Proceedings of the 2021 Winter Simulation Conference
S. Kim, B. Feng, K. Smith, S. Masoud, Z. Zheng, C. Szabo, and M. Loper, eds.

A DATA PROCESSING PIPELINE FOR CYBER-PHYSICAL RISK ASSESSMENTS OF MUNICIPAL SUPPLY CHAINS

Gabriel A. Weaver

University of Illinois at Urbana-Champaign
1308 W Main Street
Urbana, Illinois 61801, USA

ABSTRACT

Smart city technologies promise reduced congestion by optimizing transportation movements. Increased connectivity, however, may increase the attack surface of a municipality's critical functions. Increased supply chain attacks (up nearly 80 % in 2019) and municipal ransomware attacks (up 60 % in 2019) motivate the need for holistic approaches to risk assessment. Therefore, we present a methodology to quantify the degree to which supply-chain movements may be observed or disrupted via compromised smart-city devices. Our data-processing pipeline uses publicly available datasets to model intermodal commodity flows within and surrounding a municipality. Using a hierarchy tree to adaptively sample spatial networks within geographic regions of interest, we bridge the gap between grid- and network-based risk assessment frameworks. Results based on fieldwork for the Jack Voltaic exercises sponsored by the Army Cyber Institute demonstrate our approach on intermodal movements through Charleston, SC and San Diego, CA.

1 INTRODUCTION

New and emerging smart city technologies promise a variety of benefits including more efficient energy usage, approaches to reduce urban congestion and noise, and optimization of the flow of people and goods. Smart roads will have the ability to divert traffic based on number of vehicles and other conditions including weather (Langlie 2019). Moreover, networked smart city IoT devices provide a platform for data collection through which further value may be derived; initiatives such as the EU Urban Data Platform by the European Commission (2021) recognize this potential. Rapid urbanization and better resource management (natural and manmade), combined with the opportunities of a rich data ecosystem, have led some to estimate a $820.7 Billion USD global smart cities market by 2025 (Vuppuluri 2020).

From a system-of-systems perspective, however, increased connectivity also increases overall system complexity. Smart city systems introduce dependencies between communications networks and functions critical to municipalities. Such dependencies can be exploited by hackers (Engstrom 2018). Activities to manage and secure these technologies must be taken for the entire lifespan of the critical infrastructure systems they support, and this can be problematic when considering the lifespan of *Internet of Things (IoT)* technologies and companies. For example, camera data on San Diego's smart streetlamps can only be decrypted by the private company that owns the platform; cameras are not readily turned off as they are tied to the same power supply as the lights (Marx 2020). The ability to respond to disruptions within an evolving natural and manmade environment depends on coordinating activities and legal authorities across a variety of stakeholders and smart city dependencies may complicate this process. As noted by Levy-Bencheton et al. (2015) in the ENISA report on smart cities cybersecurity, such collaboration and knowledge of legislation is often lacking and not clearly defined within municipalities.

Therefore, planning tools for resilient smart cities should employ holistic threat models that help municipalities consider tradeoffs between increased infrastructure complexity and operational efficiencies.

Smart city designs that integrate such threat models are necessary considering the degree to which municipalities are already affected by cyber attacks and over 68 % of humankind is expected to live in cities by 2050 (Vuppuluri 2020). Ransomware attacks against municipalities were up 60 % in 2019 (Kapersky 2019) and the average ransom grew from $30,000 in 2017 to $380,000 in 2019 (Noble 2020). These ransoms may increase as municipalities depend more heavily on the communications sector due to emerging smart city technologies as well as remote work catalyzed by the COVID-19 pandemic response.

The intent of this paper is to motivate and present a methodology to quantify the degree to which supply-chain movements may be observed or disrupted through compromised smart city IoT devices. Global supply chain movements are essential to the economic and physical well-being of communities. *Executive Order (EO)* 14017 for more resilient supply chains (Executive Office of the President 2021), combined with upcoming global security exercises focused on targeted cyber attacks on supply chain movements, underscore the importance of resilient and secure supply chains. In fact, supply chain attacks increased by nearly 80 % in 2019 versus the previous year, a trend that continued in 2020 (BI.ZONE 2021). Modern, just-in-time supply chains designed for efficiency and reduced costs by keeping inventories low can result in high-impact disruptions. The six-day blockage of the Suez Canal by the mega-containership Ever Given resulted in an estimated $400M USD of hourly losses. The importance of resilient, secure supply chains that depend upon transportation systems surrounding municipalities is further underscored by recent hacks targeting the logistics cold chain for COVID-19 vaccines (Corera 2020).

The three intended contributions of our paper are now discussed. First, we develop a metric to quantify the degree to which assets in the communications sector can observe or influence the operation of assets within the transportation sector. Second, we integrate grid- and network-based approaches to critical infrastructure risk assessment by using a hierarchy tree to define quad-tree-like grids to induce infrastructure graphs at different spatial resolutions depending upon stakeholder interest. The resultant data processing pipeline bridges the gap between the size of transportation networks in the complex systems and network optimization literature. Finally, we present use cases, based on real-world security exercises (Mitchell et al. 2021) and open-source smart city datasets from Charleston, SC and San Diego, CA, to quantify and compare the ability to observe and disrupt sensitive supply chain movements.

This paper is structured as follows: Section 2 surveys the state of the art and practice relative to our contributions. Section 3 provides background on smart city technologies along with a threat catalog for smart city IoT devices. In Section 4, a data processing pipeline to ingest CI data, translate gridded, geospatial data into induced networks, and model the network flows relative to cyber dependencies will be discussed. Section 5 demonstrates our approach with use cases from Charleston, SC and San Diego, CA. Finally Section 6 concludes.

2 RELATED WORK AND CONTRIBUTIONS

This section surveys the state of the art and practice relative to our three contributions.

2.1 Cross-Infrastructure Risk Assessment Tools

There are a number of tools in the literature and commercially available that either help assess risk from cross-infrastructure dependencies or help design more efficient smart cities. The FutureScape tool by Deloitte consulting uses agent-based simulation models to simulate infrastructure interactions in digital twin cities. Dependencies between vehicle movements and traffic lights, their dependencies upon cell phone towers, and the tower dependencies on electrical power substations are modeled. Measures of impact include traffic congestion. Both our pipeline and the FutureScape tool consider multi-sector infrastructure dependencies. Our approach, however, integrates both simulation and optimization approaches (Weaver et al. 2021) with a cyber-physical disruption model based on historically-attested events. The *All-Hazards Analysis (AHA)* tool from *Idaho National Laboratory (INL)* performs dependency analyses on publicly available infrastructure datasets. A graph-based approach is used to encode infrastructure networks and

cross-sector dependencies. Basic facility profiles show functional processes and connections that can be queried to identify implicit dependencies within the model (Mitchell et al. 2021). Finally, commercially-available tools such as Sidewalk Labs' Replica provide an approach to urban planning based on modeling and simulation for smart cities, but do not consider cross-infrastructure disruptions.

2.2 Data Fusion and Analysis for Interconnected Critical Infrastructures

Our second contribution seeks to bridge the gap between grid and network-based risk assessment methodologies. Recent work by Batista é Silva et al. (2019) recognizes the need for cross-infrastructure risk assessments and presents a set of 22 grid maps at 1 km^2 spatial resolution that estimate the economic value of various assets in the transportation, energy, industry and social sectors. Our work extends this approach, defining a hierarchy tree (Buchsbaum and Westbrook 2000) whose levels correspond to grids of different spatial resolutions over a critical infrastructure graph. Then, induced infrastructure networks can be computed via a *view* on the hierarchy tree that specifies different resolutions for different stakeholder regions of interest. The hierarchy tree provides a data structure to translate between and integrate grid-based and graph-based analyses of a critical infrastructure network. Our approach addresses a key requirement recently identified by Xing et al. (2020), who motivated the need to rethink spatial tesselations for smart city data given their dynamic and varying spatial and temporal scales and presented a dynamic tesselation approach as an initial solution based on graphs of graphs. Our approach is similar with vertices within an induced graph that correspond to subgraphs in the original infrastructure network. Within the academic literature, researchers have noted the utility of both graph theory in *System-of-Systems Engineering (SOSE)* (Harrison 2016) as well as multilayered networks for critical infrastructure analysis (Boccaletti et al. 2014). Finally, as discussed in Sections 4 and 5, our hierarchy-tree-based approach allows for reducing the regional transportation networks from those at the scale seen within recently-surveyed complex transportation systems literature (Lin and Ban 2013) to a scale more consistent with those seen within the network optimization literature (Boland et al. 2017).

2.3 Smart City Risk Assessment Frameworks

Our research is influenced by risk assessment frameworks for smart city cybersecurity as well as security exercises focused on municipalities and supply chain security. While other documents by ENISA (Levy-Bencheton et al. 2015) and NIST's Cybersecurity and Privacy Advisory Committee (CPAC) Public Working Group (2019) have discussed high-level smart city disruption categories, the intent of Section 3 is to organize attested disruptions according to IoT system components, allowing analysts to compose threats across system components. Our approach complements and builds on the methodologies described in these guides by providing a software-based tool by which municipalities can translate high-level threat categories into specific events on actual infrastructure network models based on publicly available, open-source datasets. Such an approach may be useful to provide injects for exercises such as Jack Voltaic (Mitchell et al. 2021) and Cyber Polygon (BI.ZONE 2021), respectively.

3 BACKGROUND

This section catalogs possible threats to sensors, communications networks, and databases, and actuators by which smart city platforms help municipal stakeholders understand and manage their cities.

Smart city IoT *sensor platforms* provide multimodal, real-time data streams by which municipalities can gain situational awareness for traffic monitoring, law enforcement, and weather. For example, microphones mounted on streetlamps can help law enforcement monitor urban soundscapes for gun shots and then triangulate the location of the shooter (Eakambaram 2017). Cameras in visual and non-visual spectrums can be used by law enforcement to monitor an area of responsibility, but also to understand traffic flows. Given that the data provided by sensors are inputs to downstream smart-city algorithms upon which key city functions depend, threat models associated with sensor platforms should be considered. Compromised

camera networks can provide adversaries with valuable information. For example, in 2017, hackers were able to access the DC Metropolitan Police's surveillance camera network. Through the *Remote Desktop Protocol (RDP)*, they were able to observe activity on a remote camera; approximately 123 of 187 outdoor cameras were accessed and compromised (United States Secret Service 2017). More recently, a massive compromise of more than 150,000 security cameras allowed hackers to access live feeds from Tesla, schools, prisons, and other organizations (Harwell 2021). Unauthorized use of camera systems, even by organizations within municipalities, can raise concerns; San Diego's $30M smart streetlamp project intended to use cameras for traffic and pedestrian management were subsequently used by law enforcement, sparking controversy (Perry 2020). Given advances in hardware-based edge processing for object recognition (Bekmanis 2020), access to camera networks may provide information about high-level events in addition to raw video. The supply chain of integrated circuits and embedded devices is a known supply chain concern (King et al. 2008) and patching security flaws in widely-distributed sensing hardware often requires physically replacing the hardware.

Sensor data are transmitted via communication networks and aggregated in *databases* for further processing. One concern for smart-city communication networks are jamming or spoofing of wireless signals. GPS spoofing and jamming is of concern with the potential to interfere with vehicle movements (Burgess 2019). The increased adoption of cloud-based services within industries in critical infrastructure sectors is another source of increased risk from the perspective of reinsurance market companies such as Lloyd's Underwriting Exposure Management, CyberCube, and Guy Carpenter (2021). Within the *Maritime Transportation System (MTS)*, Octopi *Terminal Operating System (TOS)* provides a cloud-hosted solution for shipping ports to manage their container yard and gate operations. Commercial cloud outages (Zuo 2021) as well as the ability to reroute potentially sensitive communications through another country (Doffman 2020) are issues that may affect future smart cities. Open-source smart city data platforms, combined with hijacked communications and tools like Shodan, could provide adversaries with real-time municipal situational awareness. In addition, data integrity attacks, such as those at the Port of Antwerp from 2011-2013 (Bateman 2013), could allow organized crime or other actors to reroute traffic and pedestrian flows. As the value of smart city data increases, the impact of ransomware attacks on municipalities may increase given the extent to which core city services will depend upon such information, including for remote work as part of pandemic response. In 2019, ransomware attacks on municipalities increased by 60% (Kapersky 2019) with local governments facing ransoms increasing from $30,000 to $380,000 in 2020 (Noble 2020).

Actuators such as smart traffic lights and digital signage have the potential to reduce congestion by adjusting stop lights and rerouting traffic dynamically based on congestion and route conditions. As shown by the SolarWinds hack, malicious software updates can compromise the software supply chain across a wide variety of stakeholders in government and private industry (McLaughlin 2021).

4 DATA FUSION PIPELINE AND MODEL

Figure 1 illustrates our process by which a regional network model for a municipality's intermodal transportation system and other critical infrastructures upon which it depends are created. This section catalogs the primary data sources used as well as our approach to reduce the size of the transportation network using a hierarchy tree. Finally, we define a measure to prioritize communication sector assets relative to commodity flows through the transportation system and quantify the degree to which a commodity flow depends on a communication network.

4.1 Primary Data Sources

Multiple data sources in Table 1 were processed to generate critical infrastructure networks and flows. Intermodal transportation systems – both road and rail – were modeled based on data provided by the USGS National Transportation Map. Scheduled movements along those modes of transport were provided by the *Surface Deployment and Distribution Command (SDDC)*. Locations of cell-phone towers were obtained

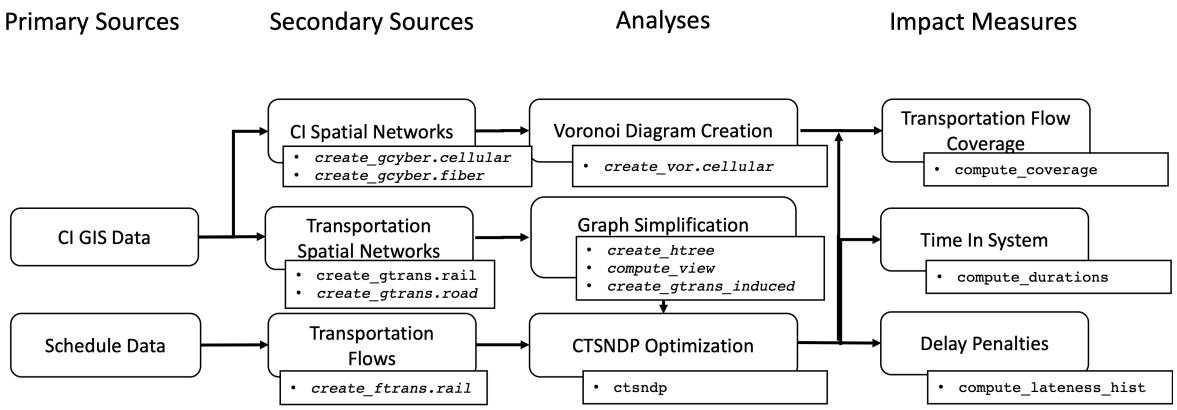


Figure 1: Workflow of the creation, calibration, and validation of the Regional MTS Models presented in this paper.

through DHS HIFLD. Finally, cross-layer dependencies that both drive efficiencies within the region and have the potential to disrupt operations were created based on the Jack Voltaic v 3.0 exercise injects and stakeholder engagement (Mitchell et al. 2021).

Table 1: Primary data sources used for the analyses presented in this section..

Data Sources						
Sector	Layer	Type	Source	Format	Resolution	ID
Transportation	Road	Network	USGS National Transportation Map	.shp	Nation	DS-TR.N-1
		Flows	SDDC Truck Movements	.ppt	Region	DS-TR.F-1
	Railway	Network	USGS National Transportation Map Federal Railroad Administration, Rail Junctions	.shp .shp	Nation Nation	DS-TR.N-3 DS-TR.N-5
		Flows	SDDC Rail Movements	.ppt	Region	DS-TR.F-3
Communication	Cellular	Network	DHS HIFLD Cell Towers	.shp	Region	DS-CM.N-3
	Cross-Layer	Network	Comms-Transportation Exercise Injects	.xls	Region	DS-CM.N-5

4.2 Secondary Data Sources

This section focuses on the representation of transportation networks to support analysis by our multicommodity network flow algorithm (Weaver et al. 2021). Additional processes at this stage of the pipeline in Figure 1 are mentioned briefly due to space. GIS files of cellular towers or wireless access points are parsed into a directed graph representation. In addition, transportation system commodity flows are instantiated from vehicle schedules.

Table 2 lists transportation network attributes and data sources used to populate *transportation networks and flows*. There are semantic, queueing, and spatial attributes associated with the transportation graph components. Semantic attributes allow for conducting analyses relative to graph component types defined within an ontology. An ontology for the transportation network defines concepts and roles that are used to specify types for vertices and edges. A more in-depth theoretical discussion of this approach, including description logics, ontologies, and graph theory, may be found in Cheh et al. (2015). Queuing attributes allow stakeholders to interpret G_{Trans} as a queueing network and, thereby, simulate the movement of vessels and containers over time. By assigning parameters for capacity, service or travel time, queue length, and queueing discipline, a queueing network can be instantiated. More details about this approach may be found in Weaver et al. (2019). Spatial attributes enable stakeholders to conduct risk assessments based on geographic regions of interest (Batista é Silva et al. 2019). Through including latitude and longitude, we

interpret and operate upon G_{Trans} as a *spatial network*, a network in which a metric is defined over the vertices (Barthélemy 2011).

4.3 Analyses

This subsection focuses on our approach to reduce the size of the transportation system graph while aggregating attributes to support our optimization algorithm. Additional analyses supported at this stage of Figure 1's pipeline, briefly mentioned due to space, include computing a Voronoi diagram for cellular towers or wireless network access points (Shamos and Hoey 1975).

4.3.1 Graph Simplification

As mentioned in Section 2, there is an apparent gap between the size of transportation systems studied within the complex networks and the transportation network optimization research. In order to make computing optimal flows on a regional network tractible for the optimizer, we need to reduce the size of the transportation network. For the Charleston, SC railway network, we needed to reduce roughly 2700 nodes to 30 nodes, or have a 100:1 node collapse ratio approximately.

Our approach to reducing the size of a critical infrastructure network builds on the grid-based approach of Silva et al.'s HARCI-EU paper (Batista é Silva, Forzieri, Herrera, Bianchi, Lavalle, and Feyen 2019) as the basis for constructing a hierarchy tree (Buchsbaum and Westbrook 2000). Given a GIS dataset encoding a network, G^D, we define a bounding box containing all points in $V[G^D]$ with sides of length b_l. This induces a grid over the bounding box, whose cells enclose a 1 km^2 area. By choosing b_l to be a power of two, we can construct a hierarchy tree T over this grid with the following properties:

- Each inner node in T represents a geographic region (bounding box) containing vertices in $V[G^D]$.
- Each leaf node in T corresponds to a vertex in $V[G^D]$.
- The level of T from root to non-leaf nodes ranges from 0 to $lg(b_l)$. This corresponds to a grid resolution ranging from 1×1 km^2 cells to $b_l \times b_l$ km^2 cells.
- At the leaf level, there is no grid.

Definition 1 A rooted tree T is a *hierarchy tree* of G if $L(T) = V(G)$, where $L(T)$ denotes the set of leaves of T. For clarity, in the remainder of this discussion, we refer to elements of $V(G)$ as *vertices* of G and to elements of $V(T)$ as *nodes* of T.

We want to be able to compute views on G^D that provide different levels of detail for different geographic regions of interest. Railway stakeholders might be more interested in the regional view of their lines whereas trucking companies might be more interested in a municipal view of the roadways. This notion of a view on a graph G is defined more formally in the following definition. Buchsbaum and Westbrook (2000) describe in more detail how, given such a view, an induced graph G can be computed, in which elements of the view $u \in U$ are vertices in G'.

Definition 2 A subset U of $V(T)$ is a *view* of G if the set $\{leaves(v) | v \in U\}$ partitions $V(G)$.

We define geographic regions of interest (e.g., port, city, and region) specifying the maximum number of vertices in $V[G^D]$ that each element of the view may span. Given a hierarchy tree T and a dictionary that maps regions of interest to upper bounds on the number of vertices per view element, we compute a view using a modified pre-order tree traversal. Figure 2 illustrates the difference in resolution at the regional and port levels for the South Carolina railway network.

4.3.2 Graph Attribute Aggregation

In order for our optimization algorithm to compute flows through the networks induced by views on the hierarchy tree, there needs to be a method to propagate network attributes in Table 2 to nodes in T that may

(a) (b)

Figure 2: Transportation network induced by a view with a higher level of detail by the Port of Charleston (Subfigure b) than the state of South Carolina as a whole (Subfigure a). The original railway GIS dataset is shown in blue while the induced graph in green.

be vertices in an induced graph G'^{D}_{Trans}. Therefore, we inductively define functions to compute queueing network parameter values for a given node $n \in V[T]$ using the rules in Table 3. We note that changing these functions may have a different effect on analysis results. These values can be populated by applying these functions in a post-order traversal of T. To be clear, $H^n \subseteq G_{Trans}$, refers to the subgraph induced by the vertices of $V[G_{Trans}]$ spanned by node $n \in V[T]$. The set $E \subseteq E[G_{Trans}]$ refers to the set of edges between two elements of the view that are consolidated into a single induced edge.

Table 2: Data attributes for intermodal transportation networks in our data processing pipeline.

Transportation Network (G_{Trans}) Attributes	
Attribute	Description
Semantic	
`rdf:type`$_i$	The type of network component $i \in V_{Trans} \cup E_{Trans}$ as defined by an critical infrastructure ontology.
Queueing	
$c_{v,e}$	Capacity, the number of entities (e.g., TEU on roadways, vessels on seaways) that can be simultaneously served at a vertex or edge.
s_v	Number of minutes to process an entity at vertex.
t_e	Travel time in minutes along an edge, computed using geodesic distance and stakeholder feedback
q_v	The maximum number of entities that can be stored while waiting for a service at a vertex.
Spatial	
lat_v, lon_v	Latitude and longitude of location $v \in V_{Trans}$.

4.3.3 Optimization and Simulation Model

Our optimization and simulation model was developed to compute the flow of commodities through a transportation network using either a discrete event queueing network simulation or a multicommodity network flow optimization algorithm. In either case, inputs to the transportation model consist of (1) a network-based representation of a transportation network, (2) vessel and commodity flows through the network, and (3) disruption, mitigation, or response events. While this paper presents results from the multicommodity network flow optimization, an in-depth discussion of the discrete event simulation model may be found in Weaver et al. (2019). For both approaches, events are specified at a time resolution of minutes. The multicommodity network flow algorithm expands the capabilities of the *Dynamic Discretization Discovery (DDD)* algorithm as a solution to the *Continuous Time Service Network Design Problem (CTSNDP)* proposed by Boland et al. (2017) and Vugrin et al. (2014). Common approaches use time-expanded networks to model movements on the network in space and time. Such discrete time approaches often have issues

Table 3: Induced transportation graph vertex and edge attributes.

Induced Network Vertex Attributes	
Attribute	Formula
$\texttt{rdf:type}_n$	`GeoRegion`
c_n	$c_n = max_{v \in V[H^n]}\{c_v\}$
s_n	$s_n = mean_{v \in V[H^n]}\{s_v\} * diameter(H^n)$
q_n	$q_v = 0$
Induced Network Edge Attributes	
$\texttt{rdf:type}_{e'}$	$\cup_{e \in E}\texttt{rdf:type}_e$
$c_{e'}$	$c_{e'} = max_{e \in E}\{c_e\}$
$t_{e'}$	$t_{e'} = mean_{e \in E}\{t_e\}$

of scalability, and so we adopt a dynamic approach to model discrete times. More details about this deterministic approach may be found in Weaver et al. (2021).

4.3.4 Graph Simplification Examples

Table 4 shows the degree to which the size of both rail networks were reduced using our hierarchy-tree based approach. A survey of real-world railway networks used in the complex transportation systems literature (Lin and Ban 2013) showed that the number of vertices in L-space representations of networks ranges from 371 to 4,853 nodes with an average of 2,512 nodes. In contrast, the number of vertices in the rail networks for the regions surrounding Charleston and San Diego were 2,732 and 4,329 vertices, respectively, close to the average and max of the networks surveyed in the literature. Applying our hierarchy-tree-based approach resulted in smaller networks consisting of 344 and 448 vertices, respectively. This is still larger than fixed graphs used as benchmarks in the transportation network optimization literature (Boland et al. 2017), which saw node sizes for 20–30 nodes and 230–700 arcs. We note, however, that the number of commodities for which results were computed were an order of magnitude smaller for this use case (1–10 commodities versus 40–100 commodities in the literature). The performance of our optimization algorithm on thousands of commodities on smaller shipping port networks is discussed in Weaver et al. (2021).

Table 4: The size of original and induced railway and roadway networks in Charleston and San Diego.

Transportation Network Sizes: Original and Induced								
	Rail		Rail-Induced		Road		Road-Induced	
	$\|V\|$	$\|E\|$	$\|V\|$	$\|E\|$	$\|V\|$	$\|E\|$	$\|V\|$	$\|E\|$
Charleston, SC	2732	2975	344	954	1245	1708	400	1196
San Diego, CA	4239	5121	448	1484	4365	6157	n/a	n/a

4.4 Impact Measures

This section presents a security metric to quantify the degree to which cyber assets can observe or influence supply chain movements within intermodal transportation networks. The intent is to provide an example of data-driven analyses to support holistic risk assessments for smart cities. The metric used for the results in this paper computes the number of communication network access points (e.g., cellular towers, Wi-Fi access points) upon which a given flow may depend. To do this, we construct a Voronoi diagram over points in a communication network and compute the Voronoi cells through which a commodity flow passes.

Definition 3 Given a finite set of points $p_i \in P$, a *Voronoi polygon (or cell)* is a convex polygon $V(i)$ with the property that p_i is the closest of the points to any $x \in V(i)$. Voronoi polygons partition the plane in a web and are called a Voronoi diagram $VOR(P)$ (Shamos and Hoey 1975).

Definition 4 A *flow* F_k for commodity $k \in K$ within network G along timebase T consists of the following components:

- An origin and earliest arrival time $(o_k, e_k) \in V[G] \times T$
- A destination and latest departure time $(d_k, l_k) \in V[G] \times T$
- An optional timed path p_k through G from (o_k, e_k) to (d_k, l_k)

In order to compute the number of communication network access points upon which a given flow F_k depends, $COV(F_k, G_{Comms})$, we construct a Voronoi Diagram over points in a communication network $VOR(G_{Comms.Cellular})$ and compute the cells $V(i)$ through which the flow passes.

Definition 5 The *communication network coverage of a transportation flow* is given by the following equation:

$$COV(F_k, G_{Comms}) = |\{V(i) | p_k \cap V(i) \neq \emptyset\}| \tag{1}$$

This measure provides a way to quantify the dependency on communication assets relative to different routes taken within a region as well as compare routes within a municipality as well as routes available in different municipalities.

5 RESULTS AND DISCUSSION

This section applies our approach to quantify the degree to which intermodal transportation flows could be surveilled or potentially disrupted by an adversary. Motivated by cyber-originating disruptions based on the Jack Voltaic 3.0 exercises sponsored by the Army Cyber Institute at West Point (Mitchell et al. 2021), we analyze the coverage of cell phone towers along baseline and alternate routes.

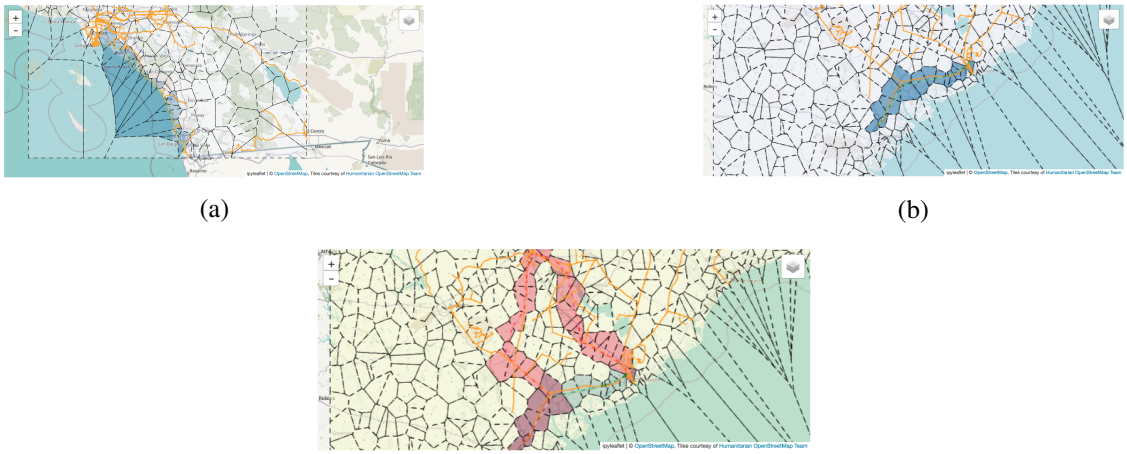
In order to evaluate communication systems' coverage relative to baseline and disrupted transportation flows, vehicle schedules, provided by the *Surface Deployment and Distribution Command (SDDC)*, were used to determine when trains left an inland distribution center and were expected to arrive at the port. SDDC was consulted in order to calibrate the choice of the route, the duration of the train route, and the speed of the train in the Charleston, SC case. The duration of the train on the route does not include staging, prepping, and loading of material. These calibration parameters were applied to the San Diego railway. Though the results from San Diego were not validated with SDDC, they provide a comparison point for evaluating the degree to which a transportation movement could be surveilled or disrupted.

5.1 Regional Rail Movements to Municipal Ports

Municipal supply chains depend upon regional rail movements to efficiently move cargo and people in an efficient manner. Railroad companies are increasingly adopting wireless communications as a cheaper and more easily maintained technology than copper wiring used in the past. For example, Norfolk Southern "uses cell phones to transmit data between field sites and central offices" (Cotey 2012). CSX Transportation has conductors and field workers communicate with each other through cell phones that include apps. These applications include reporting systems for conductors, services for track inspectors and signal maintainers, and communication with truck drivers and intermodal yard operators. Wi-Fi is also used for communication in remote locations with railroad companies building their own networks to cover regions with no cell coverage as an alternative to satellite. We compare the degree to which rail movements into two port cities – Charleston, SC and San Diego, CA – are covered by cell phone towers. Such an analysis is useful when considering the ability to intercept railroad company communication, including those during sensitive movements associated with power projection or high-value cargo.

Figure 3 illustrates the regional rail networks for the two municipalities as well as their baseline flows from an inland distribution center to port. Baseline flows were computed on induced graphs (see Figure 2, Section 4) that adaptively sampled the network for higher levels of resolution closer to the shipping ports. Cellular tower locations, provided by DHS HIFLD, were used to compute a Voronoi diagram whose cells

were filtered based on route intersection. As shown in Table 5, the Charleston and San Diego routes were comparable in terms of distance and cellular network coverage with values of 145 km, 158 km and 0.11, 0.13 towers/km, respectively.



(a) (b)

(c)

Figure 3: Regional views of primary and alternative regional rail movements (shown in green and red) from an inland distribution center shipping port via a municipality. Primary routes to San Diego, CA as well as primary and secondary routes to Charleston, SC are shown in subfigures a, b, and c, respectively. Voronoi cells highlighted in blue and red indicate cellular towers upon which communication associated with the primary and alternative rail routes depends.

Table 5: Cellular tower coverage of various regional rail routes through Charleston, SC and San Diego, CA. The baseline routes are comparable in distance and coverage. The alternate route in Charleston is much longer and less covered. No alternate route for San Diego was available..

Regional Rail Movements				
Region	Flow (F)	Distance (km)	COV$(F_k, G_{Comms.Cellular})$	$\frac{\text{COV}}{Distance}$
Charleston	Baseline	145	17	0.11
	Disrupted	408	37	0.09
San Diego	Baseline	158	21	0.13

A secondary rail route in Charleston, SC – used if the primary line became unavailable – had a significantly longer distance (408 km versus 145 km in the baseline), but similar network coverage (0.09 towers/km versus 0.11 towers/km). However, Figure 3 illustrates 29 additional cell towers upon which the secondary route depends, suggesting a potentially larger attack surface for communication to be disrupted.

6 CONCLUSION

Given the threat landscape for cyber-originating disruptions, municipalities need a holistic, system-of-systems approach to risk assessment. Surveillance and disruption models, enabled by smart cities' many interconnections, span multiple organizational and infrastructure boundaries. As a result, practitioners need tools to help integrate this domain knowledge and quantify how a supply chain attack in the communication sector could cascade to affect commodity supply chains via global, intermodal transportation systems. This paper has demonstrated a modeling approach to evaluate transportation movements relative to the degree to which they depend upon communication sector assets. Our hierarchy-tree-based data processing pipeline simultaneously allows for translating between grid- and network-based risk assessment frameworks. Use

cases in the cities of Charleston, SC and San Diego, CA demonstrate how to process openly-available data sources to model transportation flows and their dependencies on communication networks. As shown by the recent SolarWinds hack, adversary capabilities to observe critical stakeholder functions may be quickly adapted to disrupt those functions (McLaughlin 2021). Therefore, approaches to evaluate the degree to which core smart city services can be observed and subsequently disrupted, are vital to building resilient smart cities and supply chains upon which their communities depend.

ACKNOWLEDGMENTS

Research was sponsored by the U.S. Military Academy and was accomplished under Grant Number W911NF-20-1-0234, the U.S. Department of Homeland Security under Grant Number, 2015-ST-061-CIRC01 and the Dieckamp Post-Doctoral Fellowship. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the U.S. Government. The U.S. Government is authorized to reproduce and distribute reprints for Government purposes notwithstanding any copyright notation herein.

REFERENCES

Barthélemy, M. 2011. "Spatial Networks". *Physics Reports* 499(1–3):1–101.

Bateman, T. 2013. "Police Warning After Drug Traffickers' Cyber-Attack". *BBC News*. https://www.bbc.com/news/world-europe-24539417, accessed 10th April.

Batista é Silva, F., G. Forzieri, M. A. M. Herrera, A. Bianchi, C. Lavalle, and L. Feyen. 2019. "HARCI-EU, a Harmonized Gridded Dataset of Critical Infrastructures in Europe for Large-scale Risk Assessments". *Nature Scientific Data* 6(1):1–11.

Bekmanis, W. 2020. "Bringing AI at the Edge to Smart Cameras on the IoT". *QnQ Blog*. https://www.qualcomm.com/news/onq/2020/07/07/bringing-ai-edge-smart-cameras-internet-things, accessed 10th April.

BI.ZONE 2021. "Cyber Polygon – Training Description". https://cyberpolygon.com/upload/technical_training_Cyber_Polygon_2021_EN_v_1.pdf, accessed 8th April.

Boccaletti, S., G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardenes, M. Romance, I. Sendina-Nadal, Z. Wang, and M. Zanin. 2014. "The Structure and Dynamics of Multilayer Networks". *Physics Reports* 544(1):1–122.

Boland, N., M. Hewitt, L. Marshall, and M. Savelsbergh. 2017. "The Continuous-Time Service Network Design Problem". *Operations Research* 65(5):1303–1321.

Buchsbaum, A. L., and J. R. Westbrook. 2000. "Maintaining Hierarchical Graph Views". In *Proceedings of the Eleventh Annual ACM-SIAM Symposium on Discrete Algorithms (SODA)*, edited by D. B. Shmoys, 566–575. Philadelphia, Pennsylvania: Society for Industrial and Applied Mathematics.

Burgess, M. 2019. "To Protect Putin, Russia is Spoofing GPS Signals on a Massive Scale". *WIRED*. https://www.wired.co.uk/article/russia-gps-spoofing, accessed 10th April.

Cheh, C., G. A. Weaver, and W. H. Sanders. 2015. "Cyber-Physical Topology Language: Definition, Operations, and Application". In *Proceedings of the Pacific Rim International Symposium on Dependable Computing (PRDC)*, edited by G. Wang, T. Tsuchiya, and D. Xiang, 60–69. Los Alamitos, CA, USA: Institute of Electrical and Electronics Engineers, Inc.

Corera, G. 2020. "Coronavirus: Hackers Targeted COVID vVccine Supply 'Cold Chain'". *BBC News*. https://www.bbc.com/news/technology-55165552, accessed 11th April.

Cotey, A. 2012. "Railroad Communications Technology: From Cellular to Radio to Satellite to Wi-Fi". *Progressive Railroading*. https://www.progressiverailroading.com/c_s/article/Railroad-communications-technology-from-cellular-to-radio-to-satellite-to-Wi-Fi--30947, accessed 11th April.

Cybersecurity and Privacy Advisory Committee (CPAC) Public Working Group 2019. "A Risk Management Approach to Smart City Cybersecurity and Privacy". https://pages.nist.gov/GCTC/uploads/blueprints/2019_GCTC-SC3_Cybersecurity_and_Privacy_Advisory_Committee_Guidebook_July_2019.pdf, accessed 8th April.

Doffman, Z. 2020. "Russia and China 'Hijack' Your Internet Traffic: Here's What You Do". *Forbes*. https://www.forbes.com/sites/zakdoffman/2020/04/18/russia-and-china-behind-internet-hijack-risk-heres-how-to-check-youre-now-secure, accessed 10th April.

Eakambaram, M. 2017. "Smart Street Lights for Brighter Savings and Opportunities". *Intel Corporation Solution Brief*. https://www.intel.ca/content/dam/www/public/us/en/documents/solution-briefs/smart-street-lights-for-brighter-savings-solutionbrief.pdf, accessed 10th April.

Engstrom, J. 2018. "Systems Confrontation and System Destruction Warfare". Technical Report RR-1708-OSD, RAND Corporation, Santa Monica, CA.

European Commission 2021. "Urban Data Platform Plus". https://urban.jrc.ec.europa.eu/, accessed 10th April.

Executive Office of the President 2021. "Executive Order 14017, America's Supply Chains". https://www.federalregister.gov/documents/2021/03/01/2021-04280/americas-supply-chains, accessed 8th April.

Harrison, W. K. 2016. "The Role of Graph Theory in System of Systems Engineering". *IEEE Access* 4:1716–1742.

Harwell, D. 2021. "Massive Camera Hack Exposes the Growing Reach and Intimacy of American Surveillance". *The Washington Post*. https://www.washingtonpost.com/technology/2021/03/10/verkada-hack-surveillance-risk/, accessed 10th April.

Kapersky 2019. "Story of the Year 2019: Cities Under Ransomware Siege". *Securelist*. https://securelist.com/story-of-the-year-2019-cities-under-ransomware-siege/95456/, accessed 10th April.

King, S. T., J. Tucek, A. Cozzie, C. Grier, W. Jiang, and Y. Zhou. 2008. "Designing and Implementing Malicious Hardware". In *First USENIX Workshop on Large-Scale Exploits and Emergent Threats (LEET)*, edited by F. Monrose, 1–8. San Francisco, CA, USA: USENIX Association.

Langlie, K. 2019. "What Others Can Learn from Smart City Initiatives in Florida". *82 Degrees*. http://secure.thafl.com/uploads/1-29-19Whatotherscanlearn.pdf, accessed 12th April.

Levy-Bencheton, C., E. Darra, D. Bachlechner, and M. Friedewald. 2015. "Cyber security for Smart Cities: An Architecture Model for Public Transport". (TP-01-15-954-EN-N).

Lin, J., and Y. Ban. 2013. "Complex Network Topology of Transportation Systems". *Transport Reviews* 33(6):658–685.

Lloyd's Underwriting Exposure Management, CyberCube, and Guy Carpenter 2021. "Cyber risk – The Emerging Cyber Threat to Industrial Control Systems". https://www.dni.gov/files/NCSC/documents/supplychain/Lloyds_Report_-_The_Emerging_Cyber_Threat_to_Industrial_Control_Systems_Final_16022021.pdf, accessed 10th April.

Marx, J. 2020. "San Diego Can't Actually Turn Its Smart Streetlights Off". *Voice of San Diego*. https://www.voiceofsandiego.org/topics/public-safety/san-diego-cant-actually-turn-its-smart-streetlights-off/, accessed 10th April.

McLaughlin, J. 2021. "Top Biden Cyber Official: SolarWinds Breach Could Turn from Spying to Destruction 'in a Moment'". *Yahoo! News*. https://news.yahoo.com/top-biden-cyber-official-solar-winds-breach-could-turn-from-spying-to-destruction-in-a-moment-155006725.html, accessed 11th April.

Mitchell, E., D. Fletcher, E. Korn, S. Whitham, J. Hillman, R. Yearwood, C. Walker, A. Pyke, G. Weaver, B. Pugh, K. Hutton, G. Platsis, T. Klett, and R. Hruska. 2021. "Jack Voltaic 3.0 Cyber Research Report". Technical report, Army Cyber Institute at West Point. https://cyber.army.mil/Research/Jack-Voltaic/, accessed 10th April.

Noble, A. 2020. "Ransomware Attacks Demanding Larger Payouts from Local Governments". *Nextgov*. https://www.nextgov.com/cybersecurity/2020/08/ransomware-attacks-demanding-larger-payouts-local-governments/168065/, accessed 10th April.

Perry, T. S. 2020. "Cops Tap Smart Streetlights Sparking Controversy and Legislation". *IEEE Spectrum*. https://spectrum.ieee.org/view-from-the-valley/sensors/remote-sensing/cops-smart-street-lights, accessed 10th April.

Shamos, M. I., and D. Hoey. 1975. "Closest-point Problems". In *Proceedings of the 16th Annual Symposium on Foundations of Computer Science (SFCS)*, 151–162. Berkeley, California: Institute of Electrical and Electronics Engineers, Inc.

United States Secret Service 2017. "United States of America v. Mihai Alexandru Isvanca and Eveline Cismaru". http://cdn.cnn.com/cnn/2017/images/12/20/hackers.taking.over.dc.pd.cameras.affidavit.pdf, accessed 10th April.

Vugrin, E. D., M. A. Turnquist, N. J. Brown et al. 2014. "Optimal Recovery Sequencing for Enhanced Resilience and Service Restoration in Transportation Networks". *International Journal of Critical Infrastructures* 10(3/4):218–246.

Vuppuluri, P. 2020. "Investing In Innovation: The Rise of The Smart City". *Forbes*. https://www.forbes.com/sites/forbesfinancecouncil/2020/12/03/investing-in-innovation-the-rise-of-the-smart-city, accessed 13th April.

Weaver, G., M. Van Moer, and G. Salo. 2019. "Stakeholder-Centric Analyses of Simulated Shipping Port Disruptions". In *Proceedings of the 2019 Winter Simulation Conference*, edited by N. Mustafee, K.-H. G. Bae, S. Lazarova-Molnar, M. Rabe, C. Szabo, P. Haas, and Y.-J. Son, 3128–3129. Piscataway, New Jersey: Institute of Electrical and Electronics Engineers, Inc.

Weaver, G. A., B. Feddersen, L. Marla, D. Wei, A. Rose, and M. Van Moer. 2021. "Estimating Economic Losses from Cyber-Attacks on Shipping Ports: An Optimization-Based Approach". Submitted for Review. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3816659, accessed 11th April.

Xing, J., R. Sieber, and S. Roche. 2020. "Rethinking Spatial Tessellation in an Era of the Smart City". *Annals of the American Association of Geographers* 110(2):399–407.

Zuo, T. 2021. "Commercial Cloud Outages Are a Wake-Up Call". *Nextgov*. https://www.nextgov.com/ideas/2021/03/commercial-cloud-outages-are-wake-call/172731/, accessed 10th April.

AUTHOR BIOGRAPHIES

GABRIEL A. WEAVER is a Research Scientist at the Information Trust Institute at the University of Illinois at Urbana-Champaign. Weaver holds a Ph.D in Computer Science from Dartmouth College and a B.A. in Classics and Mathematics from the College of the Holy Cross. His email address is gweaver@illinois.edu.