ in the 2-dimensional hyperbolic disk $\mathbb{H}^2$ is given by the hyperbolic law of cosines and for sufficiently large r_i and r_j values is closely approximated by $\mathbf{x}_{ij} = r_i + r_j + \frac{2}{\zeta} \ln \left(\sin(\Delta \theta_{ij}/2) \right)$. We define distance from point C to geodesic $\gamma(A, B)$ as the shortest distance from C to any point on $X \in \gamma(A, B)$:

$$d(C, \gamma(A, B)) = \min \ d(C, X), \tag{62}$$

$$\text{s.t. } X \in \gamma(A, B) \tag{63}$$

The distance to the hyperbolic geodesic $d(C, \gamma(A, B))$ is closely approximated as

$$d(C, \gamma(A, B)) = \frac{1}{2} \left[d(A, C) + d(B, C) - d(A, B) \right], \tag{64}$$

provided $d(A, C)$, $d(B, C)$ and $d(A, B)$ are sufficiently large.

Network mapping or embedding into a latent space $\mathcal{M}$ is a procedure of determining the coordinates of nodes constituting the network in this space. In this work we map AS Internet and the similarity-based protein interaction network to the 2-dimensional hyperbolic

disk using the HL Embedder algorithm (Kitsak et al. 2019). Similar to other hyperbolic embedders, HL embedder maps network nodes to points $\{r_i, \theta_i\}$, $i = 1, ..., N$, in a hyperbolic disk $\mathbb{H}^2$ by maximizing the posterior probability $\mathcal{L}\left(\{r_i, \theta_i\}|a_{ij}\right)$ that the network with the adjacency matrix a_{ij} has given node coordinates and is generated as the RHG. By the Bayes' rule, $\mathcal{L}\left(\{r_i, \theta_i\}|a_{ij}\right) \propto \frac{\mathcal{L}\left(a_{ij}|\{r_i, \theta_i\}|\right)}{\text{Prob}}\left(\{r_i, \theta_i\}\right)$, where $\mathcal{L}\left(a_{ij}|\{r_i, \theta_i\}|\right)$ is the likelihood that the network a_{ij} is generated as the RHG, given node coordinates $\{r_i, \theta_i\}$, and the Prob $\left(\{r_i, \theta_i\}\right)$ is the prior probability of node coordinates generated by the RHG. Since links $\{ij\}$ in the RHG are established independently with probabilities depending on hyperbolic distances $\{\mathbf{x}_{ij}\}$ between the nodes, $\mathcal{L}\left(a_{ij}|\{\mathbf{x}_i\}\right) = \prod_{i<j}\left[p\left(x_{ij}\right)\right]^{a_{ij}}\left[1 - p\left(x_{ij}\right)\right]^{1-a_{ij}}$. The HL embedder is freely available at the DK-lab (Kitsak and Voitalov 2020).

Alternative path relevance metrics. In addition to distance to geodesic we use three network-based metric to quantify node relevance to the path of interest. The first metric is node degree, the larger is the node degree the higher is the node relevance to the geodesic. This is is the case since shortest paths often pass large degree nodes. The second metric is the brute-force networked-based path ranking which is similar to the distance to geodesic approximation in Eq. (64) except distances are network-based. $d_{\text{bf}}(C|A, B) = \ell_{A,C} + \ell_{A,B}$. The third metric is reconstructed network-based distance, which is obtained by first complementing the network with missing links added randomly, and then is given by $d_{\text{bf}}(C|A, B) = \ell_{A,C} + \ell_{A,B}$ on the resulting reconstructed network.

In Figure 42(a-c), We aim to identify nodes that lie on shortest paths between the AS5392-AS8875 pair in both the original and the incomplete Internet where 50% of links are removed uniformly at random. In Figure 42(e-f) Systematic shortest path identification tests for the AS Internet. We create three cohorts of , consisting of 100 pairs each, with angular separations closest to $\Delta\Theta_{ij} = \frac{\pi}{8}$, $\Delta\Theta_{ij} = \frac{\pi}{4}$, and $\Delta\Theta_{ij} = \frac{\pi}{2}$.
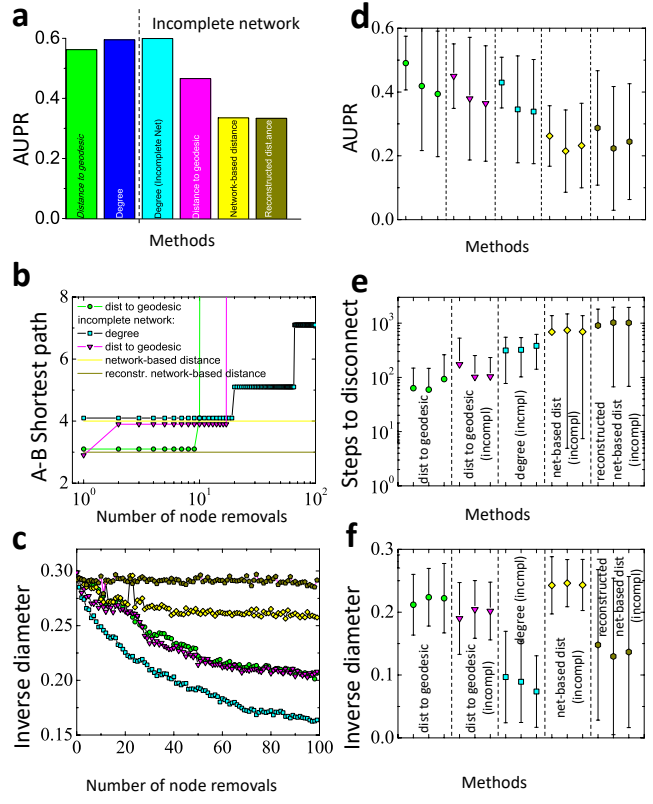
Figure 42. The accuracy of shortest path reconstructions in the AS Internet. **a**, AUPR scores for the identification of shortest path nodes using (green) distance to geodesic in the map of the original network, (blue) node degrees in the original network, (teal) node degrees in the incomplete network, (maroon) distance to geodesic in the map of the incomplete network, (yellow) network-based distance in the incomplete network, (dark yellow) reconstructed network-based distance in the incomplete network. **b**, The length of the shortest path connecting the AS5392-AS8875 pair as a function of the number of removed nodes. Nodes are removed based on their rank, using one of the methods. The methods are the same as in panel **a** save for the original network degree ranking, which is almost identical as the degree ranking in the incomplete network. It takes 12 and 19 node removals using geometric, respectively, distance to geodesic scores obtained for the full and incomplete networks. Other methods fail to disconnect the AS5392-AS8875 pair upon the removal of 100 nodes. **c** To quantify the collateral damage associated with the removal of nodes we estimate the average inverse network-based distance in the network as a function of number of removed nodes. The higher the average inverse distance, the less perturbed is the network. The color coding of the methods is the same as in panel **b**. We note that distance to geodesic methods are less invasive than the degree-based method. While network-based and reconstructed network-based methods are the least invasive, they are of little use as they cannot disconnect the AS5392-AS8875 upon the removal of 100 nodes. **d**, For each of these cohorts we evaluate the average AUPR score for path reconstruction, panel ; the average number of node removals to disconnect the AS pairs, panel **e**; the average inverse network diameter after nodes are disconnected. panel **f**.

141
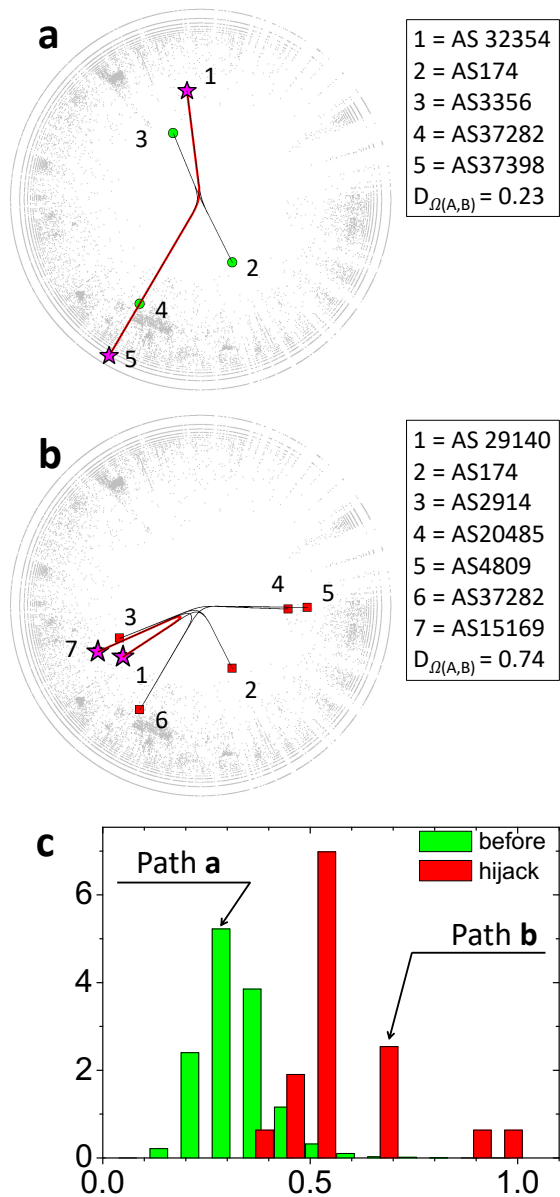
Figure 43. Anomaly detection in interdomain Internet routing. **a**, Example of a BGP path traversing the China Telecom (AS4809) announced before its prefix hijack. Green circles are ASes constituting the path, solid red line is the hyperbolic geodesic connecting the origin-destination pair. Latent-geometric stretch of the BGP path, Eq. (64), $D_{\Omega(A,B)=0.23}$. **b** Example of a BGP path traversing the China Telecom (AS4809) announced during its prefix hijack event. Red squares are ASes constituting the path, solid red line is the hyperbolic geodesic connecting the origin-destination pair. Latent-geometric stretch of the BGP path is $D_{\Omega(A,B)=0.74}$. **c** Distributions of latent-geometric stretches for BGP paths announced (green) before and (during) the China Telecom prefix hijack. Note that BGP paths announced before the hijack are characterize by significantly smaller stretch values than those announced during the hijack event.

4.3 Geometric Representation of Complementarity-Driven Networks

Network embeddings or mappings of networks to latent geometric spaces are standard tools in the arsenal of data analysis, and are routinely used in machine learning, visualization, network science, and graph theory. In general, a procedure of network embedding is a mapping network nodes to points in a suitable latent metric space, such that latent distances between connected node pairs are smaller than those between disconnected node pairs.

Latent-geometric distances are often interpreted as generalized measures of node similarities (McFarland and Brown 1973): the closer the two nodes in the latent space the more similar they are and the more likely they are to be connected in the network of interest. It is the similarity interpretation of latent distances that lies at the origin of many applications of network embeddings, including link prediction (Cannistraci et al. 2013b; Yang et al. 2015a; Xiao et al. 2015; Tang et al. 2015; Grover and Leskovec 2016; Zhu et al. 2016; Nickel and Kiela 2018; Kazemi and Poole 2018; Brew 1998; Kitsak et al. 2019; García-Pérez et al. 2019), soft community detection and clustering (Newman and Peixoto 2015; Zuev et al. 2015; Yang et al. 2015b; Sewell and Chen 2017), network navigation (Boguñá et al. 2010; Gulyás et al. 2015; Voitalov et al. 2017; García-Pérez et al. 2018), and search (Kleinberg 2006; Ratnasamy et al. 2001; Risson and Moors 2006).

Of our interest in this letter are networks that are based not on similarity but *complementarity* principles: nodes are connected in these networks if their properties are complementary. Examples of complementarity-driven networks include interdisciplinary collaboration networks, molecular interaction networks, and food webs. Indeed, individuals with complementary expertise are more likely to solve an interdisciplinary problem of interest, interactions often take place between molecular with complementary chemical properties and/or binding interfaces. While cannibalism is certainly present in food webs, in general species preferentially eat other complimentary species.

143
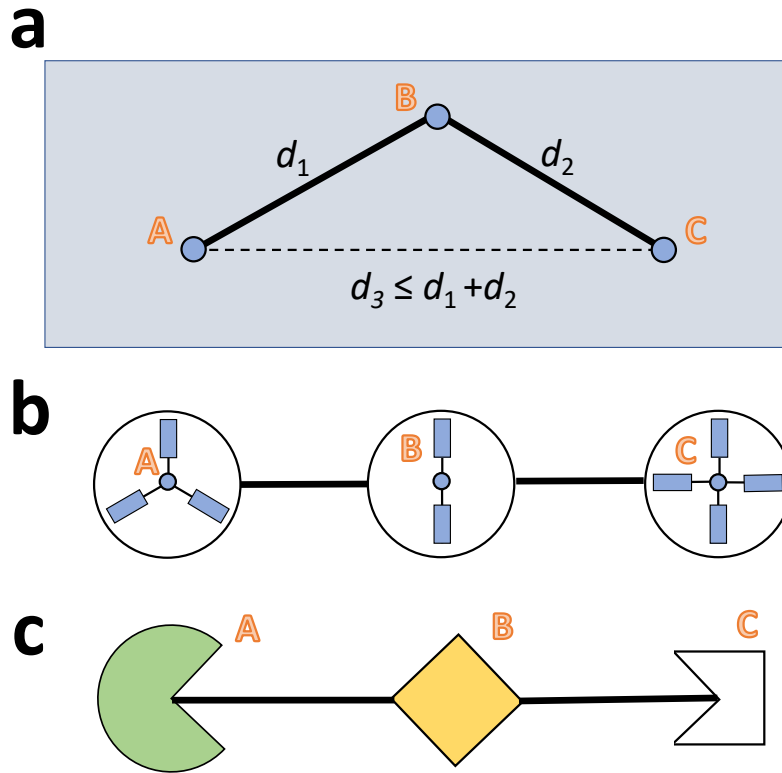
Figure 44. (**a**) Embedding of a network into a metric spaces imposes constraints on distances due to the triangle inequality. (**b**) These constraints are consistent with the similarity interpretation of latent distances since the former is transitive: if A is similar to B and B is similar to C, then A is similar C. (**c**) Triangle inequality is not consistent with complementarity: if A is complementary to B and B is complementary to C, A is not guaranteed to be complementary to C.

The major challenge behind the geometric representation of complementarity-driven networks is that unlike similarity, complementarity is not transitive. Indeed, imagine a toy complementarity-driven network consisting of three nodes A, B, and C, such that nodes A and C are both connected to B, Fig. 44**a**. Being both complementary to B, nodes A and C are not necessarily complementary to each other and, as a result, are not likely to be connected. If embedded to a metric space, however, nodes are likely to be mapped to points such that latent distances $d(A, B)$ and $d(B, C)$ are small. The metric property of the latent space then dictates that $d(A, C) \leq d(A, B) + d(B, C)$, suggesting that nodes A and C are likely to be connected, while not being complementary to each other. It is the contradiction

between the metric property of the latent space and non-transitivity of complementarity that may explain the relatively poor performance of network embedders when applied to complementarity-driven networks, Fig. 46.

This contradiction is closely related to the failure of the triadic closure principle observed in protein interaction networks (Kovács et al. 2019): proteins with a large number of common interaction partners are not likely to interact with each other. Instead, the authors of this work established that proteins are likely to interact with other proteins, that are similar to their existing interaction partners and used this principle to develop a link prediction method, based on the statistics of paths of length $\ell = 3$ and called $L3$, to predict missing protein interaction (Kovács et al. 2019). The ideas behind the $L3$ method were further developed and generalized for unipartite (Muscoloni et al. 2018) and bipartite (Daminelli et al. 2015) networks.

To enable geometric representations of complementarity-driven networks we propose a cross-geometric framework for statistical inference on complementarity-driven networks. We postulate that each node $i = 1, ..., N$ in the network is characterized by two points $\mathbf{x}^1$ and $\mathbf{x}^2$ in a latent metric space $\mathcal{M}$ and the connection probability between any two nodes takes the form of

$$p_{ij} = p\left[d\left(\mathbf{x}_i^1, \mathbf{x}_j^2\right)\right] + p\left[d\left(\mathbf{x}_i^2, \mathbf{x}_j^1\right)\right] - p\left[d\left(\mathbf{x}_i^1, \mathbf{x}_j^2\right)\right] p\left[d\left(\mathbf{x}_i^2, \mathbf{x}_j^1\right)\right], \qquad (65)$$

where $p(x)$ is any decreasing integrable function with the range of $[0, 1]$, and $d(\mathbf{x}, \mathbf{y})$ is the distance between points $\mathbf{x}$ and $\mathbf{y}$ in $\mathcal{M}$.

It is straightforward to verify that the proposed framework does not violate the transitivity constraint. Indeed, within the cross-geometric framework any node A is represented by two points $\mathbf{x}_A^1$ and $\mathbf{x}_A^2$, and connection between any two nodes A and B is higher if either cross-distance $d\left(\mathbf{x}_A^1, \mathbf{x}_B^2\right)$ or $d\left(\mathbf{x}_A^2, \mathbf{x}_B^1\right)$ is small. Within the cross-geometric framework the formation of the A-B-C toy wedge network, Fig. 45a, is possible in two cases: (i) points $\mathbf{x}_A^1$, $\mathbf{x}_B^2$, and $\mathbf{x}_C^1$ are close to each other, Fig. 45b, or (ii) point $\mathbf{x}_A^1$ is close to point $\mathbf{x}_B^2$, and $\mathbf{x}_B^1$ is close to $\mathbf{x}_C^2$, Fig. 45c. Neither case creates constraints on distances $d\left(\mathbf{x}_A^1, \mathbf{x}_C^2\right)$ and $d\left(\mathbf{x}_C^2, \mathbf{x}_C^1\right)$, relevant for the formation of the A-C link, 45b, c.
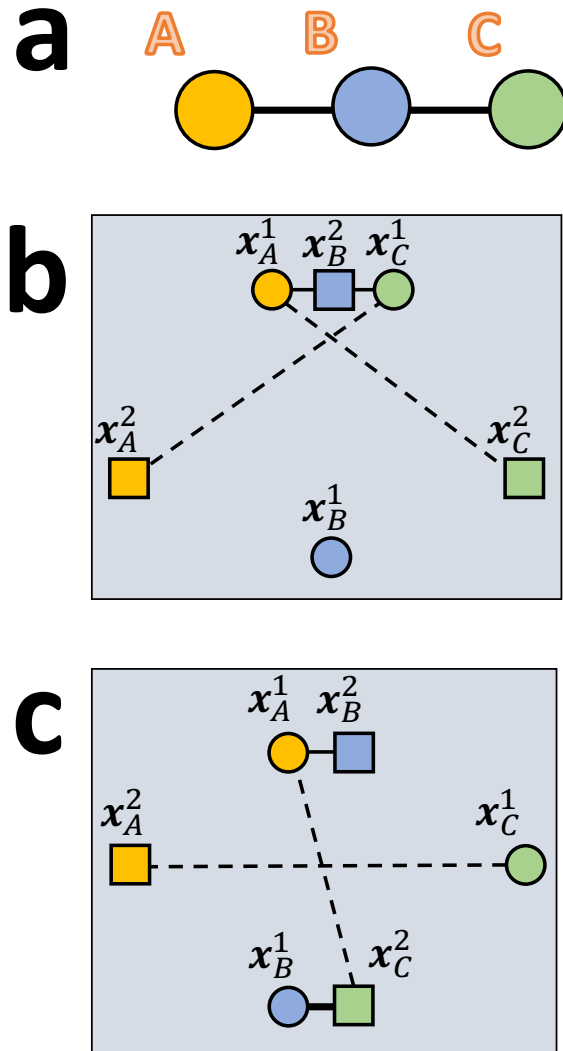
Figure 45. Cross-geometric framework does not impose triangle closure constraints. Each node in the (**a**) toy network is represented by two points in manifold $\mathcal{M}$. Connections between node pairs *A-B* and *B-C* are possible due to (**b**) latent-geometric proximity of points $\mathbf{x}_A^1$, $\mathbf{x}_B^2$, and $\mathbf{x}_C^1$ or (**c**) latent geometric proximity of point pairs $\mathbf{x}_A^1$ and $\mathbf{x}_B^2$, and $\mathbf{x}_B^1$ and $\mathbf{x}_C^2$. Neither configuration imposes constraints on distance between point pairs $\mathbf{x}_A^1$ and $\mathbf{x}_c^2$ or $\mathbf{x}_A^2$ and $\mathbf{x}_C^1$, which are relevant to the formation of the *A-C* tie. The other two configurations are identical to those in (**b**) and (**c**) upon relabeling node points $\mathbf{x}_Y^1 \rightarrow \mathbf{x}_Y^2$, where $Y = \{A, B, C\}$.

The emerging cross-geometricity of the proposed complementarity framework is easy to justify: two points corresponding to a given node characterize node's orientation with respect to relevant node properties. Examples of such properties could be scientist's expertise in two complementary disciplines, e.g., biology and statistics, in the case of a scientific collaboration network, or biochemical properties of a molecule. For some complementarity systems it might be necessary to include more than two points per node.

Cross-distances $d\left(\mathbf{x}_i^1, \mathbf{x}_j^2\right)$ and $d\left(\mathbf{x}_i^2, \mathbf{x}_j^1\right)$ in the connection probability of Eq. (65) quantify the complementarity between the two relevant node properties: the closer the two properties are, the higher is the extent of their complementarity. At first glance, it might seem unreasonable to place both node properties in the same space $\mathcal{M}$. A more intuitive setting could be a multi-space model, where node properties belong to distinct latent spaces $\mathcal{M}_1$ and $\mathcal{M}_2$, and connections are possible due to a certain mapping function $f : \mathcal{M}_1, \mathcal{M}_2 \to p \in [0, 1]$. Both single and multi-space formulations are compatible, and the former is the special case of the latter if node properties are correlated.

In this work we do not aim to learn latent spaces underlying real complementarity-driven systems, nor do we aim to learn the functional form of the probability function $p(d)$ in Eq. (65) – this will be the subject of the future research. Instead, we consider one example of the cross-geometric model that is using the hyperbolic disk as a latent space, $\mathcal{M} = \mathbb{H}^2$. We refer to this model as the Complementary-based Random Hyperbolic graph (CRHG). We demonstrate below that CHRG results in complementarity model networks with structural properties of real networks: heterogeneous distributions of node degrees and common neighbors, as well as vanishing density of short 3-loops, as captured by clustering coefficient. Further, we use the CHRG as a null-model to learn cross-geometric representations of real networks.

CRHG is based on the random hyperbolic graph (Krioukov et al. 2010) and uses the 2-dimensional hyperbolic disk $\mathbb{H}^2$ as a latent space. Points in $\mathbb{H}^2$ are described by polar coordinates $r \in [0, R]$ and $\theta \in [0, 2\pi]$, and distances between any two points $\{r^1, \theta^1\}$ and $\{r^2, \theta^2\}$ are calculated using the hyperbolic law of cosines:

$$\cosh d = \cosh r^1 \cosh r^2 - \sinh r^1 \sinh r^2 \cos \Delta\theta^{12}, \tag{66}$$

where $\Delta\theta^{12} = \pi - |\pi - |\theta^1 - \theta^2||$. Similar to the RHG, we assume that angular coordinates
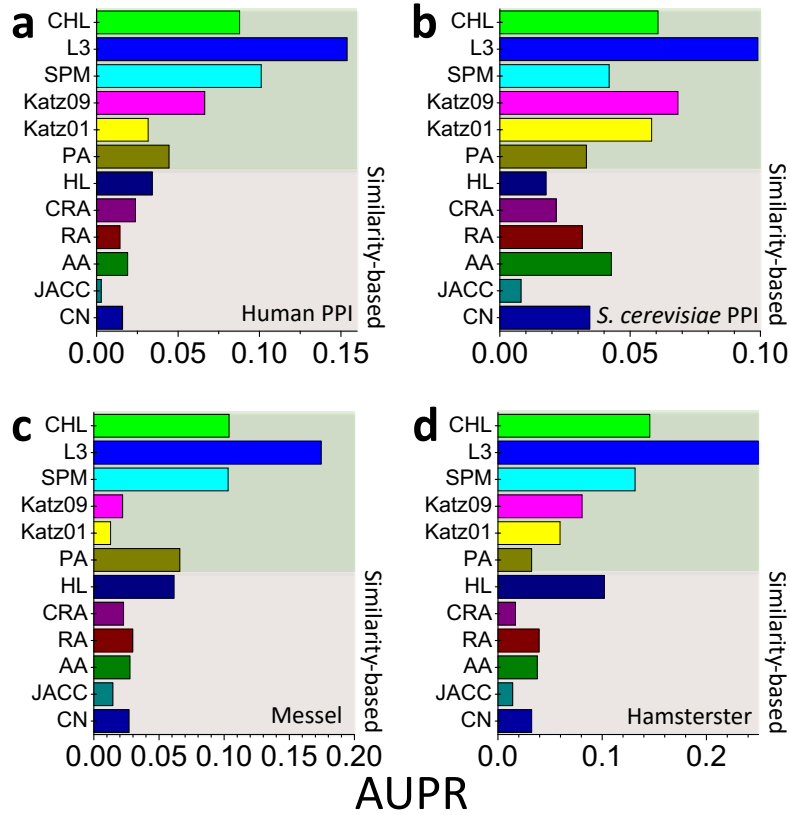
Figure 46. Link prediction with cross-geometric hyperbolic framework. Link prediction results obtained with the Complementarity HyperLink (CHL) compared to HyperLink (HL) and other representative methods in (**a**) human protein-protein interaction network, (**b**) *C. cerevisiae* protein-protein interaction network, (**c**) Messel food web, and (**d**) Hamsterster social network. All experiments correspond to the $1 - q = 0.5$ fraction of removed links. Considered link prediction methods are the (CHL) Complementarity Hyper-Link (our method), the $L3$ method (Kovács et al. 2019), the Structural Perturbation Method (SPM) (Lü et al. 2015), Katz index (Katz 1953) with parameter $\beta = 0.1$ (Katz01) and $\beta = 0.9$ (Katz09), Preferential Attachment (PA) (Barabási and Albert 1999), Hyperlink (HL) (Kitsak et al. 2019), Cannistraci Resource Allocation (CRA) (Cannistraci et al. 2013a), Resource Allocation (RA) (Zhou et al. 2009), Adamic Adar (AA) (Adamic and Adar 2003), Jaccard Index (JACC) (Jaccard 1901), and the number of common neighbors (CN) (Liben-Nowell and Kleinberg 2003).

are assigned uniformly at random, $\rho\left(\theta^t\right) = U[0, 2\pi]$, while radial coordinates are drawn from $\rho\left(r^t\right) = \frac{\sinh\left(\alpha r^t\right)}{\cosh(\alpha R) - 1}$, $r \in [0, R]$ for $t = \{1, 2\}$. Here $\alpha > \frac{1}{2}$ is the parameter controlling

the radial node density and affecting resulting degree distribution of the CRHG.

Different from the RHG, CRHG dictates that each node is characterized by two points in $\mathbb{H}^2$, and connections between nodes are established with the probability in Eq. (65), where $p(d)$ is parameterized as

$$p(d) = \frac{1}{1 + e^{\frac{d-R}{2T}}}. \tag{67}$$

Here $T \in (0, 1)$ is the temperature parameter controlling the relevance of long-distance connections, and $R > 0$ is the radius of the hyperbolic disk, controlling the expected average degree of synthetic networks.

Following the hidden variable formalism (Boguñá and Pastor-Satorras 2003), it is straightforward to express the structural properties of the CRHG in terms of those in the RHG model. Indeed, the expected degree of a node characterized by the coordinates $\{r_i^1, \theta_i^1, r_i^2, \theta_i^2\}$, to the leading order, is

$$\overline{k}\left(r_i^1, \theta_i^1, r_i^2, \theta_i^2\right) = \overline{k}^{\mathbb{H}}(r_i^1) + \overline{k}^{\mathbb{H}}(r_i^2), \tag{68}$$

where $\overline{k}^{\mathbb{H}}(r)$ is the expected degree of a node with radial coordinate r in the RHG:

$$\overline{k}^{\mathbb{H}}(r) = \frac{4N\alpha}{2\alpha - 1} \frac{T}{\sin \pi T} e^{-r/2}, \tag{69}$$

see Krioukov et al. (2010). Then, degree distribution of the CRHG is nothing else but the convolution of degree distributions in the RHG:

$$P(k) = \sum_{k'=0}^{k} P^{\mathbb{H}}(k - k') P^{\mathbb{H}}(k') \sim k^{-\gamma}, \tag{70}$$

where $\gamma = 2\alpha + 1$.

Similarly, the expected number of common neighbors between nodes i and j can be expressed, to the leading order, as

$$\overline{m}_{ij} = \overline{m}^{\mathbb{H}}\left(r_i^1, \theta_i^1; r_j^1, \theta_j^1\right) + \overline{m}^{\mathbb{H}}\left(r_i^2, \theta_i^2; r_j^2, \theta_j^2\right), \tag{71}$$

where $\overline{m}^{\mathbb{H}}\left(r_i, \theta_i; r_j, \theta_j\right)$ is the expected number of common neighbors in RHG between

149

nodes $\{r_i, \theta_i\}$ and $\{r_j, \theta_j\}$. Then, the probability $P(m)$ of two randomly chosen nodes in CRHG to have exactly m common neighbors is the convolution of those in the RHG model

$$P(m) = \sum_{m'=0}^{m} P^{\mathbb{H}}(m - m') P^{\mathbb{H}}(m') \sim m^{-\tau}, \tag{72}$$

since $P^{\mathbb{H}}(m) \sim m^{-\tau}$, where exponent $\tau = \tau(\gamma, T) > 2$, as documented both in empirical (Iamnitchi et al. 2004; Burgos et al. 2008) and theoretical (Kitsak et al. 2017) studies.

The expression for clustering coefficient in the CHRG is more involved and is omitted here for brevity, what is important is that clustering coefficient vanishes in the large N limit, as

$$\overline{c}(N) \sim N^{-1}, \tag{73}$$

consistent with the non-transitivity of complementarity.

To highlight the practical utility of the developed complementarity framework we next obtain cross-geometric hyperbolic representations of four real-world networks: the network of protein interactions in yeast (Yu et al. 2008) and human (Das and Yu 2012) and (Luck et al. 2019) cells, the Messel shell food web (Dunne et al. 2014), and the social network of the website hamsterster.com (Kunegis 2013), and use the obtained representations to predict missing links in these networks.

To this end, we remove a fraction of $1-q$ links uniformly at random in the network of interest. In doing so, we go over every existing link in the network and remove it with probability $1 - q$. We then obtain cross-geometric representation of the resulting truncated network and rank unconnected node pairs in the truncated network by the sum of the cross-distances between the corresponding points:

$$\text{rank}_{ij} = d\left(\mathbf{x}_i^1, \mathbf{x}_j^2\right) + d\left(\mathbf{x}_i^2, \mathbf{x}_j^1\right), \tag{74}$$

the smaller the rank the higher is the chance of a missing link.

Our link prediction results are summarized in Fig. 46. We observe that all similarity-based methods (lower half of each panel in Fig. 46) predict missing links in the four considered networks with accuracy lower than that of the methods that are not similarity-based (upper

half of each panel in Fig. 46). This result is expected, since real networks of interest are driven by the complementarity rule, and similar nodes are not expected to be connected. Our second observation is that the cross-geometric representation, which we label as CHL in Fig. 46, is substantially more accurate in missing link prediction that any similarity based method, including the original HyperLink. Further, CHL performs comparable with the most accurate non-similarity based methods, validating the cross-geometric framework. (Note: It has been established that the link prediction accuracy of methods that are based on hyperbolic spaces is extremely sensitive to the that of the node coordinate inference (Kitsak et al. 2019), and also depends on the measure used to quantify the accuracy of link prediction. We, therefore, expect the accuracy of link prediction to improve even further if more accurate embedding methods are developed for complementarity-based networks.)

In summary, we would like to emphasize that the proposed cross-geometric framework is not specific to hyperbolic spaces: any metric space of arbitrary dimensionality may serve as a latent space in this framework. Here we presented the *minimal* version of the cross-geometric framework, where each network node corresponds to two points in the latent space. Higher order generalizations of the framework are straightforward.

It also important to note that the proposed cross-geometric framework can not be mapped to a higher-dimensional version of the conventional latent-geometric framework, where each node is mapped to a single point. In the latter, the higher is the pairwise match between node coordinates, the smaller is the distance and the higher is the connection probability. In the former, on the other hand, the match of all node coordinates is not required: connections are established with high probability if at least one cross-distance is small. At the same time, the proposed cross-geometric framework is closely related to bipartite networks, which can be regarded as the special class of complementarity-driven networks.

The cross-geometric framework not only opens new avenues for the analysis of complementarity-driven systems such as biological networks, interdisciplinary collaboration, and food webs, it also challenges traditional approaches of network science that were initially developed for social networks and are routinely applied to other networks classes.

One example is the notion of the shortest path, which is often envisioned as a certain discrete trajectory in the network space, Fig. 47**a**. Such trajectory is possible in the cross-geometric framework: a chain of connections may form due to a spatial alignment of complementary

points into a geometric trajectory, Fig. 47**b**. While such an alignment is definitely sufficient for the formation of a network chain, it is by no means necessary: another possibility is a collection of pairwise, yet disjoint, proximities between corresponding points in the latent space, as seen in Fig. 47**c**.

Network community, in its classical formulation, is a group of nodes densely connected within and sparsely connected outside the group. Based on this definition, network communities in social sciences are often envisioned as collections of node-points localized in certain network space, Fig. 47**d-f**. While the cross-geometric framework admits this interpretation as well, Fig. 47**e**, communities that are partially localized in the latent space are also possible, Fig. 47**f**.

Shortest paths and communities have been adopted from similarity-based networks and are routinely used in the analysis of complementarity-driven networks. Network communities are routinely used to quantify disease and functional modules in biological networks (Ghiassian et al. 2015; Ahn et al. 2010; Mahmoud et al. 2014), and scientific communities in collaboration networks (Girvan and Newman 2002; Fortunato 2010). Shortest paths, on the other hand, are often used to quantify network-based separations between network modules of interest (Menche et al. 2015; Sonawane et al. 2019). While we do not question the use of paths and communities in complementarity-driven networks, we call for careful interpretation of the obtained results that should be consistent with the cross-geometric framework.
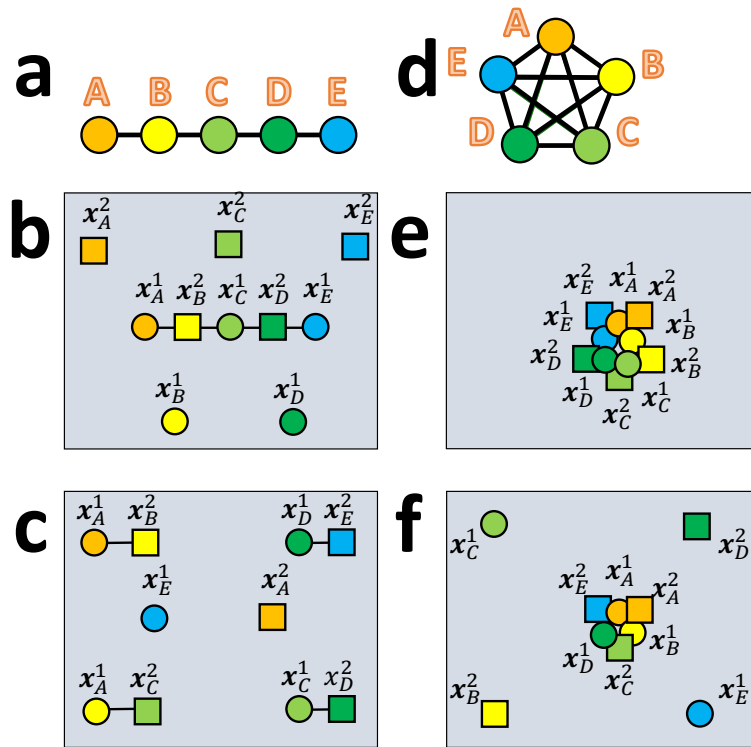
Figure 47. Paths and communities in complementarity-driven networks. Shown in (**a**) is the toy network consisting of a chain of 5 nodes as well as (**b**)-(**c**) two configuration of points in the space that might lead to it. (**b**) In this configuration the chain of connections observable in the network is achieved by aligning complementary points into a geometric trajectory in the latent space. (**c**) This configuration demonstrates how a chain of connections in (**a**) may arise from a collection of pairwise proximities between points in the latent space. In contrast, trajectory-based alignment is the only possibility for similarity-driven networks. Shown in (**d**) is the toy network consisting of a 5 node clique as well as (**e**)-(**f**) two configuration of points in the space that might lead to it. (**e**) In this configuration the clique network arises due to the complete clustering of all points. (**f**) This configuration demonstrates how the clique network in (**d**) may arise from partially clustered points in the latent space. In contrast, the complete clustering of all points in the latent space is the only possibility for the clique similarity-driven network.

5   Application to real military and civilian systems

All of the work conducted under this project has been motivated by applications in real military and civilian systems. Specific application to a maritime operations center (MOC) appears in Section 3.2. The material here is reproduced from Eisenberg et al. (2020).

Major blackouts in electric power systems are often attributed to cascading failures, where the failure of few power grid components cause large-scale loss of electricity. Despite recent innovations to model and predict cascading losses, few studies link cascades to the complex management context of blackout response. In this work, we broaden the purpose of cascading failure models to provide a heuristic way to study coordination plans to manage blackouts. We analyze the South Korean power grid with a cascading failure model and use results to generate social networks of Korean organizations that respond to national blackouts. We identify organizations critical to cascading failure response by analyzing resulting network models for power grids and social networks. Results have key implications for Korean blackout management policy as few organizations are critical hubs for coordinating almost all cascading failures (e.g., betweenness $\geq$ 0.2). We also identify highly connected organizations involved in few cascading failures that may be able to provide backup support in case critical hubs are unavailable. Together, this work contributes methods to assess how emergency coordination can extend during cascades and demonstrates methods with an initial assessment for South Korea.

5.1   Introduction

One of most important research innovations for measuring the vulnerability of electric power systems is the cascading failure model (Guo et al. 2017; Hines and Rezaei 2016). Major blackouts are often attributed to cascading failures, where the loss of few power grid components cause additional overloads, system imbalances, failures, and eventually large-scale loss of electricity. The 2003 blackout across the US and Canada was one of the first well-documented cascading failure events, where small losses caused by trees touching power lines, demand imbalances, and out of date operational models eventually led to the largest blackout in North American history (Andersson et al. 2005). Since 2003, numerous large-scale blackouts occurred across the world initiated by diverse events like natural disasters and cyber attack (Veloza and Santamaria 2016). Even with a growing number of documented cascading failure events, there is limited information to form a generalizable

understanding of cascading failures in power grids. This makes cascading failure events some of the most difficult decision-making contexts for power grid operators and managers charged with keeping systems running.

To make sense of cascades, experts turn to cascading failure models as a means to study different failure mechanisms and predict expected losses. There are three general types of models, each relating power grid vulnerability to a different aspect of power grids: network-based models that relate vulnerability to system structure and connectivity, dynamics-based models that relate vulnerability to electric power flow redistribution, and component-based models that relate vulnerability to joint failure probabilities and stochastic losses (Vaiman et al. 2012; Petersen et al. 2015). Prominent models in literature include the Crucitti-Latora-Marchiori (CLM) model (Crucitti et al. 2004; Kinney et al. 2005), Motter-Lai (Motter and Lai 2002), Oak Ridge–Pserc–Alaska (OPA) (Dobson et al. 2001; Carreras et al. 2013), Hidden Failure (Chen et al. 2005), the Manchester Model (Guo et al. 2017), Branching Process (Vaiman et al. 2012), and CASCADE (Guo et al. 2017). Several reviews summarize model development, applications, validation in power grids (Guo et al. 2017; Hines and Rezaei 2016; Cuadra et al. 2015; Vaiman et al. 2012; Baldick et al. 2008). These cascading failure models represent an important means to improve power grid management by estimating losses when cascades occur and guiding protection decisions for vulnerable components.

With a growing library of tools for predicting and controlling cascading failures, the question remains: do these models help people manage cascading failures better? Currently, the blackout management guidance offered by these models is limited because almost all studies focus on predicting failure consequences without considering human actions during cascades. While predicting consequences helps identify critical infrastructure components, it may do little to support management of failures as they happen (Eisenberg 2018). Some power grid cascades occur on a sub-second timescale faster than humans can respond. Here, hardening activities guided by cascading failure models are essential. However, the majority of large-scale failures happen on the minute and hour timescales and elicit human response (Vaiman et al. 2012). Power grid system operators are not complacent when infrastructure fails, they act according to reliability standards and protocols to arrest cascades and protect critical systems (Eisenberg et al. 2017). Common actions taken include remedial efforts to adjust power flows, protect vulnerable infrastructures, prepare backup resources, and

155

coordinate with relevant owner and regulator organizations (Knight 2001). Operators and managers are also trained in a number of worst-case scenarios and develop heuristics for managing known instabilities within their systems. Current cascading failure models do not consider these human actions and assume that failure operations are similar to normal grid operations.

However, cascading failures are not normal events, and are precisely when contingency management practices stretch to their limits. Cascades often occur because power system failures happen or propagate in unexpected and unanticipated ways. For example, cascades can happen due to surprising events that fundamentally change how system operators understand system function (Eisenberg et al. 2019). The losses experienced during the 2003 North American blackout exemplify this form of surprising event. Extreme events also cause cascades that may be anticipated, but exceed design thresholds, like Superstorm Sandy (QER Task Force 2017), Hurricanes Irma and Maria (Zorrilla 2017; Alderson et al. 2018), and the 2019 National Grid blackout in the United Kingdom (Karasz 2011). Cascading failure models should be oriented to study what people do during these surprising events to guide decision-making and update outdated emergency management practices.

Representing the human-in-the-loop is important because adaptive human actions are often the reason systems are saved or lost during crisis (Choi et al. 2019; Thomas et al. 2019; Hollnagel et al. 2006a). In best-case scenarios, humans follow established rules and remedial actions to manage anticipated failures and/or adapt to manage cascades in new and inventive ways. In worst-case scenarios, grid operators fail to follow established protocols or take adaptive actions that exacerbate cascading losses. In both situations, it is the adaptive actions taken by people that change the outcome of cascades, neither of which are considered in failure models. An important area of research is refocusing cascading failure models to study how grid operators and managers extend systems when faced with surprise.

A key focus of this work is to identify ways to *extend* interorganizational coordination plans for blackout management when experiencing a cascading failure. We define extensibility as the way people leverage new information and past experience to come up with novel solutions to unforeseen and unknown surprises (Woods 2015, 2018). Extensibility in power grids is not random — it is based on training, guidelines, and partnerships combined with situational awareness held by infrastructure operators as events occur (Eisenberg 2018). Assessing and

improving extensibility requires knowledge of the social context surrounding crisis decision-making alongside sociotechnical processes of sensing, anticipating, adapting and learning (Thomas et al. 2019; Seager et al. 2017a; Park et al. 2013a). Only very recently are studies trying to understand how human factors improve or exacerbate losses (Eisenberg et al. 2019; Veloza and Santamaria 2016). Still, there is limited knowledge of the social context that influences crisis decision-making when cascades occur, including utility customer contracts, local operational and management practices, economic constraints, regulations, and organizational culture.

In this section, we broaden the purpose of cascading failure models from measuring consequences to providing a heuristic way to study how organizations plan to manage large-scale blackouts. Specifically, we link the a cascading failure model for power grid networks with social network analysis to identify critical organizations and partnerships that may improve blackout management in South Korea. We start by studying cascades in the South Korean Power Grid (KPG) using a modified CLM model similar to that in Kim et al. (2017), but improved to include electrical information about power grids (Arianos et al. 2009; Bompard et al. 2010, 2011). We measure failure consequences at the system- and component-scale to relate cascades to the South Korean blackout management context (Eisenberg et al. 2017). These results are used to construct social networks representing the information sharing and coordination partnerships among South Korean companies during blackout response. We identify organizations critical for managing blackouts and supporting extensibility using measures of organizational betweenness (Lee et al. 2017). Results suggest new policies may help large-scale blackout management in South Korea.

5.2 Blackout Management in the South Korean Power Grid

South Korea occupies the southern portion of the Korean Peninsula. The South Korean Power Grid (KPG) is an islanded power system comprised of a primary grid that connects Korean mainland provinces and independent governing cities together (see: Figure 48). The KPG also has a separate independent power grid on Jeju-Do island that connects to the mainland via undersea power cables. For the purposes of this work, we focus on blackout management in the mainland grid and ignore blackouts on Jeju-Do. KPG infrastructure spans a wide-range of voltage classes, from Extra High Voltage (765 kV) to pre-distribution voltage (3.3 kV) powerlines, transformers, reactors, and substations (Kim et al. 2017). KPG

electricity is generated primarily from a combination of fossil fuels and nuclear power, with only ∼ 7% of power generation coming from hydroelectric or renewable sources depending on seasonal and demand variations (Park et al. 2013b).
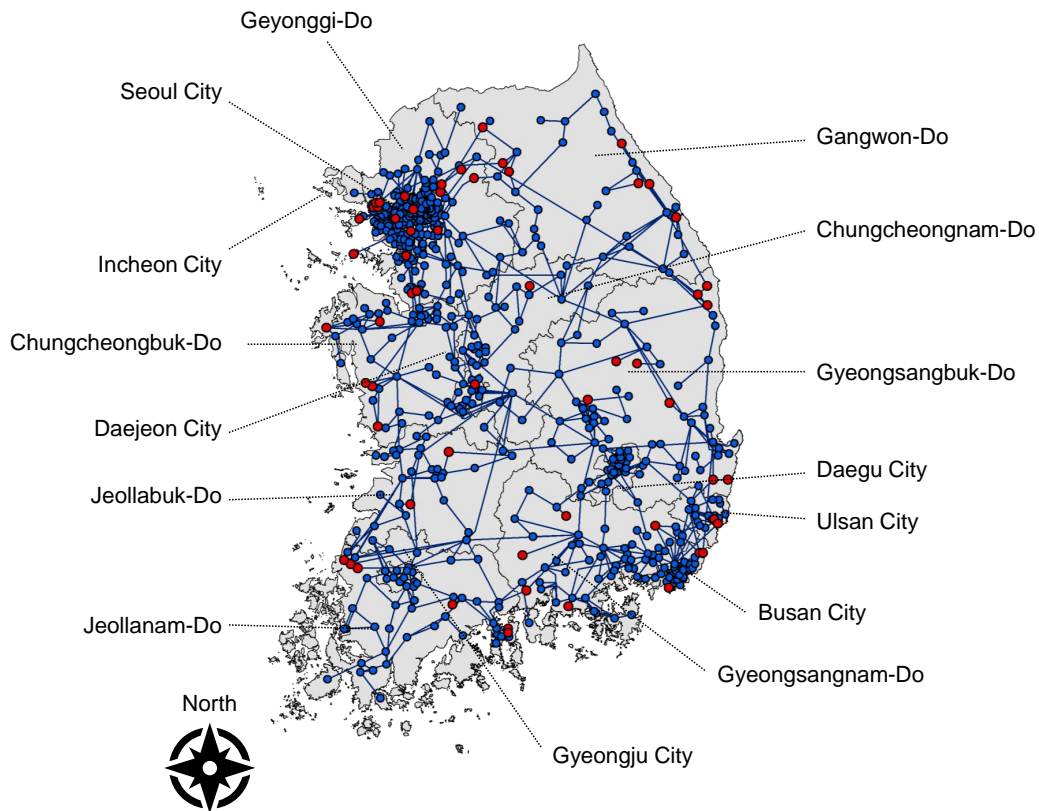
Figure 48. **Map of the 2013 South Korean Power Grid.** This map shows the connectivity of major power grid infrastructure in mainland Korea, that is, power plants (red circles), high-voltage transmission substations (dark blue circles), and power lines (dark blue lines). Transformers connect buses that are too close together to be shown in this image. South Korean provinces and major cities discussed in this work are labelled for clarity. Jeju-Do Island is excluded from the image because it is not considered in the current analysis. Figure reproduced from Eisenberg et al. (2017) under Creative Commons Attribution License (CC BY 4.0).

South Korean blackout management involves the coordination of three distinct organizational entities to respond to failures (Eisenberg et al. 2017). The first kind of entity is the

blackout response groups within KPG infrastructure owners and operators. In the KPG, essentially all power transmission and distribution infrastructure is owned by the Korean Electric Power Corporation (KEPCO) and operated by the transmission service operator, the Korea Power Exchange (KPX). Power generation facilities and associated infrastructure (e.g., transformers) are owned and operated by private corporations, where the largest 6 operate and manage ~ 97% of all generation assets. These companies are subsidiaries of KEPCO: Korea Midland Power (KOMIPO), Korea Western Power (KOWEPO), Korea East-West Power (EWP), Korea Southern Power (KOSPO), Korea Southeast Power (KOSEPO), and Korea Hydro-Nuclear Power (KHNP). Other major KPG infrastructure owners and operators include the Korea Water Administration (Kwater), that owns and operates all multipurpose dams in the country and some private power producers that generally serve heavy industries or industrial sites (e.g., POSCO Power, SK Energy, Korea District Heat Corporation, etc.).

The second kind of organizational entity is the emergency management offices that support power grid owners and operators when blackouts occur. These offices are generally firefighter operations centers or special central headquarters that coordinate emergency response to critical infrastructure failures. Each of the 8 mainland Korean provinces and 9 independent cities have their own emergency response headquarters in charge of blackout management, each with a unique name. For example, the state surrounding the Korean capital city of Seoul, Gyeonggi-Do, has a combined fire disaster headquarters (GGD-FDHQ) that supports blackout response. This HQ is distinct from the equivalent support entity for Seoul – i.e., the Seoul City Fire Department Headquarters (SC-FDHQ) – as well as for nearby provinces and cities – e.g., Gangwon-Do Fire Headquarters (GD-FHQ) and the Incheon City Fire Safety Management Office (IC-FSMO).

Finally, blackout management also includes federal agencies involved in regulation and coordination of remedial actions across sectors. The most prominent ministry involved in blackout management is the Ministry of Trade, Industry, and Energy (MOTIE). MOTIE both regulates the Korean power industry and provides decision-making support for large-scale power failures. Other related agencies at the time of analysis include the National Emergency Management Agency (NEMA) and the Ministry of Security and Public Administration (MOSPA), both of which coordinated large-scale crisis response for South Korea. These two entities are now part of the same ministry: the Ministry of Public Security and

Safety (MPSS). See Eisenberg et al. (2017) for the full list of these federal agencies and organizations.
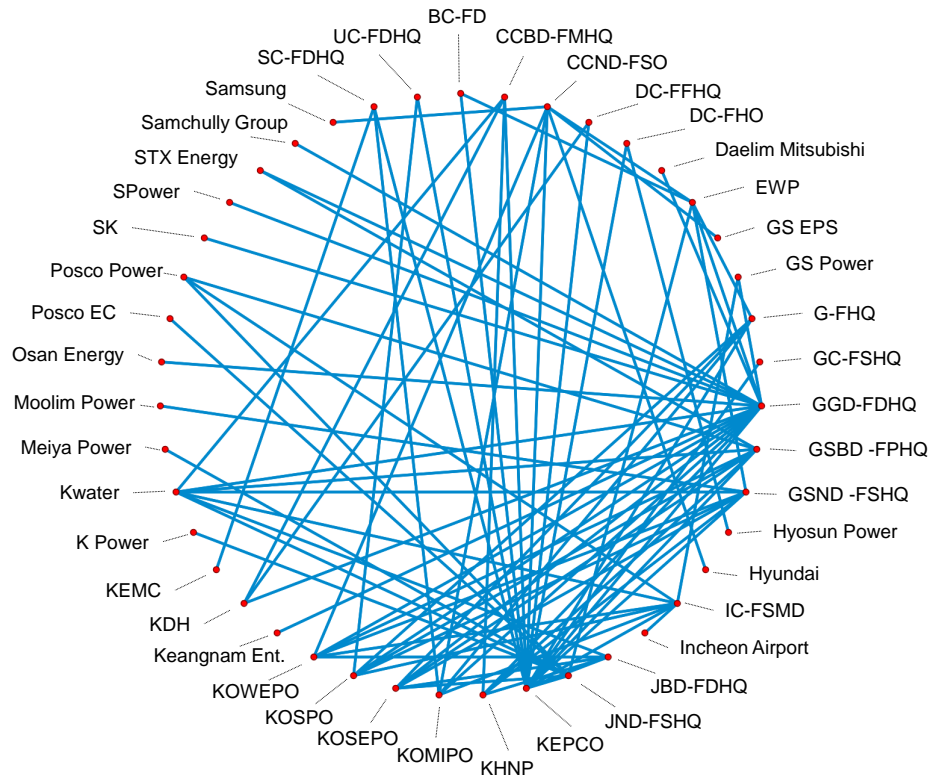
Figure 49. **Social Network of Korean Blackout Management Organizations.** Infrastructure owners, operators, and emergency responders coordinate to manage power grid failures across South Korea. Blue links represent information sharing partnerships that link organizations when problems occur as defined in national blackout management protocols harmonized across the power industry. We use this social network model as a basis for studying blackout management contexts during cascades. Note: Based on data from Eisenberg et al. (2017).

Figure 49 presents the basic social network of partnerships among Korean power and emergency management organizations prior to national emergencies (Eisenberg et al. 2017). Figure 49 is comprised of 43 companies, including transmission infrastructure owners, major power producers (i.e., with greater than 10% of national generation capacity), smaller

power companies, and state-level emergency management headquarters for organizing first responder support to power losses. Not all organizations are involved in all blackouts. Instead, there are three key factors that dictate which organizations are involved in managing a given blackout:

1. *The infrastructure directly affected by the blackout:* As cascades occur, infrastructure will fail and overload. Both automatic relays and operator actions can be taken to remove these overloaded assets from service to protect them from further damages. However, these actions also require active decision-making to turn overloaded and offline systems back on. Organizations that own and operate failed and overloaded infrastructure coordinate these actions via command centers and headquarters.

2. *The geospatial location of the blackout:* The areas affected by cascades require coordination with local emergency managers and decision-makers to determine the best course of action to protect communities and support infrastructure recovery. KPG owners and operators coordinate with local emergency management entities to ensure blackout recovery is safe and efficient.

3. *The size of the blackout and its impact on national reserve margin:* The larger the scale of the blackout, the greater the need to escalate blackout management decision-making to higher levels within the industry and government. South Korean blackout and emergency plans use *national reserve margin* as a characteristic measure for determining how to escalate decision-making. Reserve margin refers to excess available power generation in the KPG to pick up slack and fluctuations in infrastructure availability and demand. Within Korean blackout management plans, reserve margin is measured in total gigawatts (GW) of excess generation capacity, where lower values of reserve margin require higher levels of decisions-making (e.g., city, provincial, and national).

The social network in Figure 49 is an idealized representation of the interorganizational coordination context expected in South Korea prior to the onset of a large-scale failure. During cascading failures, only small portions of this social network are relevant based on the factors described above. The purpose of this work is to use cascading failure models on the KPG to generate these sub-networks and determine which coordination and information-sharing partnerships are most important for a diversity of infrastructure failures. Specifically, we study cascades caused by N-1 failures (i.e., the single failure of any

161

component within the KPG) and generate associated social networks based on Figure 49.

Note: Although federal agencies play an important role in situations where reserve margin drops to increasingly dangerous levels, the KPG is designed to be N-1 reliable so no single infrastructure failure triggers a national emergency. Hence, Figure 49 and associated sub-networks considered in this work do not include the role of agencies like MOTIE, NEMA, or MOSPA. For detailed analysis of coordination during more extreme failure situations, please refer to Eisenberg et al. (2017).

5.3 Methods

This work uses well-studied models of the South Korean power grid (KPG) network and social network of organizations involved in Korean blackout management. KPG data was provided directly by KEPCO as a PSS/E (power system simulation for engineering) file, and the model consists of 2083 power system buses and 4167 power lines and transformer links. We use the following node and link sets to generate the KPG network: power lines and transformers that transmit electricity ($L \in M_{Lines}$), buses that produce electricity ($G \in N_{Gen}$), buses that demand power ($D \in N_{Demand}$), and transmission buses that direct power flow ($T \in N_{Trans}$). A simplified structure of this complex power grid network based on publicly available data is presented in Figure 48.

Social networks were generated using Matlab and analyzed using ORA-LITE social network analysis software (Altman et al. 2017). We use the set of nodes ($O \in N_{orgs} = N_{pc} + N_{em}$) and links ($P \in M_{orgs}$) to distinguish blackout management social networks from power grid networks (see Figure 49).

Data for both models relates to the years of 2013-2014 during summertime peak load. All power flow analysis was conducted in Matlab with the Matpower package Zimmerman et al. (2010) and the GLPK optimization solver (Makhorin 2014). We refer the reader to Kim et al. (2017) and Eisenberg et al. (2017) for detailed descriptions of power grid and social network models and general metrics characterizing their structure.
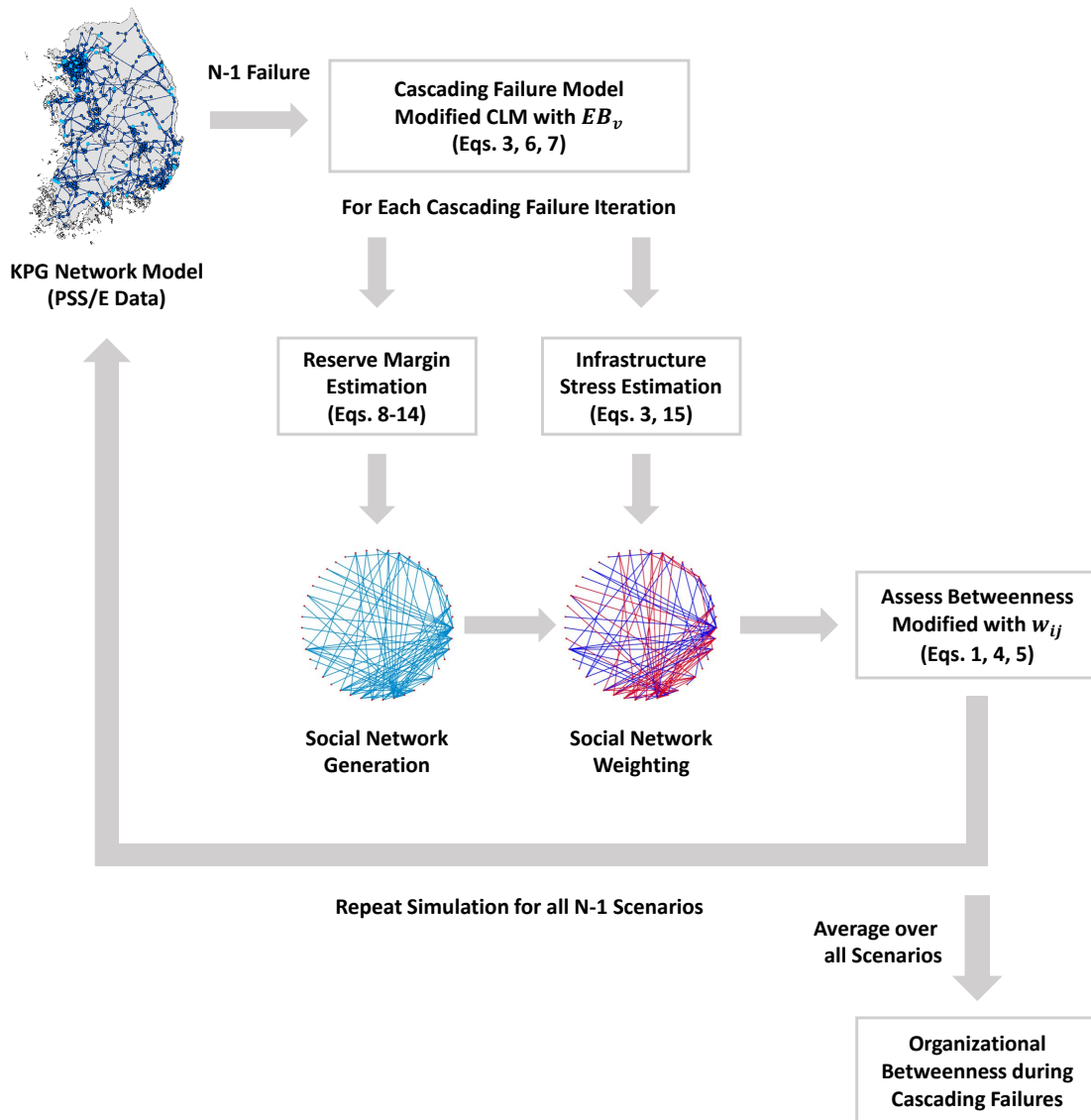
Figure 50. **Overview of Simulation Methods to Link Cascading Failures to Social Network Analysis.** This work combines methods for power grid and social network analysis into a integrated simulation approach. First, we generate an infrastructure failure and measure cascading losses. For each iteration in the cascade, we estimate failure impacts using system-level (reserve margin) and component-level (infrastructure stress) measures. These measures inform the generation and weighting of social networks for Korean blackout management. We repeat this process and average social network results across infrastructure failure scenarios.

Figure 50 presents an overview of the simulation methods used to link cascading failures to social network analysis. First, we generate infrastructure failures in the KPG network that may cause blackouts and measure their potential impacts with a cascading failure model. For each model iteration, we assess the stress on the KPG network to reveal the size of the blackout and the infrastructure components directly affected by the blackout. These two assessments inform the structure of the blackout management social network and the importance of each coordination partnership. Using the generated social network, we identify critical organizations using established social network analysis methods. The simulation is repeated and social network analysis results are averaged across all KPG failure scenarios.

Methods and simulation results were validated with expert feedback from Korean blackout management organizations. Specifically, methods to model cascading failures and measure power system stress were coordinated with experts in KEPCO. Methods to generate and analyze blackout management social networks were coordinated with experts in MPSS.

5.4 Identifying Critical Infrastructure and Organizations

5.4.1 Network Science Methods for Power Grids and Social Networks

A fundamental assumption in this work is that infrastructure and social networks correspond and can be studied in a consistent way (Eisenberg et al. 2017). We analyze the KPG with respect to the service it provides — the delivery of electricity from generation to use. This generation-demand relationship corresponds to social network theory via "package"-based flow processes (Borgatti 2005). Unlike other social process such as gossip that transfer information among actors in an unregulated way, packages are assumed to have explicit destinations. Information sharing and decision-making among infrastructure crisis managers follow a package-based flow structure. This is because command and control protocols dictate which information can and should be shared among organizations as well as the specific people and organizations who have decision-making authority (Eisenberg et al. 2018b).

164

5.4.2 Network Science-Based Measure of Critical Nodes

The package-based treatment of power grids and blackout management systems allows us to use the network-science measure of *betweenness* to identify critical infrastructure and organizations. Betweenness is a measure of the flow contribution of network elements and is an established way to rank the importance of power grid and social network nodes (Eisenberg et al. 2017). The generic form of betweenness is based on geodesic paths (or shortest paths) from nodes i to j. The set of all geodesic paths between nodes i and j is called the "minimum cut set," σ_{ij}. The betweenness of any given node v is then defined as the total number of geodesic paths that node v lies on normalized by the size of each minimum cut set (Brandes 2001):

$$B_v = \sum_{i \neq v \neq j} \frac{\sigma_{ij}^v}{\sigma_{ij}} \tag{75}$$

where σ_{ij}^v is the size of the cut set between nodes i and j that node v is on.

5.4.3 Measuring Power Grid Criticality via Electrical Betweenness

The generic form of Eq. 75 is often inappropriate for power systems vulnerability analysis as electric power does not flow based on geodesic path, it flows based on power line impedance (Ouyang 2013; Hines et al. 2010). To address this discrepancy, several authors developed power grid betweenness measures that incorporate the physics of power flow. Here, we use electrical betweenness originally proposed in Arianos et al. (2009) to determine nodes that influence power flow. This measure uses linear shift factors (also called power transfer distribution factors (Bompard et al. 2009)), f_i^{gd}, for each power line, $l \in L$, for a unit injection at generation bus, $g \in G$, and an equal increase in load at demand bus, $d \in D$. These linear shift factors determine the relationship between KPG structure and power flow for all generation-demand pairs. The linear shift factors are then used to estimate a total transfer capability for a single power line before it overloads, TTC_g^d, by assessing the maximum allowable power transfer for all g and d pairs:

$$TTC_g^d = \min \left\{ \frac{P_1^{max}}{f_1^{gd}} \cdots \frac{P_l^{max}}{f_l^{gd}} \cdots \frac{P_M^{max}}{f_M^{gd}} \right\} \tag{76}$$

The electrical betweenness of a given power system bus v, is then defined as Arianos et al. (2009); Bompard et al. (2011); Eisenberg et al. (2017):

$$EB_v = \frac{1}{2} \sum_{g \in G} \sum_{d \in D} TTC_g^d \sum_{l \in L^v} |f_l^{gd}|, g \neq v \neq d \tag{77}$$

where L_v is the set of power lines attached to bus v and the factor of $\frac{1}{2}$ accounts for power flow into and out of buses. The term $\frac{1}{2} TTC_g^d \sum_{l \in L^v} |f_l^{gd}|$ is interpreted as the security constrained contribution of power flow of bus v for a single g and d pair, where the sum over all g and d pairs measures the contribution of each bus to power flow across the entire grid.

5.4.4 Measuring Social Network Criticality via Weighted Betweenness

Like in power grids, the measure of betweenness in Eq. 75 does not necessarily measure organizational importance in social networks. This is because links in social networks are not "durative" like transmission lines in a power grid, and only exist when used. Social networks are often studied using weighted links that represent the strength of person-to-person and organization-to-organization connections (Butts 2009). For the purposes of this work, Eq. 75 must be modified to include the strength of partnerships between organizations as weights, w_{ij} on network links. In social network analysis, w_{ij} generally relates to the *strength* of a social relationship estimated as the average frequency that the relationship is used to share information. For the purposes of this work, w_{ij} represents how often a given blackout information sharing and coordination partnership is used to manage cascading failures averaged across multiple scenarios. Links with large weights, $w_{ij} \approx 1$, represent information sharing and coordination partnerships that are required for many cascading failures scenarios, where low weights, $w_{ij} \approx 0$, represent partnerships that are rarely used to coordinate blackout response. (Methods for calculating weights presented below.)
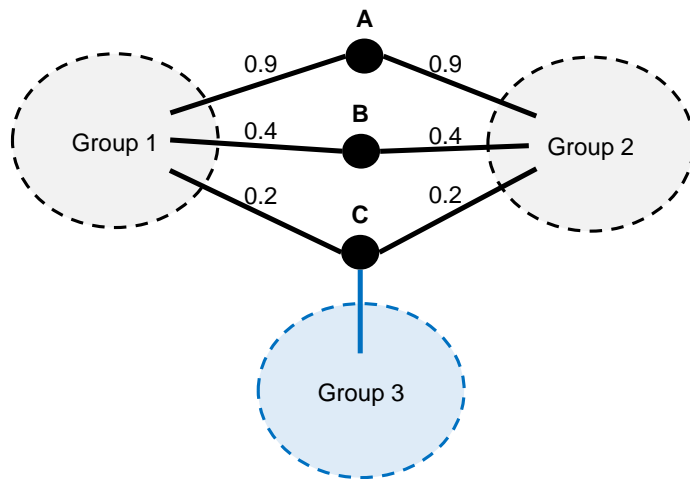
Considering Groups 1 and 2 only

$$B_A = \frac{1}{3}; B_B = \frac{1}{3}; B_C = \frac{1}{3}$$

$$B_A^1 = 1; B_B^1 = 0; B_C^1 = 0$$

$$B_A^2 = 0; B_B^2 = 0; B_C^2 = 1$$

Considering Groups 1, 2 and 3

$$B_A = \frac{1}{3}; B_B = \frac{1}{3}; B_C = 2\frac{1}{3}$$

$$B_A^1 \geq 1; B_B^1 = 0; B_C^1 = 2$$

$$B_A^2 = 0; B_B^2 = 0; B_C^2 = 3$$

Figure 51. Practical Differences among Weighted Betweenness Measures. Consider groups 1 and 2 linked by nodes A, B, and C (ignoring group 3). B_v is a purely structural measure, and nodes A, B, and C receive equal values. When measures include the link weights shown, B_v^1 denotes node A as the most between and B_v^2 denotes node C as the most between. With only two connected groups B_v^1 and B_v^2 are inverse measures of each other. However, in more complex systems where group 3 exists, B_v^1 and B_v^2 will reveal characteristically different information about the social network independent of the weight of link C to group 3. The interpretation of this more complex scenario in social network theory is B_v^1 captures optimal information sharing across well-established ties, where B_v^2 captures least optimal, yet feasible information sharing across weak ties.

We use two betweenness-based measures to identify critical organizations in weighted social networks (shown graphically in Figure 51). The first identifies organizations that connect others via strong ties, i.e., partnerships required for many cascading failure situations.

These organizations are identified with the weighted betweenness, B_v^1, that has the same form as Eq. 75, but is computed with cut sets, θ_{ij}, that use link distances based on weights, $d_{ij}^{avg} = \frac{1}{w_{ij}}$:

$$B_v^1 = \sum_{i \neq v \neq j} \frac{\theta_{ij}^v}{\theta_{ij}}. \tag{78}$$

The second identifies organizations that connect others via weak ties, i.e., partnerships that exist but are infrequently used for managing cascading failures. Organizations with these partnerships are well-connected, but rarely involved in blackout management. Thus, they can extend their current operations to support blackout coordination when needed. To identify these organizations, we use weighted betweenness, B_v^2, which has the same mathematical form as Eq. 75, but is computed with cut sets, ϕ_{ij}, based on weights, $d_{ij}^{avg} = w_{ij}$:

$$B_v^2 = \sum_{i \neq v \neq j} \frac{\phi_{ij}^v}{\phi_{ij}}. \tag{79}$$

5.4.5 Linking Power Grid Cascades to Social Networks

Network Science-Based Cascading Failure Model for Power Grids. Betweenness is a key measure used in network-based cascading failure models. In particular, the Crucitti-Latora-Marchi (CLM) cascading failure model (Kim et al. 2017; Kinney et al. 2005; Crucitti et al. 2004) centers on using betweenness to understand how congestion within a system is shifted relative to overloaded nodes. This model has been used to assess cascades on a number of power systems (Kinney et al. 2005), and was more recently used to assess the resilience of the KPG (Kim et al. 2017). The CLM model is useful for understanding how people may respond to cascades because it uses a simple process to determine how power may be shifted around after an infrastructure failure to find a stable and safe operating state.

The CLM model uses betweenness, B_v, and a measure of link efficiency, e_{ij}, to determine how remedial actions might be taken (Crucitti et al. 2004). Here, $e_{ij}(t)$ is the inverse of the distance between two nodes at a given time interval, t, where $e_{ij}(t=0) = \frac{1}{d_{ij}(0)} = 1$. The CLM model dynamically adjusts link efficiencies to mimic cascades. This cascading

failure process progresses in the following way:

First, the capacity of each node is estimated as its initial betweenness at time, $B_v(t = 0)$, tuned by a capacity parameter, α:

$$C_v = B_v(t = 0) * (1 + \alpha). \tag{80}$$

A node is removed from the network at $t = 0$ and link efficiencies are recalculated for each iteration based on the following equation:

$$e_{ij}(t) = \begin{cases} \min\left(\frac{C_i}{B_i(t)}, \frac{C_j}{B_j(t)}\right); & if \quad B_i(t) < C_i \quad or \quad B_j(t) < C_j \\ e_{ij}(t = 0); & otherwise \end{cases} \tag{81}$$

We build upon the CLM model in Eq. 81 for studying congestion-based cascades in power grids using the electrical betweenness EB_v and key engineering parameters that dictate electric power flow. Specifically, we replace B_v with EB_v and then calculate Eq. 80 assuming the power flow limits found in current KPG data are sufficient, i.e., $\alpha = 0$. We then replace link efficiency, e_{ij}, with power line impedance to force congestion re-balancing to follow the physics of power flow. For all calculations, we use the DC power flow approximation which simplifies impedance to be the power line reactance, x_{ij}, such that, $x_{ij}(t) = \frac{x_{ij}(t=0)}{e_{ij}(t)}$. The finalized model produces more realistic congestion-based power grid cascades than the original CLM model by redistributing electricity based on power flow physics (Ouyang 2013). Thus, this model provides a heuristic method to estimate the ways automated systems and power grid operators will respond to failures as they happen. See Henneaux et al. (2018) to compare our implementation with similar cascading failure models.

Linking Cascading Failure Impacts to Social Networks. Relating cascading failure results to social network generation requires we estimate the national reserve margin for South Korea. Reserve margin is a function of both the available generation capacity and the capability to transmit electricity to customers. To estimate reserve margin, we first estimate the total transfer capability (NERC 2005, 1999; Sauer 1998) of the KPG, TTC_{net}, summing over all g and d pairs from Eq. 76,

$$TTC_{net} = \frac{1}{N_{Gen}N_{Demand}} \sum_{g \in G} \sum_{d(g \neq d) \in D} TTC_g^d \tag{82}$$

We then calculate the active transfer capability, $ATC_{net}(t)$, for each timestep which is qualitatively similar to TTC_{net} but also considers congestion within the power grid estimated as active power injections,

$$ATC_g^d(t) = \min \left\{ \frac{P_1^{max} - P_1(t)}{f_1^{gd}} \cdots \frac{P_l^{max} - P_l(t)}{f_l^{gd}} \cdots \frac{P_M^{max} - P_M(t)}{f_M^{gd}} \right\} \tag{83}$$

$$ATC_{net}(t) = \frac{1}{N_{Gen}N_{Demand}} \sum_{g \in G} \sum_{d(g \neq d) \in D} ATC_g^d(t) \tag{84}$$

where $P_l(t)$ is the real power flow over line l at time t. We then estimate the transmission reserve margin , $TRM(t)$, as reductions in available generation due to system congestion that may inhibit the dispatch of additional generation resources in Korea. First, we relate reserve margin to transfer capability using the standard equation for power system planning (NERC 2005, 1999),

$$ATC = TTC - TRM \tag{85}$$

and rearrange this equation to relate reductions in generation capacity to system congestion with the following equations:

$$AGC = IC - TD \tag{86}$$

$$UGC(t) = AGC * (1 - \frac{ATC_{net}(t)}{TTC_{net}}) \tag{87}$$

$$TRM(t) = AGC - UGC(t) \tag{88}$$

where AGC is available generation capacity, IC is installed generation capacity, TD is the total demand, $UGC(t)$ is the unavailable generation capacity due to system congestion at

cascade iteration t (*note:* all calculations are in units of GW).

This measure of TRM assumes that the dominating factor preventing power dispatch are the thermal limits on power lines, which is reasonable for power system emergencies on transmission grids. Eqs. 82–88 also assume that demand, TD, remains constant throughout a cascading failure event, which is a common assumption across all cascading failure models (Hines and Rezaei 2016).

Social Network Weighting Based on Power Grid Cascades. Using the methods described above, we simulate cascading failures in the KPG with the modified CLM model and generate weighted social networks to study blackout management. We initiate a cascading failure by removing a single node in the KPG network and run the cascading failure simulation until a new stable operating state is achieved (see Figure 50). Then, we generate a subset of "active" Korean power organizations for each time step during the cascade to measure the frequency that a given blackout management partnership is used. We define "active" power organizations, $N_{orgs}^{active} \in N_{orgs}$, as organizations that own infrastructure stressed by the cascade or are geographically located in the same region as infrastructure stressed by the cascade. We define stressed power grid infrastructure as either the node removed from the network at $t = 0$ or as a power system bus v at time step t, $S_v(t) > 1$, calculated by,

$$S_v(t) = \begin{cases} \frac{EB_v(t)}{EB_v(t=0)}, & if \quad EB_v(t = 0) \leq EB_v(t); \\ 0, & otherwise \end{cases} \tag{89}$$

We use active power grid buses to generate weighted social networks. Each bus at a given cascade iteration t with $S_v(t) \geq 1$ is considered "active" and put in set $N_{bus}^{active} \in N_{Gen} + N_{Demand} + N_{Trans}$. These active buses are then related to the power industry by establishing the subset of power companies that own and operate stressed infrastructure, $N_{pc}^{active} \in N_{pc}$, and the subset of emergency management agencies that will support infrastructure response and recovery based on the geographic region stressed infrastructure is located in, $N_{em}^{active} \in N_{em}$. Then, all partnerships among active organizations, $M_{orgs}^{active} \in M_{orgs}$, are added to the social network and given equal weight, $w_{ij} = 1$. The final, weighted social networks are generated by taking the mean of each active link across all N-1 failure scenarios. Thus, all link weights are $0 \leq w_{ij} \leq 1$, where a partnership with $w_{ij} = 0$ means the partnership was

171

never relevant to N-1 blackout management, and $w_{ij} = 1$ means the partnership was active across all failure scenarios.

5.5 Results

5.5.1 Cascading Failure Interoganizational Context in South Korea

Results from our cascading failure model show that N-1 failures do not cause significant disruptions in national reserve margin (Figure 52). Korean blackout management policies use a series of reserve margin thresholds to establish which organizations serve decision-making and coordination roles, and each threshold relates to a different social network (Eisenberg et al. 2017). The KPG is N-1 reliable, and at no point does any N-1 failure lead to the activation of national crisis management thresholds that require the inclusion of federal organizations for decision-making and coordination across industries.

N-1 losses lead to a re-weighting of the social network in Figure 49 based on which organizations are involved after infrastructure is lost. Our model estimates the redistribution of electricity that forces some power infrastructure to be overloaded or near overload post failure. In most cascading failure models, these overloaded assets are then assumed to fail and cause additional damages. We assume human actions manage the additional cascading losses and allow for the system to correct these imbalances. We initially weight the "pre-failure" social network based on the amount of infrastructure each organization is involved in managing during blackouts and their position within the social network. Then, we use the shifting blackout risk during corrective actions to estimate the subset of blackout coordination partnerships that would be active during failure and cascade iterations. The combination of social and power grid data produces weighted social networks that reveal which organizations are more important for blackout management.
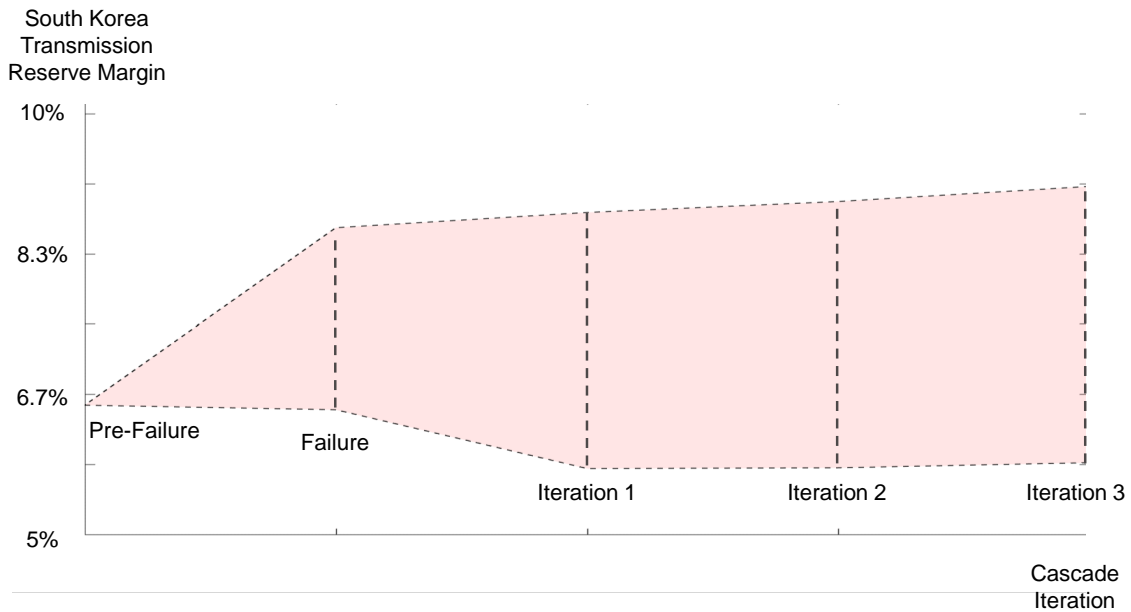
Figure 52. Cascading Failure Results for South Korean Power Grid (KPG). When a single substation or generation bus fails, electricity re-distributes across the KPG and total dispatchable reserve margin changes. The light red region in the above graph encompasses all reserve margin gains and losses from N-1 failures in the South Korean power grid, where top and bottom dashed lines are the largest gains and losses. These results are used to generate weighted social networks. Social network links form based on which organization owns and operates stressed infrastructures, where stressed infrastructures are located, and how stressed they are during cascades. The weight of each social network link using this method is the mean weight across all N-1 failure scenarios with a normalized maximum partnership strength of 1.

5.5.2 Guiding Interorganizational Coordination during Blackouts

Figure 53 shows important organizations that are involved in the KPG social network before, during, and after infrastructure cascading failures via normalized unweighted betweenness, B_v, and normalized weighted betweenness, B_v^1, measures. Figure 53 shows KEPCO and the Gyeonggi-Do Fire Disaster Headquarters (GGD-FDHQ) are the most central organizations for all cascading failures with $B_{KEPCO}^1 \approx 0.8$ and $B_{GGD-FDHQ}^1 \approx 0.4$ over the course of cascades. This is because KEPCO is involved in any event that takes a major transmission line out of service and GGD-FDHQ is involved in almost all failures that affect customers due to the large number of people living in the area surrounding Seoul.

Figure 53 also shows the shifting importance of organizations with lower betweenness values over the course of cascading failures. Prior to a failure, the Gyeongsangbuk-Do Fire Protection Headquarters (GSBD-FPHQ) and Jeollanam-Do Fire Safety Headquarters (JND-FSHQ) are two of the most central organizations for blackout coordination with $B_{GSBD-FPHQ}^1$ and $B_{JND-FSHQ}^1 \approx 0.15$. This is due to the large number of KPG infrastructure in each region. However, during cascades the betweenness for GSBD-FPHQ drops to $B_{GSBD-FPHQ}^1 \approx 0$, meaning it becomes less important relative to other regional crisis management hubs. In contrast, the Incheon City Fire Safety Management Department (IC-FSMD) and Chungcheongnam-Do Fire Safety Office (CCND-FSO) have lower betweenness before cascades, $B_{IC-FSMD}^1$ and $B_{IC-FSMD}^1 \approx 0.1$, and higher betweenness during cascades, $B_{IC-FSMD}^1 \approx 0.2$ and $B_{IC-FSMD}^1 \approx 0.18$. This is because both regions are where customers are more affected by cascading losses. JND-FSHQ remains equally important before and during cascades.

This shift in importance among organizations demonstrates a shift in the physical geography of failures as they cascade across South Korea. GSBD-FPHQ is located in the Southeastern region of Korea where GGD-FDHQ, IC-FSMD, CCND-FSO, and JND-FSHQ are all on the west coast. IC-FSMD, in particular, is one of the most Northwest regions of the country. Thus, this analysis shows that crisis coordination during cascades needs to incorporate a diversity of emergency management stakeholders on opposite ends of the country to quickly and efficiently respond to failures. Moreover, it demonstrates that failures that originate in the Southeast will likely propogate to affect people in the Northwestern Seoul, Gyeonggi-Do, and Incheon Metropolitan Area.

| | Un-Weighted (B_v) | | | Pre-Failure (B_v^1) | | | Failure (B_v^1) | | | Iteration 1 (B_v^1) | | | Iteration 2 (B_v^1) | | | Iteration 3 (B_v^1) | | |
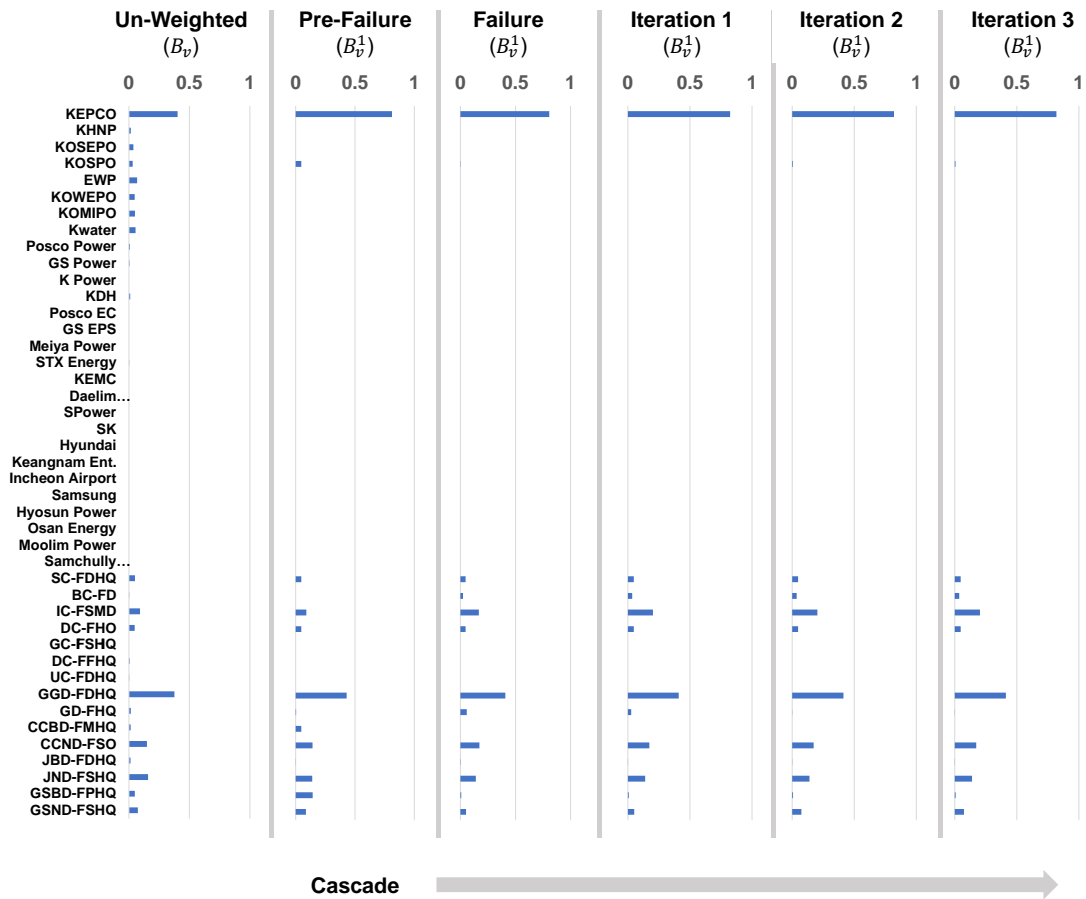
Figure 53. Critical Korean Organizations for Crisis Coordination during Cascading Failures. We measure the normalized betweenness, B_v, and normalized weighted betweenness, B_v^1, for organizations in blackout management networks weighted across each N-1 cascading failure results presented in Figure 52. Cascading failures lead to shifting importance of emergency management organizations for supporting blackout response. In particular, KEPCO, the Gyeonggi-Do Fire Department Headquarters (GGD-FDHQ), the Incheon Fire Safety Management Department (IC-FSMD), Gyeongsangbuk-Do Fire Protection Headquarters (GSBD-FPHQ), Chungcheonnam-Do Fire Safety Office (CCD-FSO), and the Jeollanam-Do Fire Safety Headquarters (JND-FSHQ) are central organizations for corrective blackout response during N-1 operations.

We reveal which organizations may be able to extend operations if critical organizations identified with B_v^1 are unavailable to coordinate blackouts using the inverse betweenness B_v^2 (Figure 54). We interpret organizations with large B_v^1 in Figure 53 to have strong partnerships and are best suited to coordinate blackout management activities. However, in surprising situations when KEPCO, GGD-FDHQ, or IC-FSMD are unable to fulfill their information sharing activities (e.g., if the KEPCO system operator room loses electricity and backup power), other organizations with low values for B_v^1 will not be able extend operations and coordinate activities. Instead, organizations with a large value of B_v^2 are better suited to take up this slack and extend operations by having many partnerships.

Figure 54 reveals at least three organizations able to support surprising coordination contexts when KEPCO, GGD-FDHQ, and IC-FSMD are unavailable. For pre-failure scenarios, the Korea Water Administration (KWater) is the most central organization for information sharing and crisis coordination based on weak ties, $B_v^2 \approx 0.6$. This is due to the location of multipurpose dams throughout South Korea and many partnerships between Kwater and governmental agencies. During cascades, provincial emergency management headquarters with large B_v^1, i.e., GGD-FDHQ, GSBD-FPHQ, and JND-FSHQ, may be supported by well-connected private power producers STX Energy, $B_v^2 \approx 0.45$ and Korea East-West Power (EWP), $B_v^2 \approx 0.22$ to coordinate emergency response activities. From a blackout management perspective, Kwater, STX Energy, and EWP should not necessarily be involved in decision-making for all cascades, but will be the most effective for acting as backup coordination hubs.

5.6 Discussion

Results reveal critical organizations for cascading failure management and ways to extend coordination activities for surprising events. Organizations like KEPCO, GGD-FDHQ, IC-FSMD, GSBD-FPHQ, and JND-FSHQ will be important for many cascading failures. This role should be made clear within policies or practices for Korean blackout management. Also, Kwater, STX Energy, and EWP can support these organizations by providing parallel coordination efforts. Taken together, our model reveals that the web of complex social relationships that change while cascades occur. This web should be considered in future, more detailed studies that try and predict the total losses expected when failures cascade beyond N-1 scenarios.

176
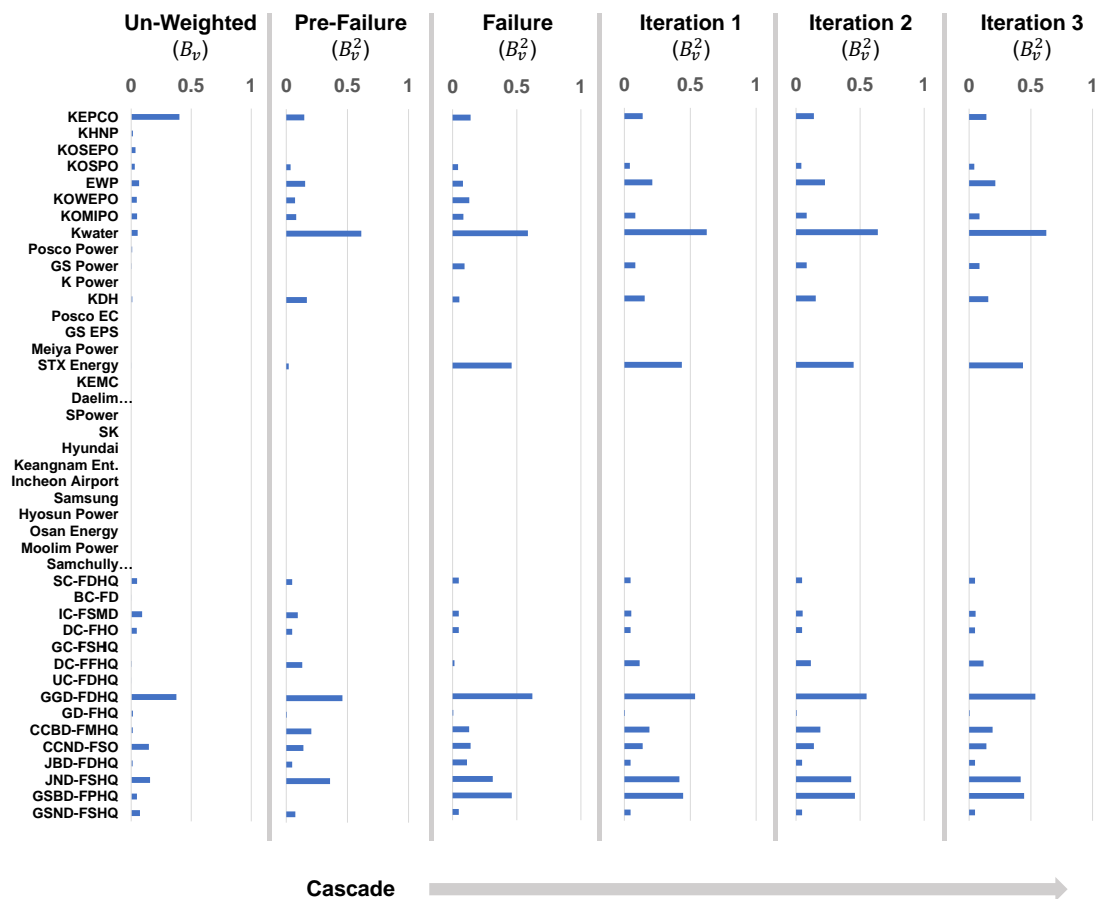
Figure 54. Critical Korean Organizations that can Serve Crisis Coordination during Coordination Failures. We measure the normalized betweenness, B_v, and the normalized inverse weighted betweenness, B_v^2 of Korean blackout management social networks to reveal unlikely organizations that are central to information sharing and coordination during cascades. Organizations with large B_v^2 may support coordination during stressed and surprising situations where central organizations identified in Figure 53 are unable to perform coordination roles. Specifically, the Korea Water Administration (Kwater), STX Energy, and Korea East-West Power (EWP) are well-suited to provide this backup role.

This study can advance national emergency management policies in South Korea to support new, important ways for organizations to manage blackouts. Several Korean infrastructure emergencies demonstrate that centralized emergency management protocols may exacerbate problems. For example, the largest Korean brownout in 2011 was exacerbated from a

contained situation outside of Seoul to a cascading loss of electricity across Seoul because of slowed coordination and information sharing activities among power organizations and regulators (Kim 2011; Eisenberg et al. 2017). Based on our results, the partnerships among specific Korean organizations make them more effective at supporting emergency management when coordination breaks down. We recommend that additional coordination roles be created among central organizations found across social network measures, e.g., creating new coordination partnerships between KEPCO and GGD-FDHQ with Kwater and STX Energy.

Our finding emphasizes the important role of managing weak ties in social or organizational networks (Granovetter 1977). While "strong" ties are generally easier to identify and define in established emergency protocols, weak ties are not typically considered in pre-disaster conditions as they are difficult to identify or many not seem relevant. Our study shows that identifying and managing the weak ties can support extensibility when strong ties are unavailable during emergencies. This new approach can also strengthen the disaster preparedness, which eventually reduces the impact from disasters and allows faster recovery. Moreover, it also increases diversity and redundancy of the functional nodes in emergency management network, improving system resilience.

Despite the potential benefits of changing policies based on our results, we acknowledge that current work is limited and requires further validation. Models and methods were validated with expert feedback regarding cascading failures and blackout management organizational networks. However, our study introduces novel simulation methods only implemented in the Korean context. Use of additional cascading failure models may provide further insight into the KPG context. Moreover, additional case studies in other social contexts would support further validation of simulation methods.

Still, we feel confident that these minor changes in blackout management protocols will lead to significant changes in the management of future blackouts based on similar changes made recently in earthquake preparedness. In 2016, earthquake response in Gyeongju (Kim et al. 2016), South Korea was exacerbated by slowed national response and information sharing to the public. The slowed response was caused by strict, hierarchical decision-making processes policies that were removed post-event to speed up coordination. The recent 2017 earthquake of similar size and velocity near Pohang (Jeong 2017), South Korea had a much

faster response due to these minor policy changes. Since our study is identifying similar ways to avoid bottlenecks in blackout management coordination, we anticipate that the small changes to national crisis management policies we recommend offer a proactive way to ensure better blackout response in the future.

References

Abdel-Aty M, Pande A, Lee C, Gayah V, Santos CD (2007) Crash risk assessment using intelligent transportation systems data and real-time intervention strategies to improve safety on freeways. *Journal of Intelligent Transportation Systems* 11:107–120, URL https://doi.org/10.1080/15472450701410395.

Adamic LA, Adar E (2003) Friends and neighbors on the Web. *Soc. Networks* 25(3):211–230, ISSN 03788733, URL http://dx.doi.org/10.1016/S0378-8733(03)00009-1.

Adger WN (2005) Social-ecological resilience to coastal disasters. *Science* 309:1036–1039, URL http://dx.doi.org/10.1126/science.1112122.

Administration NHTS (2016) Federal automated vehicles policy: Accelerating the next revolution in roadway safety (no. dot hs 812 329).

Ahn YY, Bagrow JP, Lehmann S (2010) Link communities reveal multiscale complexity in networks. *Nature* ISSN 00280836, URL http://dx.doi.org/10.1038/nature09182.

Alanis-Lobato G, Mier P, Andrade-Navarro M (2018) The latent geometry of the human protein interaction network. *Bioinformatics* 34(16):2826–2834, ISSN 1367-4803, URL http://dx.doi.org/10.1093/bioinformatics/bty206.

Alberts D (2011) *The Agility Advantage: A Survival Guide for Complex Enterprises and Endeavors* (Command Control Research Program, Washington D.C.).

Alberts D, Hayes R (2003) *Power to the Edge: Command and Control in the Information Age* (Assistant Secretary of Defense (Command Control Research Program), Washington D.C.).

Alberts D, Hayes R (2006) *Understanding Command and Control* (Command Control Research Program, Washington D.C.).

Alberts D, Huber RK, Moffat J (2010) *NATO NEC C2 Maturity Model*. ISBN 9781893723214, URL http://oai.dtic.mil/oai/oai?verb=getRecord{&}metadataPrefix=html{&}identifier=ADA555717.

Albino V, Berardi U, Dangelico RM (2015) Smart cities: Definitions, dimensions, performance, and initiatives. *Journal of Urban Technology* 22:3–21, URL https://doi.org/10.1080/10630732.2014.942092.

Aldecoa R, Orsini C, Krioukov D (2015) Hyperbolic graph generator. *Comput. Phys. Commun.* 196:492–496, ISSN 00104655, URL http://dx.doi.org/10.1016/j.cpc.2015.05.028.

Alderson D, Brown G, Carlyle W (2015) Operational models of infrastructure resilience. *Risk Analysis* 35(4):562–586.

Alderson D, Chang H, Roughan M, Uhlig S (2006) The many facets of Internet topology and traffic. *Networks and Heterogeneous Media* 1(4):569–600, ISSN 1556-1801, URL http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.158.7896.

Alderson D, Doyle J (2010) Contrasting Views of Complexity and Their Implications For Network-Centric Infrastructures. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans* 40(4):839–852.

Alderson DL, Brown GG, Carlyle WM, Cox Jr LA (2013) Sometimes there is no" most-vital" arc: Assessing and improving the operational resilience of systems. *Military Operations Research* 18(1):21–37.

Alderson DL, Bunn BB, Eisenberg DA, Howard AR, Nussbaum DA, II JT (2018) Interdependent infrastructure resilience in the u.s. virgin islands: Preliminary assessment. Technical report.

Alderson DL, Li L (2007) Diversity of graphs with highly variable connectivity. *Phys. Rev. E* 75:046102, ISSN 1063-651X, URL http://link.aps.org/doi/110.1103/PhysRevE.75.046102.

Allen WB, Liu D, Singer S (1993) Accesibility measures of us metropolitan areas. *Transportation Research Part B: Methodological* 27:439–449.

Altman N, Carley KM, Reminga J (2017) Ora user's guide 2017. *Center for the Computational Analysis of Social and Organizational System CASOS technical report* .

Amoozadeh M, Raghuramu A, Chuah Cn, Ghosal D, Zhang HM, Rowe J, Levitt K (2015) Security vulnerabilities of connected vehicle streams and their impact on cooperative driving. *IEEE Communications Magazine* 53:126–132, URL https://doi.org/10.1109/MCOM.2015.7120028.

Anderson RM, May RM, Anderson B (1992) Infectious diseases of humans: dynamics and control Wiley Online Library.

Andersson G, Donalek P, Farmer R, Hatziargyriou N, Kamwa I, Kundur P, Martins N, Paserba J, Pourbeik P, Sanchez-Gasca J, et al. (2005) Causes of the 2003 major grid blackouts in north america and europe, and recommended means to improve system dynamic performance. *IEEE transactions on Power Systems* 20(4):1922–1928.

Arianos S, Bompard E, Carbone A, Xue F (2009) Power grid vulnerability: A complex network approach. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 19(1):013119.

Atkinson SR, Moffat J (2005) *The Agile Organization: From Informal Networks to Complex Effects and Agility* (Command and Control Research Program), ISBN 978-0-12-374445-6.

Auer A, Feese S, Lockwood S (2016) History of intelligent transportation systems. *FHWA-JPO-16-329* .

Balcan D, et al. (2009a) Multiscale mobility networks and the spatial spreading of infectious diseases. *Proc. Natl. Acad. Sci.* 106:21484–21489.

Balcan D, et al. (2009b) Seasonal transmission potential and activity peaks of the new influenza a (h1n1): a monte carlo likelihood analysis based on human mobility. *BMC Med.* 7(1).

Baldick R, Chowdhury B, Dobson I, Dong Z, Gou B, Hawkins D, Huang H, Joung M, Kirschen D, Li F, et al. (2008) Initial review of methods for cascading failure analysis in electric power transmission systems ieee pes cams task force on understanding, prediction, mitigation and restoration of cascading failures. *2008 IEEE Power and Energy Society General Meeting-Conversion and Delivery of Electrical Energy in the 21st Century*, 1–8 (IEEE).

Barabási AL, Albert R (1999) Emergence of scaling in random networks. *Science* 286(5439):509–512, ISSN 00368075, URL http://dx.doi.org/10.1126/science.286.5439.509.

Barker K, Lambert JH, Zobel CW, Tapia AH, Ramirez-Marquez JE, Albert L, Nicholson CD, Caragea C (2017) Defining resilience analytics for interdependent cyber-physical-social networks. *Sustainable and Resilient Infrastructure* 2(2):59–67, ISSN 2378-9689, URL http://dx.doi.org/10.1080/23789689.2017.1294859.

Barrat A, Barthelemy M, Pastor-Satorras R, Vespignani A (2004) The architecture of complex weighted networks. *Proc. Natl. Acad. Sci.* 101:3747–3752.

Barrett CB, Constas MA (2014) Toward a theory of resilience for international development applications. *Proc. Natl. Acad. Sci.*, 14625–14630 (111).

Bartol N, Bates B, Goertzel KM, Winograd T (2009) Measuring cyber security and information assurance. *Information Assurance Technology Analysis Center* .

Beavers S (2019) *A Method for Quantifying Manpower, Personnel, Training, and Education (MPT&E) Impacts on Command and Control (C2) Resilience in Maritime Operations Centers*. Master's thesis, Naval Postgraduate School, Monterey, CA.

Belenky P (2011) *The Value of Travel Time Savings: Departmental Guidance for Conducting Economic Evaluations, Revision 2* (Washington DC: U.S. Department of Transportation).

Ben-Naim E, Krapivsky PL (2008) Phase transition with nonthermodynamic states in reversible polymerization. *Phys. Rev. E* 77(6):061132.

Berezin Y, Bashan A, Danziger MM, Li D, Havlin S (2015) Localized attacks on spatially embedded networks with dependencies. *Scientific Reports* 5, URL http://dx.doi.org/10.1038/srep08934.

Beverly K (2010) Efficient use of highway capacity. *(FHWA-HOP-* 100:10–023, texas Transportation Institute, College Station, TX p.

Bianconi G (2017) Fluctuations in percolation of sparse complex networks. *Phys. Rev. E* 96:012302, URL http://dx.doi.org/10.1103/PhysRevE.96.012302.

Boccaletti S, Bianconi G, Criado R, del Genio CI, Gómez-Gardeñes J, Romance M, Sendía-Nadal I, Wang Z, Zanin M (2014a) The structure and dynamics of multilayer networks. *Physics Reports* 544(1):1–122, ISSN 03701573, URL http://dx.doi.org/10.1016/j.physrep.2014.07.001.

Boccaletti S, Bianconi G, Criado R, Del Genio CI, Gómez-Gardenes J, Romance M, Sendina-Nadal I, Wang Z, Zanin M (2014b) The structure and dynamics of multilayer networks. *Physics Reports* 544(1):1–122.

Boccaletti S, et al. (2014c) The structure and dynamics of multilayer networks. *Phys. Rep* 544:1–122, URL http://dx.doi.org/10.1016/j.physrep.2014.07.001.

Bocchini P, Frangopol DM, Ummenhofer T, Zinke T (2014) Resilience and sustainability of civil infrastructure: toward a unified approach. *J. Infrastruct. Syst.* 20, URL http://dx.doi.org/10.1061/(ASCE)IS.1943-555X.0000177.

Boguna M, Bonamassa I, De Domenico M, Havlin S, Krioukov D, Serrano MA (2020) Network Geometry. *arXiv Prepr. arXiv2001.03241* .

Boguñá M, Papadopoulos F, Krioukov D (2010) Sustaining the Internet with hyperbolic mapping. *Nat. Commun.* 1(6):62, ISSN 2041-1723, URL http://dx.doi.org/10.1038/ncomms1063.

Boguñá M, Pastor-Satorras R (2003) Class of correlated random networks with hidden variables. *Phys. Rev. E* 68(3):036112, ISSN 1063-651X, URL http://dx.doi.org/10.1103/PhysRevE.68.036112.

Bollobás B, Riordan O (2006) *Percolation* (Cambridge University Press).

Bompard E, Napoli R, Xue F (2009) Analysis of structural vulnerabilities in power transmission grids. *International Journal of Critical Infrastructure Protection* 2(1-2):5–12.

Bompard E, Napoli R, Xue F (2010) Extended topological approach for the assessment of structural vulnerability in transmission networks. *IET Generation, Transmission & Distribution* 4(6):716–724.

Bompard E, Wu D, Xue F (2011) Structural vulnerability of power systems: A topological approach. *Electric Power Systems Research* 81(7):1334–1340.

Borgatti SP (2005) Centrality and network flow. *Social networks* 27(1):55–71.

Brandes U (2001) A faster algorithm for betweenness centrality. *Journal of mathematical sociology* 25(2):163–177.

Brew A (1998) A latent space mapping for link prediction. *Neural Inf. Process.* 1–7.

Brummitt CD, Lee KM, Goh KI (2012) Multiplexity-facilitated cascades in networks. *Phys. Rev E* 85, URL http://dx.doi.org/10.1103/PhysRevE.85.045102.

Bruneau M, et al. (2003) A framework to quantitatively assess and enhance the seismic resilience of communities. *Earthq. Spectra* 19:733–752, URL http://dx.doi.org/10.1193/1.1623497.

Buldyrev SV, Parshani R, Paul G, Stanley HE, Havlin S (2010) Catastrophic cascade of failures in interdependent networks. *Nature* 464:1025–1028.

Burgos E, Ceva H, Hernández L, Perazzo RPJ, Devoto M, Medan D (2008) Two classes of bipartite networks: nested biological and social systems. *Phys. Rev. E* 78(4):046113.

Butts CT (2009) Revisiting the foundations of network analysis. *science* 325(5939):414–416.

Byres E, Lowe J (2004) The myths and facts behind cyber security risks for industrial control systems. *Proceedings of the VDE 2004 Congress* 6, berlin, Germany.

Calabrese F, Diao M, Lorenzo GD, Ferreira J, Ratti C (2013) Understanding individual mobility patterns from urban sensing data: A mobile phone trace example. *Transportation Research Part C: Emerging Technologies* 26:301.

Calabrese F, Lorenzo D, G L, L R (2011) Estimating origin-destination flows using mobile phone location data. *IEEE Pervasive Computing* 36–44, URL https://doi.org/10.1109/MPRV.2011.41.

Callaway DS, Newman MEJ, Strogatz SH, Watts DJ (2000) Network robustness and fragility: Percolation on random graphs. *Phys. Rev. Lett.* 85(25):5468.

Cannistraci CV, Alanis-Lobato G, Ravasi T (2013a) From link-prediction in brain connectomes and protein interactomes to the local-community-paradigm in complex networks. *Sci. Rep.* 3:1613, ISSN 20452322, URL http://dx.doi.org/10.1038/srep01613.

Cannistraci CV, Alanis-Lobato G, Ravasi T (2013b) Minimum curvilinearity to enhance topological prediction of protein interactions by network embedding. *Bioinformatics* 29(13):i199–i209, ISSN 1367-4803, URL http://dx.doi.org/10.1093/bioinformatics/btt208.

Carey M, Hendrickson C (1984) Bounds on expected performance of networks with links subject to failure. *Networks.* 14:439.

Carreras BA, Newman DE, Dobson I, Degala NS (2013) Validating opa with wecc data. *2013 46th Hawaii International Conference on System Sciences*, 2197–2204 (IEEE).

Carvalho R, et al. (2014) Resilience of natural gas networks during conflicts, crises and disruptions. *PLoS ONE* 9, URL http://dx.doi.org/10.1371/journal.pone.0090265.

Cascetta E, Pagliara F, Papola A (2007) Alternative approaches to trip distribution modelling: A retrospective review and suggestions for combining different approaches. *Papers in Regional Science* 86:597.

Casey HJ (1955) Applications to traffic engineering of the law of retail gravitation. *Traffic Quaterly. IX pp* 23.

Catanzaro M, Boguñá M, Pastor-Satorras R (2005) Generation of uncorrelated random scale-free networks. *Phys. Rev. E* 71, URL http://dx.doi.org/10.1103/PhysRevE.71.027103.

Cebrowski AK, Garstka JJ (1998) Network-Centric Warfare: Its Origin and Future. *US Naval Institute Proceedings* 124(1):28–35.

Cerrudo C (2014) Hacking us traffic control systems. *Defcon* 22.

Chang SE, Nojima N (2001) Measuring post-disaster transportation system performance: the 1995 kobe earthquake in comparative perspective. *Transportation Research Part A: Policy and Practice* 35:475–494.

Chen J, Thorp JS, Dobson I (2005) Cascading dynamics and mitigation assessment in power system disturbances via a hidden failure model. *International Journal of Electrical Power & Energy Systems* 27(4):318–326.

Chen T, Kockelman K, Khan M (2013) Locating electric vehicle charging stations: Parking-based assignment method for seattle, washington. *Transportation Research Record: Journal of the Transportation Research Board* 2385:28–36, URL https://doi.org/10.3141/2385-04.

Chen W, D'Souza RM (2011) Explosive percolation with multiple giant components. *Phys. Rev. Lett.* 106(11):115701.

Chester MV, Allenby B (2019) Toward adaptive infrastructure: flexibility and agility in a non-stationarity age. *Sustainable and Resilient Infrastructure* 4(4):173–191.

Cho E, Myers SA, Leskovec J (2011) Friendship and mobility: user movement in location-based social networks. *Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (ACM Press, San Diego)* 1082.

Cho J, Gordon P, Moore II, J E P, Q P, J R (2015) Transniemo: economic impact analysis using a model of consistent inter-regional economic and network equilibria. *Transportation Planning and Technology* 38:483–502, URL https://doi.org/10.1080/03081060.2015.1039230.

Cho S, Elhourani T, Ramasubramanian S (2012) Independent directed acyclic graphs for resilient multipath routing. *IEEE/ACM Trans. Netw.* 20:153–162, URL http://dx.doi.org/10.1109/TNET.2011.2161329.

Choi J, Naderpajouh N, Yu DJ, Hastak M (2019) Capacity building for an infrastructure system in case of disaster using the system's associated social and technical components. *Journal of Management in Engineering* 35(4):04019013.

Cimellaro GP, Reinhorn AM, Bruneau M (2010) Framework for analytical quantification of disaster resilience. *Eng. Struct.* 32:3639–3649, URL http://dx.doi.org/10.1016/j.engstruct.2010.08.008.

Cohen R, Havlin S (2010) *Complex networks* (Cambridge University Press), ISBN 9780511780356, URL http://dx.doi.org/10.1017/CBO9780511780356.

Cohick DS (2018) *A First-Principles Approach to Measuring Robustness in Command and Control Systems*. Master's thesis, Naval Postgraduate School, Monterey, CA.

Çolak S, Lima A, González MC (2016) Understanding congested travel in urban areas. *Nature communications* 7(1):1–8.

Colizza B V, A B, M AJ Valleron, Vespignani A (2007) Modeling the worldwide spread of pandemic influenza: baseline case and containment interventions. *PLoS Med 4* e13.

Colizza V, Vespignani A (2007) Invasion threshold in heterogeneous metapopulation networks. *Phys. Rev. Lett.* 99:148701.

Colizza V, Vespignani A (2008) Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *J. Theor. Biol.* 251:450–467.

Collier ZA, Panwar M, Ganin AA, Kott A, Linkov I (2016) Security metrics in industrial control systems. M K, Colbert E, eds., *J*, 167–185 (Cham: Cyber-Security of SCADA and Other Industrial Control Systems. Springer International Publishing).

Committee IS (2015) Presidential policy directive 21 implementation 21.

Como G, Savla K, Acemoglu D, Dahleh MA, Frazzoli E (2013) Robust distributed routing in dynamical networks - part ii: strong resilience, equilibrium selection and cascaded failures. *IEEE Trans. Autom. Control* 58:333–348, URL http://dx.doi.org/10.1109/TAC. 2012.2209975.

Coogan S, Arcak M (2015) A compartmental model for traffic networks and its dynamical behavior. *IEEE Transactions on Automatic Control* 60:2698–2703, URL https://doi.org/ 10.1109/TAC.2015.2411916.

Cooper BS, Pitman RJ, Edmunds WJ, Gay NJ (2006) Delaying the international spread of pandemic influenza. *PLoS Med 3* e212.

Cornell Department of Mathematics (2004) English letter frequency (based on a sample of 40,000 words), https://www.math.cornell.edu/~mec/2003-2004/cryptography/subs/ frequencies.html.

Corominas-Murtra B, Goni J, Sole RV, Rodriguez-Caso C (2013) On the origins of hierarchy in complex networks. *Proc. Natl. Acad. Sci. 110*, 13316–13321.

Council NR (2012) *Disaster Resilience: A National Imperative*. URL https://doi.org/10. 17226/13457.

Cox A, Prager F, Rose A (2011) Transportation security and the role of resilience: A foundation for operational metrics. *Transport Policy* 18:307–317.

Crucitti P, Latora V, Marchiori M (2004) Model for cascading failures in complex networks. *Physical Review E* 69(4):045104.

Cuadra L, Salcedo-Sanz S, Del Ser J, Jiménez-Fernández S, Geem Z (2015) A critical review of robustness in power grids using complex networks concepts. *Energies* 8(9):9211–9265.

Daeinabi A, Rahbar AG (2013) Detection of malicious vehicles (dmv) through monitoring in vehicular ad-hoc networks. *Multimed Tools Appl* 66:325–338, URL https://doi.org/10. 1007/s11042-011-0789-y.

Daft R (2012) *Organization Theory and Design* (Nelson Education, Scarborough, Ontario, Canada).

Daganzo CF (1994) The cell transmission model: A dynamic representation of highway traffic consistent with the hydrodynamic theory. *Transportation Research Part B: Methodological* 28:269.

D'Agostino G, Scala A (2014) Networks of networks: the last frontier of complexity. *Springer* .

Daminelli S, Thomas JM, Durán C, Vittorio Cannistraci C (2015) Common neighbours and the local-community-paradigm for topological link prediction in bipartite networks. *New J. Phys.* 17(11), ISSN 13672630, URL http://dx.doi.org/10.1088/1367-2630/17/11/ 113037.

Das J, Yu H (2012) HINT: High-quality protein interactomes and their applications in understanding human disease. *BMC Syst. Biol.* 6(1):92, ISSN 1752-0509, URL http: //dx.doi.org/10.1186/1752-0509-6-92.

De Domenico M, Lancichinetti A, Arenas A, Rosvall M (2015) Identifying modular flows on multilayer networks reveals highly overlapping organization in interconnected systems. *Phys. Rev* 5, URL http://dx.doi.org/10.1103/PhysRevX.5.011027.

De Domenico M, Solé-Ribalta A, Cozzo E, Kivelä M, Moreno Y, Porter MA, Gómez S, Arenas A (2014) Mathematical formulation of multilayer networks. *Physical Review X* 3(4):1–15, ISSN 21603308, URL http://dx.doi.org/10.1103/PhysRevX.3.041022.

Degli Atti MLC, et al. (2008) Mitigation measures for pandemic influenza in italy: an individual based model considering different scenarios. *PloS One 3* e1790.

Dekker A (2008) A taxonomy of network centric warfare architectures. Technical report, Defense Science and Technology Organisation, Canberra (Australia).

D'Este GM, Zito R, Taylor MA (1999) Using gps to measure traffic system performance. *Computer-Aided Civil and Infrastructure Engineering* 14:255–265.

Diaz DO (2019) *An optimization-based approach to measuring robustness in command and control networks*. Master's thesis, Naval Postgraduate School, Monterey, CA.

Dobson I, Carreras B, Lynch V, Newman D (2001) An initial model for complex dynamics in electric power system blackouts. *hicss*, 2017 (Citeseer).

Dorogovtsev S (2010) *Lectures on complex networks* (Oxford University Press), ISBN 9780199548927.

Doyle JC, Alderson DL, Li L, Low S, Roughan M, Shalunov S, Tanaka R, Willinger W (2005) The "robust yet fragile" nature of the internet. *Proc. Nat. Acad. Sci. USA* 102(41):14497–14502.

Dunne JA, Labandeira CC, Williams RJ (2014) Highly resolved early Eocene food webs show development of modern trophic structure after the end-Cretaceous extinction. *Proc. R. Soc. B Biol. Sci.* 281(1782):20133280, URL http://dx.doi.org/10.1098/rspb.2013.3280.

Eisenberg D, Alderson D, Kitsak M, Ganin A, Linkov I (2018a) Network Foundation for Command and Control (C2) Systems: Literature Review. *IEEE Access* 6:68782–68794.

Eisenberg D, Seager T, Alderson DL (2019) Rethinking resilience analytics. *Risk Analysis* .

Eisenberg DA (2018) *How to think about resilient infrastructure systems*. Ph.D. thesis, Arizona State University.

Eisenberg DA, Alderson DL, Kitsak M, Ganin A, Linkov I (2018b) Network foundation for command and control (c2) systems: Literature review. *IEEE Access* 6:68782–68794.

Eisenberg DA, Park J, Seager TP (2017) Sociotechnical network analysis for power grid resilience in south korea. *Complexity* 2017.

Eisenberg DA, Park J, Seager TP (2020) Linking cascading failure models and organizational networks to manage large-scale blackouts in South Korea. *Journal of Management in Engineering* 36(5):04020067.

Epstein JM, et al. (2007) Controlling pandemic flu: the value of international air travel restrictions. *PloS One 2* e401.

Erdős P, Rényi A (1960) On the evolution of random graphs. *Publ. Math. Inst. Hung. Acad. Sci* 5(1):17–60.

ESRI (2012) Arcmap.

Eubank S, et al. (2004) Modelling disease outbreaks in realistic urban social networks. *Nature* 429:180–184.

Ferguson NM, et al. (2003) Planning for smallpox outbreaks. *Nature* 425:681–685.

Ferguson NM, et al. (2006) Strategies for mitigating an influenza pandemic. *Nature* 442:448–452.

Fewell M, Hazen M (2003) Network-Centric Warfare-Its Nature and Modelling. Technical report, Defence Science and Technology Organisation, Systems Sciences Lab, Salisbury, Australia, https://apps.dtic.mil/dtic/tr/fulltext/u2/a420257.pdf.

Florin MV, Linkov I, eds. (2016) *IRGC resource guide on resilience* (Lausanne, Switzerland: EPFL International Risk Governance Center (IRGC)), ISBN 978-2-9701188-0-0, URL irgc.epfl.ch.

Flynn SE, Burke SP (2012) Critical transportation infrastructure and societal resilience. *Center for National Policy* .

Fortunato S (2010) Community detection in graphs. *Phys. Rep.* 486(3-5):75–174, ISSN 03701573, URL http://dx.doi.org/10.1016/j.physrep.2009.11.002.

Forum WE (2015) Global risks. Technical report, Technical report, URL http://www3. weforum.org/docs/WEF_Global_Risks_2015_Report15.pdf.

Fosgerau M, Hjorth K, Lyk-Jensen SV (2007) *The Danish Value of Time Study (No. 5)* (Copenhagen, Denmark: The Danish Transport Research Institute).

Fosgerau M, Karlstr"om A (2010) The value of reliability. *Transportation Research Part B: Methodological* 44:38–49, URL https://doi.org/10.1016/j.trb.2009.05.002.

Fotheringham AS (1983) Some theoretical aspects of destination choice and their relevance to production-constrained gravity models. *Environment and Planning A*. 15:1121.

Fraser C, Riley S, Anderson RM, Ferguson NM (2004) Factors that make an infectious disease outbreak controllable. *Proc. Natl. Acad. Sci. U.S.A.* 101:6146–6151.

Freckleton D, Heaslip K, Louisell W, Collura J (2012) Evaluation of transportation network resiliency with consideration for disaster magnitude. *Transportation Research Board 91st Annual Meeting, Washington, DC* 91.

Fullerton T (2014) *Game Design Workshop: A Playcentric Approach to Creating Innovative Games* (CRC Press, Boca Raton, FL).

Funk S, Salathé M, Jansen VA (2010) Modelling the influence of human behaviour on the spread of infectious diseases: a review. *J. R. Soc. Interface* .

Ganin A, Massaro E, Gutfraind A, Steen N, Keisler J, Kott A, Mangoubi R, Linkov I (2016) Operational resilience: concepts, design and analysis. *Scientific Reports* 6:19540, ISSN 2045-2322, URL http://dx.doi.org/10.1038/srep19540.

Ganin AA, Kitsak M, Marchese D, Keisler JM, Seager T, Linkov I (2017) Resilience and efficiency in transportation networks. *Sci. Adv.* ISSN 23752548, URL http://dx.doi.org/ 10.1126/sciadv.1701079.

Ganin AA, Mersky AC, Jin AS, Kitsak M, Keisler JM, Linkov I (2019) Resilience in Intelligent Transportation Systems (ITS). *Transp. Res. Part C Emerg. Technol.* 100:318– 329, ISSN 0968090X, URL http://dx.doi.org/10.1016/j.trc.2019.01.014.

Gao J, Barzel B, Barabási AL (2016) Universal resilience patterns in complex networks. *Nature* 530:307–312.

Gao J, Buldyrev SV, Havlin S, Stanley HE (2011) Robustness of a network of networks. *Phys. Rev. Lett.* 107:195701.

García-Pérez G, Aliakbarisani R, Ghasemi A, Serrano MÁ (2019) Predictability of missing links in complex networks .

García-Pérez G, Boguñá M, Serrano MA (2018) Multiscale unfolding of real networks by geometric renormalization. *Nat. Phys.* 1, ISSN 1745-2473, URL http://dx.doi.org/10.1038/s41567-018-0072-5.

Gen M, Cheng R, Lin L (2008) *Network Models and Optimization: Multiobjective Genetic Algorithm Approach (Decision Engineering)* (Springer-Verlag London Limited, London).

Gerdes RM, Winstead C, Heaslip K (2013) Cps: An efficiency-motivated attack against autonomous vehicular transportation. *Proceedings of the 29th Annual Computer Security Applications Conference, ACSAC '13, ACM, New York, NY, USA, pp* 99, URL https://doi.org/10.1145/2523649.2523658.

Germann TC, Kadau K, Longini IM, Macken CA (2006) Mitigation strategies for pandemic influenza in the united states. *Proc. Natl. Acad. Sci.* 103:5935–5940.

Ghena B, Beyer W, Hillaker A, Pevarnek J, Halderman JA (2014) Green lights forever: Analyzing the security of traffic infrastructure. *8th USENIX Workshop on Offensive Technologies (WOOT 14)* .

Ghiassian SD, Menche J, Barabási AL (2015) A DIseAse MOdule Detection (DIAMOnD) Algorithm Derived from a systematic analysis of connectivity patterns of disease proteins in the human interactome. *PLoS Comput. Biol.* 11(4), ISSN 15537358, URL http://dx.doi.org/10.1371/journal.pcbi.1004120.

Girvan M, Newman MEJ (2002) Community structure in social and biological networks. *Proc. Natl. Acad. Sci. U. S. A.* 99(12):7821–6, ISSN 0027-8424, URL http://dx.doi.org/10.1073/pnas.122653799.

Goldman D (2013) *The Internet's most dangerous sites* (CNN.com).

Gomes MF, et al. (2014) Assessing the international spreading risk associated with the 2014 west african ebola outbreak. *PLOS Curr. Outbreaks* .

González MC, Hidalgo CA, Barabási AL (2008) Understanding individual human mobility patterns. *Nature.* 453:779.

Grabowski M, Roberts KH (2015) Reliability seeking virtual organizations: Challenges for high reliability organizations and resilience engineering. *Safety Science* ISSN 18791042, URL http://dx.doi.org/10.1016/j.ssci.2016.02.016.

Grad S (2009) *Engineers who hacked into L.A. traffic signal computer, jamming streets, sentenced. Los Angeles Times.* URL http://latimesblogs.latimes.com/lanow/2009/12/engineers-who-hacked-in-la-traffic-signal-computers-jamming-traffic-sentenced.html.

Grais RF, Ellis JH, Glass GE (2003) Assessing the impact of airline travel on the geographic spread of pandemic influenza. *Eur. J. Epidemiol.* 18:1065–1072.

Granovetter MS (1977) The strength of weak ties. *Social networks*, 347–367 (Elsevier).

Greenshields BD, Bibbins JR, Channing WS, Miller HH, Crum RW (1934) A study of traffic capacity. *Proceedings of the Fourteenth Annual Meeting of the Highway Research Board (Washington DC)* 1.

Gross D, Gross D, Eds (2008) *Fundamentals of queueing theory* (series in probability and statistics: Wiley, Hoboken), 4th edition.

Grover A, Leskovec J (2016) node2vec: Scalable Feature Learning for Networks. *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min. - KDD '16*, 855–864 (New York, New York, USA: ACM Press), ISBN 9781450342322, URL http://dx.doi.org/10.1145/2939672.2939754.

Gulyás A, Bíró JJ, Körösi A, Rétvári G, Krioukov D (2015) Navigable networks as Nash equilibria of navigation games. *Nat. Commun.* 6:7651, ISSN 2041-1723, URL http://dx.doi.org/10.1038/ncomms8651.

Guo H, Zheng C, Iu HHC, Fernando T (2017) A critical review of cascading failure analysis and modeling of power system. *Renewable and Sustainable Energy Reviews* 80:9–22.

Haaren RV (2012) Assessment of electric cars' range requirements and usage patterns based on driving behavior recorded in the national household travel survey of 2009. *Solar Journey USA* .

Haklay M, Weber P (2008) Openstreetmap: User-generated street maps. *IEEE Pervasive Computing* 7:12, URL https://doi.org/10.1109/MPRV.2008.80.

Hallegatte S, Rentschler J, Rozenberg J (2019) *Lifelines: the resilient infrastructure opportunity, Sustainable infrastructure series* (Washington: World Bank Group).

Ham H, Kim TJ, Boyce D (2005) Assessment of economic impacts from unexpected events with an interregional commodity flow and multimodal transportation network model. *Transportation Research Part A: Policy and Practice* 39:849–860, URL https://doi.org/10.1016/j.tra.2005.02.006.

Hamilton KE, Pryadko LP (2014) Tight lower bound for percolation threshold on an infinite graph. *Phys. Rev. Lett.* 113(20):208701.

Havlin S, Kenett DY, Bashan A, Gao J, Stanley HE (2014) Vulnerability of network of networks. *Eur. Phys. J. Spec. Top.* 223:2087–2106, URL http://dx.doi.org/10.1140/epjst/e2014-02251-6.

Henneaux P, Ciapessoni E, Cirio D, Cotilla-Sanchez E, Diao R, Dobson I, Gaikwad A, Miller S, Papic M, Pitto A, et al. (2018) Benchmarking quasi-steady state cascading outage analysis methodologies. *2018 IEEE International Conference on Probabilistic Methods Applied to Power Systems (PMAPS)*, 1–6 (IEEE).

Hines P, Cotilla-Sanchez E, Blumsack S (2010) Do topological models provide good information about electricity infrastructure vulnerability? *Chaos: An Interdisciplinary Journal of Nonlinear Science* 20(3):033122.

Hines PD, Rezaei P (2016) Cascading failures in power systems. *Smart Grid Handbook* 1–20.

Holling CS (1973) Resilience and stability of ecological systems. *Annu. Rev. Ecol. Syst* 4:1–23, URL http://dx.doi.org/10.1146/annurev.es.04.110173.000245.

Holling CS (1996) Engineering resilience versus ecological resilience. Schulze PC, ed., *Engineering within ecological constraints*, 31–44 (Washington, DC: The National Academies Press).

Hollnagel E, Woods DD, Leveson N (2006a) *Resilience engineering: Concepts and precepts* (Ashgate Publishing, Ltd.).

Hollnagel E, Woods DD, Leveson N, Eds (2006b) Resilience engineering: concepts and precepts .

Hoogendoorn R, van Arem B, Hoogendoorn S (2014) Automated driving and traffic flow efficiency and human factors: Literature review. *Transportation Research Record: Journal of the Transportation Research Board. 2422* 113–120.

Hufnagel L, Brockmann D, Geisel T (2004) Forecast and control of epidemics in a globalized world. *Proc. Natl. Acad. Sci.* 101:15124–15129, u.S.A.

Hughes JF, Healy K (2014) *Measuring the resilience of transport infrastructure* (Wellington, New Zealand: NZ Transport Agency).

Iamnitchi A, Ripeanu M, Foster I (2004) Small-world file-sharing communities. *INFOCOM 2004. Twenty-third Annu. Jt. Conf. IEEE Comput. Commun. Soc.*, volume 2, 952–963 (IEEE).

J de D Ortúzar S, Willumsen LG (2011) Modelling transport. *John Wiley & Sons, Chichester, West Sussex, United Kingdom* .

Jaccard P (1901) Étude comparative de la distribution florale dans une portion des Alpes et des Jura. *Bull. del la Société Vaudoise des Sci. Nat.* 37(JANUARY 1901):547–579.

Jansen W (2009) Directions in security metrics research. *National Institute of Standards and Technology* .

Jeong S (2017) Summary of 2017 Pohang Earthquake (2017.11. 15., ML= 5.4). 2017 Special Forum of EESK: Pohang earthquake and building damage: 2017 Nov 27. *Seoul, Korea* .

Jiang X, Adeli H (2003) Freeway work zone traffic delay and cost optimization model. *Journal of Transportation Engineering* 129:230.

Kahan JH, Allen AC, George JK (2009) An operational framework for resilience. *J. Homel. Secur. Emerg. Manag.* 6, URL http://dx.doi.org/10.2202/1547-7355.1675.

Karasz P (2011) U.k. seeks answers after biggest power failure in a decade. URL https://www.nytimes.com/2019/08/10/world/europe/uk-power-cut.html, accessed: 16 August 2019.

Kargl F, Maier J, Weber M (2001) Protecting web servers from distributed denial of service attacks. *Proceedings of the 10th international conference on World Wide Web* 514–524, URL http://dx.doi.org/10.1145/371920.372148, aCM Press.

Katz L (1953) A new status index derived from sociometric analysis. *Psychometrika* 18(1):39–43, ISSN 00333123, URL http://dx.doi.org/10.1007/BF02289026.

Kazemi SM, Poole D (2018) SimplE Embedding for Link Prediction in Knowledge Graphs. *Adv. Neural Inf. Process. Syst.* ISSN 10495258, URL http://arxiv.org/abs/1802.04868.

Keeling MJ, Rohani P (2008) Modeling infectious diseases in humans and animals. *Princeton University Press* .

Kermanshah A, Derrible S (2016) *Robustness of road systems to extreme flooding: using elements of GIS, travel demand, and network science* (Natural Hazards), URL http://dx.doi.org/10.1007/s11069-016-2678-1.

Khondaker B, Kattan L (2015) Variable speed limit: A microscopic analysis in a connected vehicle environment. *Transportation Research Part C: Emerging Technologies* 58:146, URL https://doi.org/10.1016/j.trc.2015.07.014.

Khurana H, Hadley M, Frincke DA (2010) Smart-grid security issues. *IEEE Security & Privacy Magazine* 8:81, URL https://doi.org/10.1109/MSP.2010.49.

Kim DH, Eisenberg DA, Chun YH, Park J (2017) Network topology and resilience analysis of south korean power grid. *Physica A: Statistical Mechanics and its Applications* 465:13–24.

Kim T (2011) Blackouts hit korea nationwide. URL http://www.koreatimes.co.kr/www/news/biz/2011/10/123_94847.html, accessed: 08 August 2019.

Kim Y, Rhie J, Kang TS, Kim KH, Kim M, Lee SJ (2016) The 12 september 2016 gyeongju earthquakes: 1. observation and remaining questions. *Geosciences Journal* 20(6):747–752.

King D, Shalaby A (2016) Performance metrics and analysis of transit network resilience in toronto. *Transportation Research Board 95th Annual Meeting* 95.

Kinney R, Crucitti P, Albert R, Latora V (2005) Modeling cascading failures in the north american power grid. *The European Physical Journal B-Condensed Matter and Complex Systems* 46(1):101–107.

Kitsak M, Ganin AA, Eisenberg DA, Krapivsky PL, Krioukov D, Alderson DL, Linkov I (2018) Stability of a giant connected component in a complex network. *Physical Review E* 97(1):012309.

Kitsak M, Papadopoulos F, Krioukov D (2017) Latent geometry of bipartite networks. *Phys. Rev. E* 95(3):032309, ISSN 2470-0045, URL http://dx.doi.org/10.1103/PhysRevE. 95.032309.

Kitsak M, Voitalov I (2020) HyperLink: an algorithm to embed undirected unweighted networks in two-dimensional hyperbolic space. https://bitbucket.org/dk-lab/2020_code_ hyperlink/src/master/.

Kitsak M, Voitalov I, Krioukov D (2019) Link prediction with hyperbolic geometry URL http://arxiv.org/abs/1903.08810.

Kivelä M, Arenas A, Barthelemy M, Gleeson JP, Moreno Y, Porter MA (2014) Multilayer networks. *Journal of Complex Networks* 2(3):203–271, ISSN 20511329, URL http://dx. doi.org/10.1093/comnet/cnu016.

Kleinberg J (2006) Complex networks and decentralized search algorithms. *Proc. Int. Congr. Math. Madrid, August 22–30, 2006*, 1019–1044 (Zuerich, Switzerland: European Mathematical Society Publishing House), URL http://dx.doi.org/10.4171/022-3/50.

Knight UG (2001) *Power systems in emergencies: From contingency planning to crisis management* (Wiley).

Kolisch R (1995) *Project Scheduling Under Resource Constraints: Efficient Heuristics for Several Problem Classes* (Physica-Verlag, Heidleberg).

Kovács IA, Luck K, Spirohn K, Wang Y, Pollis C, Schlabach S, Bian W, Kim DK, Kishore N, Hao T, Calderwood MA, Vidal M, Barabási AL (2019) Network-based prediction of

protein interactions. *Nat. Commun.* 10(1):1240, ISSN 2041-1723, URL http://dx.doi.org/10.1038/s41467-019-09177-y.

Krioukov D, Papadopoulos F, Kitsak M, Vahdat A, Boguñá M (2010) Hyperbolic geometry of complex networks. *Phys. Rev. E* 82(3):036106, ISSN 1539-3755, URL http://dx.doi.org/10.1103/PhysRevE.82.036106.

Kuehn R, Rogers T (2017) Heterogeneous micro-structure of percolation in sparse networks. *Europhysics Letters* 118(6):68003.

Kunegis J (2013) KONECT - The koblenz network collection. *WWW 2013 Companion - Proc. 22nd Int. Conf. World Wide Web*, ISBN 9781450320382.

Kurth M, Kozlowski W, Ganin A, Mersky A, Leung B, Dykes J, Kitsak M, Linkov I (2020) Lack of resilience in transportation networks: Economic implications. *Transportation Research Part D: Transport and Environment* 86:102419.

Lam TC, Small KA (2001) The value of time and reliability: measurement from a value pricing experiment. *Transportation Research Part E: Logistics and Transportation* .

Laney D (2001) 3d data management: Controlling data volume, velocity and variety. *META group research note* 6(70):1.

Laszka A, Potteiger B, Vorobeychik Y, Amin S, Koutsoukos X (2016) Vulnerability of transportation networks to traffic-signal tampering. *ICCPS ACM/IEEE 7th International Conference on Cyber-Physical Systems* 2016, URL https://doi.org/10.1109/ICCPS.2016.7479122.

Lee CY, Chong HY, Liao PC, Wang X (2017) Critical review of social network analysis applications in complex project management. *Journal of Management in Engineering* 34(2):04017061.

Li D (2016) Resilience of spatial networks. Yu X, Chen G, Yu W, eds., *Complex Systems and Networks* (Berlin, Heidelberg pp. 79–106: Springer Berlin Heidelberg).

Li T, Wernersson R, Hansen RB, Horn H, Mercer J, Slodkowicz G, Workman CT, Rigina O, Rapacki K, Stærfeldt HH, Brunak S, Jensen TS, Lage K (2016) A scored human

protein–protein interaction network to catalyze genomic interpretation. *Nat. Methods* 14(1):61–64, ISSN 1548-7091, URL http://dx.doi.org/10.1038/nmeth.4083.

Liben-Nowell D, Kleinberg J (2003) The Link Prediction Problem for Social Networks. *Proc. Twelfth Annu. ACM Int. Conf. Inf. Knowl. Manag.* 556–559, ISSN 1532-2882, URL http://dx.doi.org/10.1002/asi.v58:7.

Lima A, Stanojevic R, Papagiannaki D, Rodriguez P, González MC (2016) Understanding individual routing behaviour. *Journal of The Royal Society Interface* 13(20160):021.

Linkov I, Bridges T, Creutzig F, Decker J, Fox-Lent C, Kröger W, Lambert JH, Levermann A, Montreuil B, Nathwani J, et al. (2014a) Changing the resilience paradigm. *Nature Climate Change* 4:407–409.

Linkov I, Fox-Lent C, Keisler J, Della Sala S, Sieweke J (2014b) Risk and resilience lessons from venice. *Environ. Syst. Decis.* 34:378–382.

Linkov I, Palma-Oliveira JM, eds. (2017a) *Resilience and Risk*. NATO Science for Peace and Security Series C: Environmental Security (Dordrecht: Springer Netherlands), ISBN 978-94-024-1122-5 978-94-024-1123-2, URL http://link.springer.com/10.1007/978-94-024-1123-2.

Linkov I, Palma-Oliveira JM (2017b) Resilience and risk: methods and application in environment, cyber and social domains Springer, New York.

Linkov I, et al. (2013a) Measurable resilience for actionable policy. *Environ. Sci* 47:10108–10110, URL http://dx.doi.org/10.1021/es403443n.

Linkov I, et al. (2013b) Measurable resilience for actionable policy. *Environ. Sci. Technol.* 47:10108–10110.

Linkov I, et al. (2013c) Resilience metrics for cyber systems. *Environ. Syst. Decis.* 33:471–476, URL http://dx.doi.org/10.1007/s10669-013-9485-y.

Linkov I, et al. (2014c) Changing the resilience paradigm. *Nat. Clim. Change* 4:407–409.

Lipsitch M, Finelli L, Heffernan RT, Leung GM, Redd SC (2011) Improving the evidence base for decision making during a pandemic: the example of 2009 influenza a/h1n1. *Biosecurity Bioterrorism Biodefense Strategy Pract. Sci.* 9:89–115.

Lomax T, et al. (2011) Real-timing the 2010 urban mobility report. *Texas Transportation Institute, College Station TX* URL http://utcm.tamu.edu/publications/final_reports/Lomax_10-65-55.pdf).

Lombardi P, Giordano S, Farouh H, Yousef W (2012) Modelling the smart city performance. *Innovation: The European Journal of Social Science Research* 25:137, URL https://doi.org/10.1080/13511610.2012.660325.

Lu D, Yang S, Zhang J, Wang H, Li D (2017) Resilience of epidemics for sis model on networks. *Chaos Interdiscip. J. Nonlinear Sci.* 27:083105.

Lü L, Pan L, Zhou T, Zhang YC, Stanley HE (2015) Toward link predictability of complex networks. *Proc. Natl. Acad. Sci.* 112(8):2325–2330, ISSN 0027-8424, URL http://dx.doi.org/10.1073/pnas.1424644112.

Luck K, Kim DK, Lambourne L, *et al* (2019) A reference map of the human protein interactome. *bioRxiv* URL http://dx.doi.org/10.1101/605451.

Mackie PJ, Wardman M, Fowkes AS, Whelan G, Nellthorp J, Bates J (2003) *Values of Travel Time Savings UK (No. 567)* (Leeds, UK: The University of Leeds).

Mahmassani HS (2016) Autonomous vehicles and connected vehicle systems: Flow and operations considerations. *Transportation Science* 50:1140, URL https://doi.org/10.1287/trsc.2016.0712.

Mahmoud H, Masulli F, Rovetta S, Russo G (2014) Community Detection in Protein-Protein Interaction Networks Using Spectral and Graph Approaches. *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*, volume 8452 LNBI, 62–75 (Springer Verlag), ISBN 9783319090412, URL http://dx.doi.org/10.1007/978-3-319-09042-9_5.

Majdandzic A, Braunstein LA, Curme C, Vodenska I, Levy-Carciente S, Stanley HE, Havlin S (2016) Multiple tipping points and optimal repairing in interacting networks. *Nature Communications* 7:10850.

Majdandzic A, et al. (2013) Spontaneous recovery in dynamical networks. *Nat. Phys* 10:34–38, URL http://dx.doi.org/10.1038/nphys2819.

Majdandzic A, et al. (2014) Spontaneous recovery in dynamical networks. *Nat. Phys.* 10:34–38.

Makarem L, Gillet D (2012) Fluent coordination of autonomous vehicles at intersections. *IEEE International Conference on Systems, Mana, and Cybernetics (SMC)* (. IEEE, pp. 2557–2562).

Makhorin A (2014) Glpk (gnu linear programming kit), 2000. *B B* .

Marathe M, Vullikanti AKS (2013) Computational epidemiology. *Commun. ACM* 56:88–96.

Martin Wa, McGuckin Na (1998) Travel estimation techniques for urban planning. *NCHRP Report* 365.

Massaro E, Bagnoli F (2014) Epidemic spreading and risk perception in multiplex networks: a self-organized percolation method. *Phys. Rev. E* 90:052817.

Massaro E, Ganin A, Perra N, Linkov I, Vespignani A (2018) Resilience management during large-scale epidemic outbreaks. *Scientific reports* 8(1):1–9.

Mattsson LG, Jenelius E (2015) Vulnerability and resilience of transport systems – a discussion of recent research. *Transportation Research Part A: Policy and Practice* 81:16, URL https://doi.org/10.1016/j.tra.2015.06.002.

McChrystal S, Collins T, Fussell C, Silverman D (2015) *Team of Teams: New Rules of Engagement for a Complex World* (Penguin, London, United Kingdom).

McFarland DD, Brown DJ (1973) Social distance as a metric: a systematic introduction to smallest space analysis. *Bond. Plur. Form Subst. Urban Soc. Networks* 6:213–252.

Melkote S, Daskin MS (2001) An integrated model of facility location and transportation network design. *Transportation Research Part A: Policy and Practice* 35:515.

Meltzer MI, et al. (2014) Estimating the future number of cases in the ebola epidemic—liberia and sierra leone. *MMWR Surveill Summ 2014–2015* 63:1–14.

Menche J, Sharma A, Kitsak M, Ghiassian SD, Vidal M, Loscalzo J, Barabási AL (2015) Uncovering disease-disease relationships through the incomplete interactome. *Science* 347(6224):1257601, ISSN 1095-9203, URL http://dx.doi.org/10.1126/science.1257601.

Merler S, Ajelli M (2010) The role of population heterogeneity and human mobility in the spread of pandemic influenza. *Proc. R. Soc. Lond. B Biol. Sci.* 277:557–565.

Merler S, et al. (2015) Spatiotemporal spread of the 2014 outbreak of ebola virus disease in liberia and the effectiveness of non-pharmaceutical interventions: a computational modelling analysis. *Lancet Infect. Dis.* 15:204–211.

Molloy M, Reed B (1995) A critical point for random graphs with a given degree sequence. *Random structures & algorithms* 6(2-3):161–180.

Morone F, Makse HaA (2015) Influence maximization in complex networks through optimal percolation. *Nature* 524(7563):65–68, ISSN 0028-0836, URL http://dx.doi.org/10.1038/nature14604.

Motter AE, Lai YC (2002) Cascade-based attacks on complex networks. *Physical Review E* 66(6):065102.

Murray-Tuite PM (2006) A comparison of transportation network resilience under simulated system optimum and user equilibrium conditions. *Proceedings of the 2006 Winter Simulation Conference (Monterey)* 1398.

Muscoloni A, Abdelhamid I, Cannistraci CV (2018) Local-community network automata modelling based on length-three-paths for prediction of complex network structures in protein interactomes, food webs and more. *bioRxiv* 346916, URL http://dx.doi.org/10.1101/346916.

NERC (1999) Transmission capability margins and their use in atc determination. North American Electric Reliability Corporation (NERC), Accessed 31 July 2019: https://www.pjm.com//media/etools/oasis/references/tcmnercwhitepaper.ashx?la=en.

NERC (2005) Standard mod-001-0 — documentation of ttc and atc calculation methodologies introduction. North American Electric Reliability Corporation (NERC), Accessed 31 July 2019: https://www.nerc.com/files/MOD-001-0.pdf.

Newman M (2010) *Networks: An Introduction* (Oxford University Press), ISBN 0199206651, URL http://dx.doi.org/10.1093/acprof:oso/9780199206650.001.0001.

Newman MEJ, Peixoto TP (2015) Generalized communities in networks. *Phys. Rev. Lett.* 115(8):088701, ISSN 10797114, URL http://dx.doi.org/10.1103/PhysRevLett.115.088701.

Newman MEJ, Strogatz SH, Watts DJ (2001) Random graphs with arbitrary degree distributions and their applications. *Phys. Rev. E* 64(2):026118.

Nickel M, Kiela D (2018) Learning continuous hierarchies in the Lorentz model of hyperbolic geometry. *35th Int. Conf. Mach. Learn. ICML 2018*, ISBN 9781510867963.

Nussbaum L (2014) *Debian packaging tutorial*. URL https://www.debian.org/doc/manuals/packaging-tutorial/packaging-tutorial.en.pdf.

Omer M, Mostashari A, Nilchiani R (2013) Assessing resilience in a regional road-based transportation network. *International Journal of Industrial and Systems Engineering* 13:389, URL https://doi.org/10.1504/IJISE.2013.052605.

Onieva E, Milanés V, Villagrá J, Pérez J, Godoy J (2012) Genetic optimization of a vehicle fuzzy decision system for intersections. *Expert Systems with Applications* 39:13148, URL https://doi.org/10.1016/j.eswa.2012.05.087.

OpenStreetMap contributors (2017) Planet dump retrieved from https://planet.osm.org, URL https://planet.openstreetmap.org, Washington, DC.

Ouyang M (2013) Comparisons of purely topological model, betweenness based model and direct current power flow model to analyze power grid vulnerability. *Chaos: An Interdisciplinary Journal of Nonlinear Science* 23(2):023114.

Ouyang M, Due nas-Osorio L (2012) Time-dependent resilience assessment and improvement of urban infrastructure systems. *Chaos Interdiscip. J. Nonlinear Sci* 22, URL http://dx.doi.org/10.1063/1.4737204.

Ouyang M, Due nas-Osorio L, Min X (2012) A three-stage resilience analysis framework for urban infrastructure systems. *Struct. Saf.* 36:23–31, URL http://dx.doi.org/10.1016/j.strusafe.2011.12.004.

Ouyang M, Wang Z (2015) Resilience assessment of interdependent infrastructure systems: with a focus on joint restoration modeling and analysis. *Reliab. Eng. Syst. Saf* 141:74–82, URL http://dx.doi.org/10.1016/j.ress.2015.03.011.

O'Leary JG (2009) *LA Traffic: Knowledge + Malice = Chaos*.

Pant SB (2012) *Transportation network resiliency: A study of self-annealing* (Civil and Environmental Engineering Commons (available at: Utah State University), URL http://digitalcommons.usu.edu/etd/1434/).

Park J, Seager T, Rao P, Convertino M, Linkov I (2013a) Integrating risk and resilience approaches to catastrophe management in engineering systems. Risk Analysis, Society for Risk Analysis, McLean, VA, 33(3):356–367.

Park J, Seager TP, Rao PSC (2011) Lessons in risk- versus resilience-based design and management. *Integrated Environmental Assessment and Management* 7:396–399.

Park NB, Yun SJ, Jeon EC (2013b) An analysis of long-term scenarios for the transition to renewable energy in the korean electricity sector. *Energy Policy* 52:288–296.

Parshani R, Buldyrev SV, Havlin SIn (2010) reducing the coupling strength leads to a change from a first to second order percolation transition. *Phys. Rev. Lett* 105, URL http://dx.doi.org/10.1103/PhysRevLett.105.048701.

Pelling M, Özerdem A, Barakat S (2002) The macro-economic impact of disasters. *Progress in Development Studies* 2:283, URL https://doi.org/10.1191/1464993402ps042ra.

Perra N, Balcan D, Gonçalves B, Vespignani A (2011) Towards a characterization of behavior-disease models. *PloS One 6* e23084.

Petersen PF, Jóhannsson H, Nielsen AH (2015) Investigation of suitability of cascading outage assessment methods for real-time assessment. *2015 IEEE Eindhoven PowerTech*, 1–5 (IEEE).

Petit J, Shladover SE (2015) Potential cyberattacks on automated vehicles. *IEEE Transactions on Intelligent Transportation Systems* 16:546, URL https://doi.org/10.1109/TITS.2014.2342271.

Petrenj B, Lettieri E, Trucco P (2013) Information sharing and collaboration for critical infrastructure resilience - A comprehensive review on barriers and emerging capabilities. *International Journal of Critical Infrastructures* 9(4):304–329, ISSN 14753219, URL http://dx.doi.org/10.1504/IJCIS.2013.058171.

Pimm SL (1984) The complexity and stability of ecosystems. *Nature* 307:321–326, URL http://dx.doi.org/10.1038/307321a0.

Pinha D (2015) *Short-Term Resource Allocation and Management*. Doctoral dissertation, Department of Industrial and Management Systems Engineering, Statler College of Engineering and Mineral Resources at West Virginia University, Morgantown, WV.

Pinha D, Ahluwalia R, Senna P (2016) The combinatorial multi-mode resource constrained multi-project scheduling problem. *International Journal of Supply and Operations Management* 3(3):1391–1412, http://dx.doi.org/10.22034/2016.3.03.

Press TNA (2012) Disaster resilience: A national imperative .

Pritsker A, Watters L, Wolfe P (1969) Multiproject scheduling with limited resources: A zero-one programming approach. *Management Science* 16(1):93–108, https://www.jstor.org/stable/2628369.

QER Task Force (2017) Quadrennial energy review, transforming the nation's electricity system: The second installment of the qer. *DOE Office of Policy* .

Ratnasamy S, Francis P, Handley M, Karp R, Schenker S (2001) A scalable content-addressable network. *ACM SIGCOMM Comput. Commun. Rev.* 31(4):161–172, ISSN 01464833, URL http://dx.doi.org/10.1145/964723.383072.

Reed DA, Kapur KC, Christie RD (2009) Methodology for assessing the resilience of networked infrastructure. *IEEE Syst. J.* 3:174–180, URL http://dx.doi.org/10.1109/JSYST.2009.2017396.

Regional Economic Models Inc (2017) *Model Equations.* (MA).

Regional Economic Models Inc (2018) Transight 4.2 user guide, amherst, MA.

Reilly J, Sebastien M, Payer M, Bayen AM (2015) On cybersecurity of freeway control systems: Analysis of coordinated ramp metering attacks. *Transportation Research Board 94th Annual Meeting* 19, washington, DC.

Rekhter Y, Li T, Hares S, et al. (1994) A border gateway protocol 4 (bgp-4).

Renn AM (2011) The cost of congestion, the value of transit. *Urbanophile LLC* URL http://www.urbanophile.com/2011/10/06/the-cost-of-congestion-the-value-of-transit/).

Richardson T (2016) Evening snow brings traffic gridlock across washington. *The Washington Post, Washington DC* URL https://www.washingtonpost.com/news/dr-gridlock/wp/2016/01/20/evening-snow-brings-traffic-gridlock-across-washington/?utm_term=.6aada266526a).

Riley S (2007) Large-scale spatial-transmission models of infectious disease. *Science* 316:1298–1301.

Risson J, Moors T (2006) Survey of research towards robust peer-to-peer networks: Search methods. *Comput. Networks* 50(17):3485–3521, ISSN 13891286, URL http://dx.doi.org/10.1016/j.comnet.2006.02.001.

Rivers C, Lofgren E, Marathe M, Eubank S, Lewis B (2014) Modeling the Impact of Interventions on an Epidemic of Ebola in Sierra Leone and Liberia. *PLOS Curr. Outbreaks* .

Rodrigues VS, Stantchev D, Potter A, Naim M, Whiteing A (2008) Establishing a transport operation focused uncertainty model for the supply chain. *International Journal of Physical Distribution & Logistics Management* 38:388–411.

Rogers T (2015) Assessing node risk and vulnerability in epidemics on networks. *Europhysics Letters* 109(2):28005.

Rose A (2017) *Defining and measuring economic resilience from a societal, environmental and security perspective, Integrated disaster risk management* (Singapore: Springer).

Rose A, Dormady N (2018) Advances in analyzing and measuring dynamic economic resilience. Trump BD, Florin MV, Linkov I, eds., *IRGC resource guide on resilience (vol*

2): Domains of resilience for complex interconnected systems (Lausanne, CH: EPFL
International Risk Governance Center).

Sabyasachee M, Wang Y, Zhu X, Moeckel R, Mahapatra S (2013) Comparison between
gravity and destination choice models for trip distribution in maryland. *TRB 92nd Annual
Meeting Compendium of Papers (Washington)* 22.

Sami J, Pascal F, Younes B (2013) Public road transport efficiency: A stochastic frontier
analysis. *Journal of Transportation Systems Engineering and Information Technology*
13:64–71.

Samuelsson A, Tilanus B (1997) A framework efficiency model for goods transporta-
tion, with an application to regional less-than-truckload distribution. *Transport Logistics*
1:139–151.

Santos Bernardes E, Hanna MD (2009) A theoretical review of flexibility, agility and respon-
siveness in the operations management literature. *International Journal of Operations &
Production Management* 29(1):30–53, ISSN 0144-3577, URL http://dx.doi.org/10.1108/
01443570910925352.

Sauer PW (1998) Alternatives for calculating transmission reliability margin (trm) in avail-
able transfer capability (atc). *Proceedings of the Thirty-First Hawaii International Con-
ference on System Sciences*, volume 3, 89–vol (IEEE).

Schrank D, Eisele B, Lomax T, Bak J (2015) *2015 Urban Mobility Scorecard* (College
Station, TX (available at: Texas A&M Transportation Institute and INRIX), URL http:
//d2dtl5nnlpfr0r.cloudfront.net/tti.tamu.edu/documents/mobility-scorecard-2015.pdf).

Seager TP, Clark SS, Eisenberg DA, Thomas JE, Hinrichs MM, Kofron R, Jensen CN,
McBurnett LR, Snell M, Alderson DL (2017a) Redesigning resilient infrastructure re-
search. *Resilience and risk*, 81–119 (Springer).

Seager TP, et al. (2017b) Redesigning resilient infrastructure research. Linkov I, Palma-
Oliveira JM, eds., *Risk and Resilience* (Springer).

Sedjelmaci H, Senouci SM, Abu-Rgheff MA (2014) An efficient and lightweight intrusion
detection mechanism for service-oriented vehicular networks. *IEEE Internet of Things
Journal* 1:570, URL https://doi.org/10.1109/JIOT.2014.2366120.

Sen C (2014) How 2 inches of snow created a traffic nightmare in atlanta. *The Atlantic, Washington DC* URL https://www.theatlantic.com/business/archive/2014/01/how-2-inches-of-snow-created-a-traffic-nightmare-in-atlanta/283434/).

Sermpezis P, Kotronis V, Dainotti A, Dimitropoulos X (2018) A survey among network operators on bgp prefix hijacking. *ACM SIGCOMM Computer Communication Review* 48(1):64–69.

Serrano MÁ, Boguñá M (2003) Topology of the world trade web. *Phys. Rev. E* 68(1):015101, ISSN 1063-651X, URL http://dx.doi.org/10.1103/PhysRevE.68.015101.

Sewell DK, Chen Y (2017) Latent Space Approaches to Community Detection in Dynamic Networks. *Bayesian Anal.* 12(2):351–377, ISSN 1936-0975, URL http://dx.doi.org/10.1214/16-BA1000.

Shaman J, Yang W, Kandula S (2014) Inference and forecast of the current west african ebola outbreak in guinea, sierra leone and liberia. *PLOS Curr. Outbreaks* .

Shapiro JF (2007) Modeling the supply chain (thomson-brooks/cole) .

Sherehiy B, Karwowski W, Layer JK (2007) A review of enterprise agility: Concepts, frameworks, and attributes. *International Journal of Industrial Ergonomics* 37(5):445–460, ISSN 01698141, URL http://dx.doi.org/10.1016/j.ergon.2007.01.007.

Shladover S, Su D, Lu XY (2012) Impacts of cooperative adaptive cruise control on freeway traffic flow. *Transportation Research Record: Journal of the Transportation Research Board* 2324:63, URL https://doi.org/10.3141/2324-08.

Simonovic S (2016) From risk management to quantitative disaster resilience – a paradigm shift. *International Journal of Safety and Security Engineering* 6:85.

Sinai MB, Partush N, Yadid S, Yahav E (2014) Exploiting social navigation. arxiv, preprint.

Skiena S (2008) *The algorithm design manual* (London: Springer-Verlag), 2nd edition, URL http://dx.doi.org/10.1007/978-1-84800-070-4.

Smulders S (1990) Control of freeway traffic flow by variable speed signs. *Transportation Research Part B: Methodological* 24:111.

SN D, MH K, CR A, Linkov I (2016) Disease epidemics: lessons for resilience in an increasingly connected world. *J. Public Health* .

Snyder C (2013) *A User's Manual to the PMBOK® Guide,* 5th ed. (John Wiley & Sons, Inc., Hoboken, New Jersey).

Sonawane AR, Weiss ST, Glass K, Sharma A (2019) Network medicine in the age of biomedical big data. URL http://dx.doi.org/10.3389/fgene.2019.00294.

Stanton NA, Baber C, Walker GH, Houghton RJ, McMaster R, Stewart R, Harris D, Jenkins D, Young MS, Salmon PM (2008) Development of a generic activities model of command and control. *Cognition, Technology and Work* 10(3):209–220, ISSN 14355558, URL http://dx.doi.org/10.1007/s10111-007-0097-5.

Stark C, Breitkreutz BJ, Reguly T, Boucher L, Breitkreutz A, Tyers M (2006) BioGRID: a general repository for interaction datasets. *Nucleic Acids Res.* 34(Database issue):D535–9, ISSN 1362-4962, URL http://dx.doi.org/10.1093/nar/gkj109.

Stauffer D, Aharony A (1994) *Introduction to percolation theory, Rev., 2nd ed.* (London; Bristol, PA: Taylor & Francis).

Sterbenz JP, Hutchison D, Çetinkaya EK, Jabbar A, Rohrer JP, Schöller M, Smith P (2010) Resilience and survivability in communication networks: Strategies, principles, and survey of disciplines. *Computer Networks* 54:1245.

Suzuki J, Hirao T, Sasaki Y, Maeda E (2003) Hierarchical directed acyclic graph kernel: methods for structured natural language data. *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics* 1:32–39, URL http://dx.doi.org/10.3115/1075096.1075101.

Tang J, Qu M, Wang M, Zhang M, Yan J, Mei Q (2015) LINE: Large-scale information network embedding. *WWW 2015 - Proc. 24th Int. Conf. World Wide Web*, ISBN 9781450334693, URL http://dx.doi.org/10.1145/2736277.2741093.

Tatano H, Tsuchiya S (2008) A framework for economic loss estimation due to seismic transportation network disruption: a spatial computable general equilibrium approach. *Nat Hazards* 44:253, URL https://doi.org/10.1007/s11069-007-9151-0.

Team ICSCER (2012) *ICE-CERT Monitor* (Washington, DC: U.S. Department of Homeland Security).

Team WER (2014) Ebola virus disease in west africa—the first 9 months of the epidemic and forward projections. *N Engl J Med* 1481–1495.

The White House (2009) Presidential proclamation for national preparedness month.

The White House (2013) *Presidential Policy Directive 21 - Critical Infrastructure Security and Resilience* (U.S. Office of the Press Secretary).

Thiagarajan A, et al. (2009) Vtrack: accurate energy-aware road traffic delay estimation using mobile phones. Berkeley CAp, ed., *Proceedings of the 7th ACM Conference on Embedded Networked Sensor Systems* (ACM Press).

Thomas JE, Eisenberg DA, Seager TP, Fisher E (2019) A resilience engineering approach to integrating human and socio-technical system capacities and processes for national infrastructure resilience. *Journal of Homeland Security and Emergency Management* 16(2).

Thulasiraman K (1992) Graphs: Theory and algorithms. *Wiley* .

Tizzoni M, et al. (2012) Real-time numerical forecast of global epidemic spreading: case study of 2009 a/h1n1pdm. *BMC Med.* 10:1.

Top500org (2014) Operating system family / linux, URL http://www.top500.org/statistics/details/osfam/1, technical report.

Transportation Research Board D Washington (2007) *Metropolitan travel forecasting: current practice and future direction* (Special report).

Transportation Research Board D Washington (2010) Highway capacity manual 5th edition .

Turnbull K (2016) Transportation resilience: Adaptation to climate change. *Transportation Research Board, National Academies of Sciences, Engineering, and Medicine* .

US Census Bureau (2010) 2010 census data, URL ftp://ftp2.census.gov/geo/tiger/TIGER2010DP1/Tract_2010Census_DP1.zip.

US Census Bureau (2012a) 2012 cartographic boundary file, 2010 census urban area for united states, 1:500,000, URL http://www2.census.gov/geo/tiger/TIGER2010/UA/2010/tl_2010_us_uac10.zip), Washington, DC.

US Census Bureau (2012b) Growth in urban population outpaces rest of nation, census bureau reports, URL https://www.census.gov/newsroom/releases/archives/2010_census/cb12-50.html), Washington, DC.

US Department of Homeland Security (2009) National infrastructure protection plan, URL http://www.dhs.gov/xlibrary/assets/NIPP_Plan.pdf, technical report. Available at:.

US Department of Homeland Security (2015) Transportation systems sector-specific plan. Technical report, URL https://doi.org/10.1002/ejoc.201200111.

US Department of Homeland Security, Industrial Control Systems Cyber Emergency Response Team (2014) Sensys networks traffic sensor vulnerabilities (update a) (no icsa-14-247-01a) 1, washington, DC.

US Department of Transportation (2009) 2009 national household travel survey. URL http://nhts.ornl.gov.

US Department of Transportation (2016) *The Value of Travel Time Savings: Departmental Guidance for Conducting Economic Evaluations Revision 2 (2016 Update)* (Washington, DC: U.S. Department of Transportation).

US Department of Transportation (2017) *Smart City Challenge*.

Vaiman M, Bell K, Chen Y, Chowdhury B, Dobson I, Hines P, Papic M, Miller S, Zhang P (2012) Risk assessment of cascading outages: Methodologies and challenges. *IEEE Transactions on Power Systems* 27(2):631.

Van Kerkhove MD, Ferguson NM (2012) Epidemic and intervention modelling: a scientific rationale for policy decisions? lessons from the 2009 influenza pandemic. *Bull. World Health Organ.* 90:306–310.

Van Kerkhove MD, et al. (2010) Studies needed to address public health challenges of the 2009 h1n1 influenza pandemic: insights from modeling. *PLoS Med 7* e1000275.

Vanhoucke M (2012) *Project Management with Dynamic Scheduling: Baseline Scheduling, Risk Analysis and Project Control* (Springer-Verlang, Berlin Heidelberg).

Veloza OP, Santamaria F (2016) Analysis of major blackouts from 2003 to 2015: Classification of incidents and review of main causes. *The Electricity Journal* 29(7):42–49.

Verma PK, Verma R, Prakash A, Agrawal A, Naik K, Tripathi R, Alsabaan M, Khalifa T, Abdelkader T, Abogharaf A (2016) Machine-to-machine (m2m) communications: A survey. *Journal of Network and Computer Applications* 66:83, URL https://doi.org/10.1016/j.jnca.2016.02.016.

Vespignani A (2010) Complex networks: the fragility of interdependency. *Nature* 464:984–985, URL http://dx.doi.org/10.1038/464984a.

Vijayakumar P, Azees M, Kannan A, Deborah LJ (2016) Dual authentication and key management techniques for secure data transmission in vehicular ad hoc networks. *IEEE Transactions on Intelligent Transportation Systems* 17:1015, URL https://doi.org/10.1109/TITS.2015.2492981.

Voitalov I, Aldecoa R, Wang L, Krioukov D (2017) Geohyperbolic Routing and Addressing Schemes. *ACM SIGCOMM Comput. Commun. Rev.* 47(3):11–18, URL http://dx.doi.org/10.1145/3138808.3138811.

Voronoi G (1908) Nouvelles applications des paramètres continus à la théorie des formes quadratiques. *Deuxième mémoire. Recherches sur les parallélloèdres primitifs. Journal für die reine und angewandte Mathematik (Crelle's Journal)* 1908:198, URL https://doi.org/10.1515/crll.1908.134.198.

Vugrin ED, Warren DE, Ehlen MA, Camphouse CR (2010) in sustainable and resilient critical infrastructure systems simulation, modeling, and intelligent engineering (eds. 77–116 (Springer).

Walker GH, Stanton Na, Salmon PM, Jenkins DP (2008) A review of sociotechnical systems theory: a classic concept for new command and control paradigms. *Theoretical Issues in Ergonomics Science* 9(6):479–499, ISSN 1463-922X, URL http://dx.doi.org/10.1080/14639220701635470.

Walker GH, Stanton NA, Salmon PM, Jenkins DP (2012) *Command and Control: The sociotechnical perspective* (Ashgate).

Wauters T, Kinable J, Smet P, Vancroonenburg W, Vanden Berghe G, Verstichel J (2016) The multi-mode resource-constrained multi-project scheduling problem: The mista 2013 challenge. *Journal of Scheduling* 19(3):271–283, https://doi.org/10.1007/s10951-014-0402-0.

Whaiduzzaman M, Sookhak M, Gani A, Buyya R (2014) A survey on vehicular cloud computing. *Journal of Network and Computer Applications* 40:325, URL https://doi.org/10.1016/j.jnca.2013.08.004.

Wilf HS (2006) *Generating functionology, 3rd ed. ed* (Wellesley, Mass: A K Peters).

Willinger W, Govindan R, Jamin S, Paxson V, Shenker S (2002) Scaling phenomena in the Internet: Critically examining criticality. *Proceedings of the National Academy of Sciences* 99(Supplement 1):2573–2580, ISSN 0027-8424, URL http://dx.doi.org/10.1073/pnas.012583099.

Woodard D, Nogin G, Koch P, Racz D, Goldszmidt M, Horvitz E (2017) Predicting travel time reliability using mobile phone gps data. *Transportation Research Part C: Emerging Technologies* 75:30, URL https://doi.org/10.1016/j.trc.2016.10.011.

Woods DD (2015) Four concepts for resilience and the implications for the future of resilience engineering. *Reliability Engineering & System Safety* 141:5–9.

Woods DD (2018) The theory of graceful extensibility: basic rules that govern adaptive systems. *Environment Systems and Decisions* 38(4):433–457.

Woodside AG, Lysonski S, General A (1989) Model of traveler destination choice. *Journal of Travel Research* 27:8.

World Health Organization (2009) Who guide to identifying the economic consequences of disease and injury.

Worton KE (2012) Using socio-technical and resilience frameworks to anticipate threat. *2012 Workshop on Socio-Technical Aspects in Security and Trust* 2012, URL https://doi.org/10.1109/STAST.2012.16.

Wu JT, Riley S, Fraser C, Leung GM (2006) Reducing the impact of the next influenza pandemic using household-based public health interventions. *PLoS Med 3* e361.

Xiao H, Huang M, Zhu X (2015) From One Point to A Manifold: Knowledge Graph Embedding For Precise Link Prediction. *IJCAI Int. Jt. Conf. Artif. Intell.* ISSN 10450823, URL http://arxiv.org/abs/1512.04792.

Yamashita T, Izumi K, Kurumatani K (2004) Car navigation with route information sharing for improvement of traffic efficiency. *Proceedings of the 7th International IEEE Conference on Intelligent Transportation Systems (IEEE, , DC*, 465–470.

Yan G, Zhou T, Hu B, Fu ZQ, Wang BH (2006) Efficient routing on complex networks. *Physical Review E.* 73:046108.

Yan KK, Fang G, Bhardwaj N, Alexander RP, Gerstein M (2010) Comparing genomes to computer operating systems in terms of the topology and evolution of their regulatory control networks. *Proc. Natl. Acad. Sci. 107*, 9186–9191.

Yang B, tau Yih W, He X, Gao J, Deng L (2015a) Embedding entities and relations for learning and inference in knowledge bases. *3rd Int. Conf. Learn. Represent. ICLR 2015 - Conf. Track Proc.*

Yang L, Cao X, Jin D, Wang X, Meng D (2015b) A Unified Semi-Supervised Community Detection Framework Using Latent Space Graph Regularization. *IEEE Trans. Cybern.* ISSN 21682267, URL http://dx.doi.org/10.1109/TCYB.2014.2377154.

Yu H, Braun P, Yildirim MA, Lemmens I, Venkatesan K, Sahalie J, Hirozane-Kishikawa T, Gebreab F, Li N, Simonis N, Hao T, Rual JF, Dricot A, Vazquez A, Murray RR, Simon C, Tardivo L, Tam S, Svrzikapa N, Fan C, de Smet AS, Motyl A, Hudson ME, Park J, Xin X, Cusick ME, Moore T, Boone C, Snyder M, Roth FP, Barabási AL, Tavernier J, Hill DE, Vidal M, Yildirim M (2008) High-quality binary protein interaction map of the yeast interactome network. *Science* 322(5898):104–110, ISSN 1095-9203, URL http://dx.doi.org/10.1126/science.1158684.

Yu H, Gerstein M (2006) Genomic analysis of the hierarchical structure of regulatory networks. *Proc. Natl. Acad. Sci. 103*, 14724–14731.

Zetter K (2014) Hackers can mess with traffic lights to jam roads and reroute cars. *Wired*, URL https://www.wired.com/2014/04/traffic-lights-hacking/.

Zhao J, Li D, Sanhedrai H, Cohen R, Havlin S (2016) Spatio-temporal propagation of cascading overload failures in spatially embedded networks. *Nature Communications* 7:10094.

Zhou T, Lü L, Zhang YC (2009) Predicting missing links via local information. *Eur. Phys. J. B* 71(4):623–630, ISSN 14346028, URL http://dx.doi.org/10.1140/epjb/e2009-00335-8.

Zhu L, Guo D, Yin J, Steeg GV, Galstyan A (2016) Scalable temporal latent space inference for link prediction in dynamic social networks. *IEEE Trans. Knowl. Data Eng.* ISSN 10414347, URL http://dx.doi.org/10.1109/TKDE.2016.2591009.

Zimmerman RD, Murillo-Sánchez CE, Thomas RJ (2010) Matpower: Steady-state operations, planning, and analysis tools for power systems research and education. *IEEE Transactions on power systems* 26(1):12–19.

Zorrilla CD (2017) The view from puerto rico—hurricane maria and its aftermath. *New England journal of medicine* 377(19):1801–1803.

Zuev K, Boguñá M, Bianconi G, Krioukov D (2015) Emergence of soft communities from geometric preferential attachment. *Sci. Rep.* 5:9421, ISSN 2045-2322, URL http://dx.doi.org/10.1038/srep09421.

Initial Distribution List

1. Defense Technical Information Center
   Ft. Belvoir, Virginia

2. Dudley Knox Library
   Naval Postgraduate School
   Monterey, California