https://dx.doi.org/10.1016/j.trc.2017.08.015

AN OPTIMIZATION APPROACH FOR DERIVING UPPER AND LOWER BOUNDS OF TRANSPORTATION NETWORK VULNERABILITY UNDER SIMULTANEOUS DISRUPTIONS OF MULTIPLE LINKS

Xiangdong Xu [a], Anthony Chen [b,a] *, Chao Yang [a]

[a] Key Laboratory of Road and Traffic Engineering of the Ministry of Education, Tongji University, Shanghai, China

[b] Department of Civil and Environmental Engineering, The Hong Kong Polytechnic University, Kowloon, Hong Kong

ABSTRACT

This paper aims to develop an optimization approach for deriving the upper and lower bounds of transportation network vulnerability under simultaneous disruptions of multiple links without the need to evaluate all possible combinations as in the enumerative approach. Mathematically, we formulate the upper and lower bounds of network vulnerability as a binary integer bi-level program (BLP). The upper-level subprogram maximizes or minimizes the remaining network throughput under a given number of disrupted links, which corresponds to the upper and lower vulnerability bounds. The lower-level subprogram checks the connectivity of each origin-destination (O-D) pair under a network disruption scenario without path enumeration. Two alternative modeling approaches are provided for the lower-level subprogram: the virtual link capacity-based maximum flow problem formulation and the virtual link cost-based shortest path problem formulation. Computationally, the BLP model can be equivalently reformulated as a single-level mixed integer linear program by making use of the optimality conditions of the lower-level subprograms and linearization techniques for the complementarity conditions and bilinear terms. Numerical examples are also provided to systematically demonstrate the validity, capability, and flexibility of the proposed optimization model. The vulnerability envelope constructed by the upper and lower bounds is able to effectively consider all possible combinations without the need to perform a full network scan, thus avoiding the combinatorial complexity of enumerating multi-disruption scenarios. Using the vulnerability envelope as a network performance assessment tool, planners and managers can more cost-effectively plan for system protection against disruptions, and prioritize system improvements to minimize disruption risks with limited resources.

Keywords: network vulnerability; vulnerability envelope; upper and lower bounds

* Corresponding author. Email addresses: xiangdongxu@tongji.edu.cn (X. Xu); anthony.chen@polyu.edu.hk (A. Chen); tongjiyc@tongji.edu.cn (C. Yang).

1 INTRODUCTION

Vulnerability is the susceptibility of a system to threats and incidents that results in operational degradation. The core of transportation network vulnerability analysis is to identify the critical/vulnerable/important components (e.g., links and nodes), whose disruptions could have a significant impact on travelers' behaviors and network performance. This topic has received a great deal of attention in the past decade (see, e.g., Berdica, 2002; Chen *et al*., 2007a,c; Murray and Grubesic, 2007; Nagurney and Qiang, 2010; Chen *et al*., 2012; Ho *et al*., 2013; Zhao *et al*., 2013; Jenelius and Mattsson, 2015; Bell *et al*., 2017). Identification of critical components in a network has many potential applications in both the pre-disaster planning and post-disaster management (e.g., targeted protection or retrofitting, strategic location of rapid response and repair stations to facilitate the network recovery and mitigation, evacuation routes planning, and evacuation network monitoring) to ensure that the critical components are adequately monitored (Murray-Tuite and Wolshon, 2013; He *et al*., 2015; Wang *et al*., 2016).

In the literature, the majority of existing methodologies for transportation network vulnerability analysis belong to the *disruption scenario enumeration approach* (via enumeration without or with pre-scanning, or random simulation schemes). At each enumerated scenario, *one link/node is removed or degraded at a time*, and the impact of each individual link/nodal removal or degradation is evaluated and ranked according to different indictors. Interested readers are directed to a comprehensive review by Mattsson and Jenelius (2015) on the evaluation methods and indictors. With the increase of the number of enumerated scenarios, this type of approach is able to consider a range of potential disruption scenarios. However, the *enumeration* approach has a combinatorial complexity especially when considering the simultaneous disruption of multiple links/nodes at the same time. Let m and n denote the total number of links and the number of simultaneously disrupted links, respectively. Then, the number of potential scenarios with n disrupted links is C_n^m. One can envision the computational burden when applying it to large-scale networks with simultaneous disruptions. Although many well connected networks could be resilient enough to a single-link (or node) failure, **simultaneous disruptions** can be very problematic, resulting in disruption propagations and widespread disruptions. On the other hand, the *simulation* scheme may miss some important scenarios (e.g., the best or the worst case) due to the limited number of samples. Also, it may miss some hidden/phantom vulnerability scenarios (see, e.g., Jenelius, 2010) that are not apparent to analysts due

to the large scale and complex network structure. These issues render an incomplete understanding of all potential disruption scenarios and their associated impacts. Recently, Wang *et al*. (2016) recognized the combinatorial complexity of considering multiple disruptions, and provided a global optimization approach to identifying critical links for multiple disruptions, without the need to perform a complete network scan. However, only the worst case situation was considered, which may not provide a useable or cost-effective strategy for managing/protecting the identified critical links.

On a different line of research, the *game theoretic approach* has been developed to assess the transportation network vulnerability, such as Bell (2000) of considering a two-player zero-sum non-cooperative game, Bell and Cassir (2002) of considering a multiplayer game, and Szeto *et al*. (2007) of considering multiple network-specific demons. In this approach, the evil entity or demon seeks to maximize the total network cost by damaging links in the network, while network users seek a route to minimize their travel costs. The critical links in the network are likely to be destroyed by the demon as a consequence of the game. Since the demon is allowed to destroy any link in the network, the game theoretic approach also only considers the worst-case scenario, and therefore provides a pessimistic evaluation of network performance.

In addition, *sensitivity and uncertainty analyses* have also been used to identify critical links that affect the system performance the most (e.g., Nicholson and Du, 1997; Chen *et al*., 2002; Luathep *et al*., 2011; Yang *et al*., 2013). A weak link with higher capacity variability may not necessarily be a critical link. Instead, a critical link must be one that is both important (i.e., substantial impact on system performance) and weak (i.e., large capacity variability). The critical links should be the prime candidates for strengthening, rather than those that are merely weak. The critical index of a link indicates the proportion of the overall uncertainty of performance measure contributed by the uncertainty of its link capacity. This approach is able to consider simultaneous link degradations. However, since sensitivity analysis is only valid locally for minor perturbations of inputs and parameters, this approach may not be applicable to large perturbations in some disruption scenarios.

When considering multiple simultaneous disruptions (e.g., n links disrupted), there may have a large number of possible scenarios corresponding to different location combinations, and each scenario has an *unknown* occurrence possibility. Hence, a way to avoid the unknown occurrence possibility of disruption scenarios is to consider the *range* of all potential disruption scenarios and their associated impacts. To the best of

our knowledge, there is no analytical approach of transportation network vulnerability with a systematic consideration and quantification of all possible simultaneous disruptions. **This study attempts to develop an optimization approach for deriving the upper and lower bounds of transportation network vulnerability –a vulnerability envelope, while circumventing the need of enumerating all possible disruption scenarios**. This is different from the conservative (or pessimistic) consideration of the worst-case scenario in Wang *et al*. (2016) or the *game theoretic approach* in Bell (2000), Bell and Cassir (2002), and Szeto *et al*. (2007). The upper and lower bounds provide the most optimistic and pessimistic quantification of network vulnerability range (i.e., the least to the most disruptive cases). The single consideration of either the least or the most disruptive cases may lead to biased (overestimated or underestimated) network performance assessment. Instead, with the vulnerability envelope, network planners and managers can more cost-effectively plan for system protection against disruptions, and prioritize system improvements to minimize disruption risks with limited resources. Particularly, those links appeared in both the upper-bound and lower-bound scenarios deserve more resources and actions to protect in the pre-disaster network planning stage. A large range between the upper and lower bounds indicates that the network is more susceptible or less resilient against disruptions, and a substantial percentage of trips can be affected by these different combinations of multiple simultaneous disruptions. On the other hand, a small range could mean that the network is very vulnerable as it could be easily disconnected (i.e., both upper and lower bounds are quite small and similar), or it is not vulnerable at all as it could be highly connected (i.e., both bounds are quite large and similar).

A unified optimization framework requires the same (or a common) objective function to derive the upper and lower bounds of network vulnerability. The proposed framework is flexible in the sense that it allows the use of different specifications of vulnerability measures (i.e., objective functions) for different modeling purposes. Examples include the remaining travel throughput of the network (a measure of network capacity) after disruptions, and the remaining route diversity (a measure of network redundancy) after disruptions. The more remaining travel throughput and/or the alternative routes are preserved after disruptions, the more robust is the network to disruptions.

Mathematically, we formulate the upper and lower bounds of transportation network vulnerability as a binary integer bi-level program (BLP). The basic decision variables are binary variables indicating whether a link is disrupted or not, and the basic equality

constraint is imposed on the total number of disrupted links to represent the level of network disruption. The larger the number of simultaneously disrupted links, the more severe is the network disruption. The upper-level subprogram maximizes or minimizes the remaining network travel throughput after disruptions under a given number of disrupted links, which corresponds to the upper and lower vulnerability bounds, respectively. The lower-level subprogram determines the O-D connectivity/accessibility under a disruption scenario from the upper-level subprogram. Two alternative modeling approaches are provided for the lower-level subprogram without path enumeration but with an explicit consideration of traveler's route acceptance criterion: the virtual link capacity-based maximum flow problem formulation and the virtual link cost-based shortest path problem formulation. If an O-D pair is connected, then the virtual maximum flow is larger than zero and the virtual minimum cost is smaller than travelers' acceptable elongation threshold (i.e., a usable path exists). To solve the proposed BLP model, we can reformulate it as a single-level mixed integer linear program (MILP) by making use of the Karush-Kuhn-Tucker (KKT) conditions of the lower-level subprogram and various linearization techniques for the complementarity conditions and bilinear terms. The desirable MILP structure makes it solvable by existing solvers in commercial software packages for its globally optimal solution.

In summary, the main contribution of this paper is the development of an optimization framework for deriving the exact upper and lower bounds of transportation network vulnerability envelope under simultaneous disruptions of multiple links, without the need to evaluate all possible combinations as in the enumerative approach for assessing network efficiency (e.g., Nagurney and Qiang, 2010). The proposed modeling framework is flexible for three reasons: (1) it allows flexible specifications of a common objective function for different modeling purposes, (2) it allows flexible modeling approaches to check the O-D connectivity without path enumeration, and (3) it allows users to specify thresholds to implicitly define route availability/usability. The vulnerability bound envelope constructed by the upper and lower bounds under the optimization framework is capable of considering all possible combinations without the need to perform a full network scan, thus avoiding the combinatorial complexity of enumerating multi-disruption scenarios.

The remainder of this paper is organized as follows. Section 2 and Section 3 present the mathematical formulation and solution algorithm of the vulnerability bounds model, respectively. Section 4 uses a set of numerical examples to demonstrate the features of

the proposed model. Finally, some concluding remarks and future research directions are provided in Section 5.

2 MATHEMATICAL FORMULATION

In this section, the BLP formulation is provided to derive the exact upper and lower bounds of network vulnerability envelope under multiple simultaneous disruptions.

2.1 Upper-Level Subprogram

Consider the following binary integer linear program:

$$\text{max or min } f_n(\mathbf{x}, \mathbf{z}) = \sum_{w \in W} q^w z^w \tag{1}$$

$$s.t. \quad \sum_{a \in A} x_a = n, \tag{2}$$

$$-M\left(1 - z^w\right) + \varepsilon \le V^w \le M z^w, \ \forall w \in W, \tag{3}$$

$$x_a = \{0, 1\}, \ \forall a \in A, \tag{4}$$

$$z^w = \{0, 1\}, \ \forall w \in W, \tag{5}$$

where A denotes the set of directed links; W is the set of O-D pairs; q^w is the travel demand between O-D pair w; M is a very large positive constant; and ε is a very small positive constant. The two types of binary *decision variable*s (**x** and **z** in the vector form) are defined as follows:

x_a=1 Link a is disrupted

z^w=1 O-D pair w is connected

The objective function in Eq. (1) maximizes or minimizes the remaining network throughput after disruptions, which corresponds to the upper and lower vulnerability bounds (i.e., the optimistic and pessimistic cases), respectively. Eq. (2) is the basic equality constraint imposed on the total number of disrupted links (n), representing the level of network disruption or the ability of attackers to destroy the network components. The larger the number of simultaneously disrupted links, the more severe is the network disruption. If the purpose is not to obtain the entire vulnerability envelope (i.e., the upper and lower bounds of all values of n), we can replace Eq. (2) by an inequality, which corresponds to the upper and lower bounds under the simultaneous disruptions of no more than n links. Even a single link disruption could significantly deteriorate the connection of multiple O-D pairs. To link the objective function (in terms of z^w) and the basic decision variables (in terms of x_a), we introduce another set

of auxiliary variables V^w in Eq. (3) to denote the virtual maximum flow of O-D pair w. V^w (or denoted by V_{rs}) can be determined from the following lower-level subprogram for each O-D pair w (or denoted by rs). In other words, we can adapt the concept of maximum flow problem to check the O-D connectivity under each disruption scenario ($\mathbf{x}$). The meaning of Eq. (3) will be further explained in Eq. (12) after presenting the lower-level subprogram.

The network vulnerability analysis using a predefined travel demand pattern q^w can be considered as a conservative estimation (i.e., identical importance of all O-D pairs) since no available information is known about the future growth or decline under different disruption scenarios. The objective function in Eq. (1) specifies the remaining network travel throughput after disruptions as the common objective function of the vulnerability envelope, which is different from the total travel time measure typically used in vulnerability analysis. Note that the network travel throughput has been used to characterize network capacity (see, e.g., Wong and Yang (1997); Yang *et al*. (2000); Gao and Song (2002) and Chen and Kasikitwiwat (2011) with different assumptions and choice dimensions). Herein, for simplicity and also as an initial development of the vulnerability envelope, we only consider the O-D connectivity (i.e., whether an O-D pair is connected or not) to measure the O-D pair travel throughput while ignoring the travelers' rerouting effect. When an O-D pair is connected, its travel demand/throughput could be realized despite that the level of service may be degraded under disruptions. To a certain degree, this is justifiable since vulnerability analysis is mainly used to assist in planning/designing/managing a network against disruptions. However, the proposed optimization framework is flexible in the sense that different lower-level subprograms can be incorporated to include travelers' rerouting behavior. One possible model is the partial user equilibrium (PUE) model developed by Sumalee and Watling (2003), which can be used to consider travelers' adaptive route choice behaviors under disruptions. Note that this PUE model has been also adopted in Faturechi and Miller-Hooks (2014) to determine travel time resilience under disaster. In this extension, the O-D connectivity based network travel throughput could be replaced by congestion and behavior based network throughput. The O-D specific binary auxiliary variable in the upper-level objective function could be replaced by a continuous variable to represent the realized proportion of travel demand between this O-D pair.

2.2 Lower-Level Subprogram

[Alternative I: Virtual Link Capacity-Based Maximum Flow Problem]

We use the following linear program to determine V^w (V_{rs}) for each O-D pair.

$$\max_{\mathbf{y}} V^w \left(\text{or } V_{rs}\right) \tag{6}$$

$$s.t. \quad \sum_{a\in O_i} y_a - \sum_{a\in I_i} y_a = \begin{cases} +V^w, & \text{if } i = r \\ -V^w, & \text{if } i = s \\ 0, & \forall i \in N, i \neq r, s \end{cases}, \tag{7}$$

$$0 \le y_a \le 1, \ \forall a \in A, \tag{8}$$

$$y_a \le 1 - x_a, \ \forall a \in A, \tag{9}$$

where y_a is a binary decision variable: if y_a=1, link a is selected; and 0 otherwise; O_i is the set of links emanating from node i; I_i is the set of links going into node i; N is the set of nodes. The objective function in Eq. (6) is the virtual maximum flow between O-D pairs rs. If the virtual maximum flow is equal to 0, then this O-D pair is disconnected (i.e., does not exist any feasible route) under the disruption scenario **x**. Eq. (7) is the flow conservation constraint at origin, destination, and non-centroid nodes. In Eq. (8), we set the virtual link capacity as one unit for the BLP model solvability. In addition, we add Eq. (9) to characterize the mutually exclusive relationship between x_a and y_a,

$$y_a \le 1 - x_a \Leftrightarrow \begin{cases} \text{if } x_a = 1 \left(\text{disrupted}\right), & y_a = 0 \quad \left(\text{not usable in any path}\right) \\ \text{if } x_a = 0 \left(\text{connected}\right), & y_a = 0 \text{ or } 1 \quad \left(\text{usable}\right) \end{cases}. \tag{10}$$

Considering x_a is a binary integer variable, Eqs. (8) and (9) can be simplified as

$$0 \le y_a \le 1 - x_a, \ \forall a \in A. \tag{11}$$

One can readily observe that the lower-level subprogram in Eqs. (6)-(9) is a simple linear program (LP). According to the Integral Flow Theorem (Ford and Fulkerson, 1956), if each link has integral capacity, there exists an integral maximal flow. From Eq. (8), one can see if an O-D pair is connected, the virtual maximum flow determined by the lower-level subprogram is integral (at least 1); otherwise, it is zero.

With the lower-level subprogram, we can verify Eq. (3) by substituting the binary values of z^w.

$$\begin{aligned} & -M\left(1 - z^w\right) + \varepsilon \le V^w \le M z^w, \ \forall w \in W \\ \Leftrightarrow & \begin{cases} z^w = 1 \Leftrightarrow \varepsilon \le V^w \underbrace{\le M}_{\text{unrestrictive}} \Leftrightarrow V^w \ge \varepsilon > 0 \quad \left(\text{i.e., } V^w > 0\right) \\ z^w = 0 \Leftrightarrow \underbrace{-M + \varepsilon \le}_{\text{unrestrictive}} V^w \le 0 \Leftrightarrow V^w \le 0 \quad \left(\text{i.e., } V^w = 0\right) \end{cases}. \end{aligned} \tag{12}$$

This relationship is consistent with the definition of z^w. If $V^w>0$ (there is at least one feasible route between O-D pair w), then $z^w=1$ (i.e., this O-D pair is connected); if $V^w=0$, then $z^w=0$ (i.e., this O-D pair is disconnected).

As an initial development, Eq. (3) only uses the virtual maximum flow to check the O-D connectivity. As long as a route is available (i.e., $V^w>0$), this O-D pair is treated as connected. To account for capacity of the O-D connectivity, we may use the concept of volume/capacity (V/C) ratio to check if the maximum flow under disruptions is able to satisfy a minimum percentage of travel demand. For example, Eq. (3) can be replaced by

$$-M\left(1-z^w\right)+\varepsilon \le lc^w \cdot V^w - \underline{q}^w \le Mz^w,\ \ \forall w \in W\ , \tag{13}$$

where lc^w is the maximum link capacity among all links going out of the origin and going into the destination of w; $\underline{q}^w$ can be considered as the minimum travel demand that should be satisfied, e.g., 50% of q^w; $lc^w \times V^w$ can be considered as an upper bound approximation of the virtual maximum flow between O-D pair w, i.e., $\sum_{a\in O_r} c_a \cdot y_a \le \sum_{a\in O_r} lc^w \cdot y_a = lc^w \cdot V^w$, where c_a is the capacity of link a. If $lc^w \times V^w > \underline{q}^w$, then $z^w=1$ (this O-D pair is considered as connected); otherwise, $z^w=0$ (disconnected).

[Alternative II: Virtual Link Cost-Based Shortest Path Problem]

Other than using the virtual maximum flow problem (Eqs. (6)-(9)) as the lower-level subprogram, we can also check the O-D connectivity by using a virtual link cost-based shortest path problem. We introduce another auxiliary variable u^w to denote the minimum cost of O-D pair w, which can be determined by the following lower-level subprogram for each O-D pair w (or denoted by rs).

$$\min_{\mathbf{y}} u^w = \sum_{a\in A}\left(t_a + M\cdot x_a\right) y_a \tag{14}$$

s.t.

$$\sum_{a\in O_i} y_a - \sum_{a\in I_i} y_a \begin{cases} =1 & \text{, if } i=r \\ =-1 & \text{, if } i=s \\ =0, & \text{otherwise} \end{cases}, \tag{15}$$

$$y_a \ge 0,\ \ \forall a \in A\ , \tag{16}$$

where t_a is the travel time or cost of link a; y_a is a binary decision variable: $y_a=1$ if link a is used in the minimum cost route, and 0 otherwise. Eq. (14) minimizes the total travel cost for the selected links. Eq. (15) is the node conservation constraint. Note that the objective function is different from the classical shortest path problem, since it imposes a virtual link cost of using the disrupted links. If link a is disrupted ($x_a=1$), its virtual

link cost or usage penalty is $t_a + M$; otherwise, its usage cost is t_a. For a given disruption scenario **x**, the lower-level subprogram is a simple LP with respect to **y**. This continuous LP has a desirable integral property; therefore, every basic optimal solution has all variables equal to 0 or 1 (Ahuja *et al.*, 1993). The main reason is that the constraint coefficient matrix of Eq. (15) (i.e., a (+1, 0, -1) matrix, and each column has entries of exactly one 1 and one -1) is totally unimodular, which enables to relax the binary integer constraint as a continuous interval [0, 1] (Hoffman and Kurskal, 1956). Furthermore, since we deal with directed graphs without negative cycles, the solution variables can be relaxed as a simple non-negative variable. This nice property is quite valuable in rendering the lower-level subprogram with a simple non-negativity constraint and reformulating the BLP into a more solvable form.

When specifying the lower-level subprogram as Eqs. (14)-(16), Eq. (3) in the upper-level subprogram should be replaced by the following inequality, while Eqs. (1), (2), (4) and (5) keep intact.

$$-Mz^w + \varepsilon \le u^w - \overline{u}^w \le M\left(1 - z^w\right), \ \forall w \in W \ . \tag{17}$$

where $\overline{u}^w$ could be uniformly specified as a constant or defined as $\overline{u}^w = \theta u_0^w$. Note that θ (>1) denotes the travelers' allowable elongation ratio of accepting a detoured route relative to a reference cost u_0^w (e.g., the shortest route cost between O-D pair w in normal situation). When θ approaches positive infinity, travelers could accept all available routes to serve this O-D pair; when θ approaches 1, travelers only accept a very tight route elongation ratio relative to the reference cost. Typically, not-too-long routes with an acceptable travel cost are more likely to be considered by travelers as a reasonable alternative when the primary or secondary route is not available under disruptions. Eq. (17) corresponds to the following equivalent 'if-then' conditions:

$$\begin{cases} z^w = 1 \ \Leftrightarrow \underbrace{-M + \varepsilon \le}_{\text{unrestrictive}} u^w - \theta u_0^w \le 0 \Leftrightarrow u^w \le \theta u_0^w & \text{exist usable path} \\ z^w = 0 \ \Leftrightarrow \varepsilon \le u^w - \theta u_0^w \ \underbrace{\le M}_{\text{unrestrictive}} \Leftrightarrow u^w > \theta u_0^w & \text{no usable path} \ . \end{cases} \tag{18}$$

If an O-D pair is connected (z^w=1), the minimum cost u^w will be smaller than the travelers' acceptable elongated cost θu_0^w; otherwise, the O-D pair is considered as inaccessible for two possible reasons: (i) it is physically connected, but the minimum cost u^w is greater than the travelers' acceptable elongated cost θu_0^w, and (ii) it is physically disconnected (i.e., no path available), so the minimum path cost with virtual link cost is greater than a large positive number M (i.e., x_a=y_a=1, link a is disconnected but it is still virtually used under the additional penalty M).

Instead of using the objective function in Eq. (14), one may think of using $1-y_a \geq x_a$ (i.e., if link a is disrupted $x_a=1$, then it is not usable in any path $y_a=0$; if link a is not disrupted $x_a=0$, then it is usable $y_a=0$ or 1) to characterize the mutually exclusive relationship between x_a and y_a. Then, one may use $\sum_{a \in A} y_a$ (i.e., with an identical link cost) as the objective function subject to Eqs. (15)-(16) and $1-y_a \geq x_a$. However, when there is no available route (i.e., $y_a=0$ for all links), the constraint set in Eq. (15) at the origin and destination is violated. Hence, the above trial is infeasible.

The proposed optimization framework is flexible, because it allows the use of different specifications of vulnerability measures (i.e., objective functions) for different modeling purposes. Examples include the remaining network throughput (a measure of network capacity) and the remaining route diversity (a measure of network redundancy) after disruptions. The more remaining travel throughput and/or the more alternative routes are preserved, the more robust is the network to disruptions. In addition, within the BLP framework, the lower-level subprogram is also flexible. As shown in Alternatives I and II, we can use either the virtual link capacity based maximum flow problem or the virtual link cost based shortest path problem to effectively check the O-D connectivity under a network disruption scenario without path enumeration.

In a broader picture, the lower bound problem under a special case of $\theta=M$ (i.e., as long as there exists a physically connected route, this O-D pair is considered to be connected, without considering route cost constraint or route diversity requirement in checking route usability) is similar to the screen line-based traffic counting location (TCL) problem (Yang *et al*., 2001, 2006; Chootinan *et al.*, 2005; Chen *et al.*, 2007b). Yang *et al.* (2001) provided the O-D separation rule to optimally locate a given number of counting locations to separate as many O-D pairs as possible, and used a genetic algorithm (GA) to solve the TCL problem without the need of path enumeration. Yang *et al*. (2006) further formulated the TCL problem as an integer linear programming that embodies the shortest path calculation as a column generation procedure combined with a branch and bound technique. Chootinan *et al.* (2005) considered the bi-objective TCL problem, and developed a distance-based GA to solve for the non-dominated solutions that consider the tradeoff between quality and cost of coverage, while Chen *et al.* (2007b) developed strategies in the screen line-based TCL problem for selecting additional traffic counts for improving O-D trip table estimation. Bell *et al*. (2017) developed a capacity weighted spectral partitioning to identify potential flow

bottlenecks (i.e., potent sources of vulnerability) without reference to O-D matrix, paths or path assignment. Specifically, they identified the network cut with the least normalized capacity, which is closely related to the maximum flow problem but without requiring the identification of origins and destinations.

Note that Matisziw *et al*. (2007) developed a network interdiction model to identify bounds for network connectivity vulnerability in terms of facilities associated with worst-case (or best-case) impacts. The model is a binary integer linear program, which is solvable theoretically. However, it relies on path enumeration and/or path storage. On the one hand, path generation/enumeration is a non-trivial task, especially for large-scale networks. On the other hand, the quality of the generated route set (e.g., behavioral realism or reasonableness) may affect the resultant evaluation of vulnerability bounds. For example, the generated route set may exclude routes that are unattractive but become attractive after implementing some management policies (Watling *et al*., 2015). Subsequently, Matisziw and Murray (2009) developed a path-enumeration-free approach to model the O-D path availability to support disaster vulnerability assessment of network infrastructure. Specifically, a path aggregation constraint structure was proposed to eliminate path enumeration in the worst-case scenario of the flow interdiction model, which provides computational benefit over the original network interdiction model of Matisziw *et al*. (2007). Other than the conservative consideration of only the worst-case scenario, we observe that the path aggregation constraint may not be applicable to urban transportation systems, since it is not able to differentiate path length when checking the O-D connectivity. It implicitly assumes that all connected routes between an O-D pair are viable and usable, regardless of their travel distance or cost. The negative consequence of this unrealistic assumption is that the O-D connectivity could be significantly overestimated under disruptions. Travelers do not treat all simple routes (without any upper bound on distance or cost) as usable alternatives in either normal or disruption scenarios. Instead, they have some tolerance for taking alternative routes that deviate from their normal best routes. On the contrary, our proposed model explicitly captures the travelers' tolerance for accepting these alternative routes when checking the O-D connectivity under disruptions. Specifically, this modeling realism is attributed to the virtual link capacity based maximum flow formulation together with Eq. (13) or the virtual link cost based shortest path formulation together with Eq. (17). Also, Eq. (13) and Eq. (17) allow users to flexibly specify thresholds to implicitly define route availability.

3 SOLUTION ALGORITHM

Recall that the proposed model is a BLP, which is not directly solvable. The upper-level subprogram is a binary integer linear program (BILP). Both alternatives I and II of the lower-level subprogram (for a given **x**) are a continuous LP. Below, we reformulate the BLP as a single-level MILP by substituting the lower-level subprogram with its KKT conditions along with some linearization techniques.

Consider the lower-level subprogram alternative I. The Lagrangian function is (Luenberger and Ye, 2008)

$$\begin{aligned} L(\mathbf{y};\lambda;\boldsymbol{\mu};\boldsymbol{\tau}_1;\boldsymbol{\tau}_2) = &-\left(\sum_{a\in O_r} y_a - \sum_{a\in I_r} y_a\right) \\ &+\lambda\left(\left(\sum_{a\in O_r} y_a - \sum_{a\in I_r} y_a\right)+\left(\sum_{a\in O_s} y_a - \sum_{a\in I_s} y_a\right)\right) \\ &+\sum_{\substack{i\in N\\ i\neq r,s}} \mu_i\left(\sum_{a\in O_i} y_a - \sum_{a\in I_i} y_a\right) - \sum_a \tau_{a1} y_a + \sum_a \tau_{a2}\left(y_a - 1 + x_a\right), \end{aligned} \tag{19}$$

where the objective function has been replaced by the first condition of Eq. (7); λ and μ_i are the dual variables associated with the second and third conditions of Eq. (7); and τ_{a1} and τ_{a2} are the dual variables associated with $-y_a \leq 0$ and $y_a-(1-x_a)\leq 0$ in Eq. (11). Then, the KKT conditions of the lower-level subprogram can be derived as follows.

$$-\left(\delta_a^{r+}-\delta_a^{r-}\right)+\lambda\left(\delta_a^{r+}-\delta_a^{r-}+\delta_a^{s+}-\delta_a^{s-}\right)+\sum_{\substack{i\in N\\ i\neq r,s}}\mu_i\left(\delta_a^{i+}-\delta_a^{i-}\right)-\tau_{a1}+\tau_{a2}=0,\ \forall a \tag{20}$$

,

$$\left(\sum_{a\in O_r} y_a - \sum_{a\in I_r} y_a\right)+\left(\sum_{a\in O_s} y_a - \sum_{a\in I_s} y_a\right)=0, \tag{21}$$

$$\sum_{a\in O_i} y_a - \sum_{a\in I_i} y_a = 0,\ \forall i\in N, i\neq r,s, \tag{22}$$

$$\tau_{a1}\geq 0,\ y_a\geq 0,\ \tau_{a1}y_a=0,\ \forall a, \tag{23}$$

$$\tau_{a2}\geq 0,\ 1-x_a-y_a\geq 0,\ \tau_{a2}\left(1-x_a-y_a\right)=0,\ \forall a, \tag{24}$$

where $\delta_a^{i+}=1$ if link a is leaving out of node i, and 0 otherwise; $\delta_a^{i-}=1$ if link a is going into node i, and 0 otherwise. Note that Eqs. (20)-(22) are all linear equations. Eqs. (23) and (24) are complementarity conditions, making the KKT conditions being a nonlinear

system. Below, we reformulate the complementarity conditions in Eqs. (23) and (24) as a *linear system* by making full use of the problem characteristics. As mentioned before, the lower-level subprogram alternative I has an integral optimum solution (i.e., y_a at optimality will be either 0 or 1). Considering x_a is a binary integer variable, then $1\text{-}x_a \geq y_a$ in Eq. (23) implies that $1\text{-}x_a\text{-}y_a$ is also binary. With these observations, the complementarity conditions in Eqs. (23) and (24) can be reformulated as the following simple linear inequalities. We should point out that this linearization technique with a much smaller number of variables and constraints is different from the typical linearization technique of complementarity conditions (e.g., Wang and Lo, 2010).

$$
\begin{aligned}
&\begin{cases} \tau_{a1} \le M \cdot (1 - y_a) \\ \tau_{a1} \ge 0 \\ y_a \in \{0,1\} \end{cases} \\
\Leftrightarrow &\begin{cases} M \cdot (1 - y_a) > 0, \ \text{if } \tau_{a1} > 0 \\ M \cdot (1 - y_a) \ge 0, \ \text{if } \tau_{a1} = 0 \end{cases} \\
\Leftrightarrow &\begin{cases} y_a = 0, & \text{if } \tau_{a1} > 0 \\ 0 \le y_a \le 1, & \text{if } \tau_{a1} = 0 \end{cases} \Leftrightarrow 0 \le \tau_{a1} \perp y_a \ge 0, \forall a \ ,
\end{aligned}
\tag{25}
$$

$$
\begin{aligned}
&\begin{cases} \tau_{a2} \le M \cdot (1 - (1 - x_a - y_a)) \\ 1 - x_a - y_a \ge 0 \\ \tau_{a2} \ge 0 \\ y_a \in \{0,1\} \ (\text{implies } 1 - x_a - y_a \in \{0,1\}) \end{cases} \\
\Leftrightarrow &\begin{cases} M \cdot (1 - (1 - x_a - y_a)) > 0, \text{if } \tau_{a2} > 0 \\ M \cdot (1 - (1 - x_a - y_a)) \ge 0, \text{if } \tau_{a2} = 0 \end{cases} \\
\Leftrightarrow &\begin{cases} 1 - x_a - y_a = 0, & \text{if } \tau_{a2} > 0 \\ 0 \le 1 - x_a - y_a \le 1, & \text{if } \tau_{a2} = 0 \end{cases} \Leftrightarrow 0 \le \tau_{a2} \perp (1 - x_a - y_a) \ge 0, \ \forall a \ .
\end{aligned}
\tag{26}
$$

With the above equivalent reformulation in Eqs. (25)-(26), the KKT conditions becomes a linear system. Along with the BILP structure of the upper-level subprogram in Eqs. (1)-(5), the proposed BLP model has been reformulated as a single-level MILP. This desirable feature permits a number of existing algorithms in commercial software packages for its global optimal solution. For completeness, below we present the reformulated MILP for the virtual link capacity based maximum flow problem as the lower-level subprogram. Since the lower-level subprogram works for each O-D pair, we add superscript w to decision variable $\mathbf{y}$ and dual variables $\boldsymbol{\lambda}$, $\boldsymbol{\mu}$, and $\boldsymbol{\tau}$ of the lower-level subprogram.

$$
\begin{cases}
\max \text{ or } \min \sum_{w \in W} q^w z^w \\
\text{subject to}: \\
\quad \sum_{a \in A} x_a = n \\
\quad -M\left(1-z^w\right)+\varepsilon \le V^w \le M z^w, \ \forall w \\
\quad x_a = \{0,1\}, \ \forall a \\
\quad z^w = \{0,1\}, \ \forall w \\
\quad \begin{cases}
V^w = \sum_{a \in O_r} y_a^w - \sum_{a \in I_r} y_a^w \\
-\left(\delta_a^{r+} - \delta_a^{r-}\right) + \lambda^w \left(\delta_a^{r+} - \delta_a^{r-} + \delta_a^{s+} - \delta_a^{s-}\right) \\
\quad + \sum_{i \in N, i \ne r,s} \mu_i^w \left(\delta_a^{i+} - \delta_a^{i-}\right) - \tau_{a1}^w + \tau_{a2}^w = 0, \ \forall a \\
\left(\sum_{a \in O_r} y_a^w - \sum_{a \in I_r} y_a^w\right) + \left(\sum_{a \in O_s} y_a^w - \sum_{a \in I_s} y_a^w\right) = 0 \\
\sum_{a \in O_i} y_a^w - \sum_{a \in I_i} y_a^w = 0, \ \forall i \in N, i \ne r, s \\
\tau_{a1}^w \le M \cdot \left(1 - y_a^w\right), \ \forall a \\
\tau_{a2}^w \le M \cdot \left(1 - \left(1 - x_a - y_a^w\right)\right), \ \forall a \\
1 - x_a - y_a^w \ge 0, \ \tau_{a1}^w \ge 0, \ \tau_{a2}^w \ge 0, \ y_a^w \in \{0,1\}, \ \forall a
\end{cases}, \ \forall w
\end{cases}
\tag{27}
$$

Similarly, the BLP model using the lower-level subprogram alternative II in Eqs. (14)-(16) can be reformulated as the following MILP:

$$
\begin{cases}
\max \text{ or } \min \sum_{w \in W} q^w z^w \\
\text{subject to:} \\
\quad \sum_{a \in A} x_a = n \\
\quad -Mz^w + \varepsilon \le u^w - \overline{u}^w \le M\left(1 - z^w\right), \ \forall w \\
\quad x_a = \{0,1\}, \ \forall a \\
\quad z^w = \{0,1\}, \ \forall w \\
\quad \begin{cases}
u^w = \sum_a t_a y_a^w + M p_a^w \\
p_a^w \le x_a, \ p_a^w \le y_a^w, \ p_a^w \ge x_a + y_a^w - 1, \ p_a^w \ge 0, \ \forall a \\
\left(t_a + M x_a\right) + \mu_r^w \left(\delta_a^{r+} - \delta_a^{r-}\right) + \mu_s^w \left(\delta_a^{s+} - \delta_a^{s-}\right) \\
\quad + \sum_{i \in N, i \ne r,s} \mu_i^w \left(\delta_a^{i+} - \delta_a^{i-}\right) - \tau_a^w = 0, \ \forall a \\
\sum_{a \in O_i} y_a^w - \sum_{a \in I_i} y_a^w \begin{cases} = 1 \ , & \text{if } i = r \\ = -1, & \text{if } i = s \\ = 0, & \text{otherwise} \end{cases} \\
\tau_a^w \le M \cdot \left(1 - y_a^w\right), \ \tau_a^w \ge 0, \ y_a^w \in \{0,1\}, \ \forall a
\end{cases}, \ \forall w
\end{cases}
\tag{28}
$$

where μ_r^w, μ_s^w, and μ_i^w are the dual variables associated with the three conditions of Eq. (15); and τ_a^w is the dual variable associated with Eq. (16). The same logic in Eq. (25) is used to reformulate the complementarity condition $\tau_a^w y_a^w = 0$ as a linear form. We add another set of variables p_a^w in order to eliminate the bilinear term when calculating $u^w = \sum_a \left(t_a + M x_a\right) y_a^w$. There are multiple ways to deal with the bilinear terms (e.g., Hanson and Martin, 1990). Note that $x_a y_a^w = x_a$ when $y_a^w = 1$; and $x_a y_a^w = 0$ when $y_a^w = 0$. Eq. (28) uses the four inequalities associated with $p_a^w = x_a y_a^w$ together with the binary constraint of y_a^w to calculate u^w:

$$
\begin{cases} p_a^w \le x_a, & p_a^w \ge x_a + y_a^w - 1 \\ p_a^w \le y_a^w, & p_a^w \ge 0 \end{cases} \Leftrightarrow \begin{cases} y_a^w = 1 \ \Leftrightarrow \ p_a^w = x_a \\ y_a^w = 0 \ \Leftrightarrow \ p_a^w = 0 \end{cases}. \tag{29}
$$

In addition, some practical constraints may be used to enhance the computational efficiency. The maximization case seeks to have $z^w = 1$, while the minimization case seeks to have $z^w = 0$. As mentioned in Eq. (18), if $z^w = 1$, $u^w \le \theta u_0^w$; there are two possible reasons for an O-D pair not to be considered as connected: $\theta u_0^w < u^w < M$ (physically connected, but exceeds the travelers' acceptable elongated cost), or $u^w > M$ (physically disconnected). To differentiate $u^w > \theta u_0^w$ and $u^w > M$, we could define another set of binary variables h^w: $h^w = 1$ for physically connected case and 0 otherwise.

$$
\begin{aligned}
&-Mh^w + \varepsilon \le u^w - M \le M\left(1-h^w\right),\ \forall w \\
&\Leftrightarrow \begin{cases} h^w = 1 \Leftrightarrow -M + \varepsilon \le u^w - M \le 0 \ \Leftrightarrow u^w \le M \ \left(\text{physically connected}\right) \\ h^w = 0 \Leftrightarrow \quad\ \ \varepsilon \le u^w - M \le M \Leftrightarrow u^w > M \ \left(\text{physically disconnected}\right). \end{cases}
\end{aligned} \tag{30}
$$

Then, the relationship between z^w and h^w is as follows

$$
h^w \ge z^w,\ \ \forall w \ \Leftrightarrow \begin{cases} z^w = 1 \ \Rightarrow\ h^w = 1 \ \left(u^w < \theta u_0^w \ \Rightarrow\ u^w < M\right) \\ z^w = 0 \ \Rightarrow h^w \ge 0 \\ h^w = 1 \ \Rightarrow z^w \le 1 \\ h^w = 0 \ \Rightarrow z^w = 0 \ \left(u^w > M \quad \Rightarrow\ u^w > \theta u_0^w\right). \end{cases} \tag{31}
$$

When h^w=1, $p_a{}^w$ is 0 for all links (i.e., no disrupted link is used; otherwise, there exists a link with $p_a{}^w$=1); when h^w=0, at least one disrupted link is used (i.e., $p_a{}^w$=1). This relationship could be expressed as

$$
1-\sum_a p_a^w \le h^w \le 1 - p_a^w \Leftrightarrow \begin{cases} h^w = 1 \ \Leftrightarrow \ \underbrace{\sum_a p_a^w \ge 0}_{\text{unrestrictive}},\ p_a^w = 0,\ \forall a \\ h^w = 0 \ \Leftrightarrow \ \underbrace{p_a^w \le 1}_{\text{unrestrictive}},\ \sum_a p_a^w \ge 1\ . \end{cases} \tag{32}
$$

The proposed optimization model for deriving the vulnerability envelope is capable of guaranteeing a globally optimal solution from two perspectives: (1) it enables to derive both the exact upper and lower bounds (i.e., the best and worst cases) of network vulnerability under multiple simultaneous disruptions, while circumventing the need to perform a full network scan, but with an implicit consideration of all possible combinations of multiple disruptions (via the optimization approach); (2) from the optimization perspective, the equivalently reformulated MILP has a globally optimal solution, which is a desirable property for both theoretical exploration and algorithm development.

4 NUMERICAL EXAMPLES

In this section, numerical examples are provided to demonstrate the validity, capability, and flexibility of the proposed optimization framework for deriving the upper and lower vulnerability bounds. Example 1 verifies the correctness of the optimal upper and lower bounds by comparing with the complete enumeration approach, and also explores the implication of the vulnerability envelope. Example 2 demonstrates the flexibility of the proposed framework in terms of allowing different models in the lower-level subprogram, and implicitly defining route availability without path enumeration.

Example 3 compares the vulnerability envelope between individual link disruption and pairwise link disruption. A simple network shown in Figure 1 is used for ease of results exposition. This network consists of 6 nodes and 16 directed links. Examples 1 and 2 consider 14 O-D pairs (i.e., (1, 2), (1, 3), (1, 4), (2, 1), (2, 3), (2, 4), (3, 1), (3, 2), (3, 4), (4, 1), (4, 2), (4, 3), (5, 6), and (6, 5)), and Example 3 considers all 30 (i.e., 6×5) O-D pairs. For simplicity, all O-D pairs are assumed to have the same travel demand of 1 unit. Accordingly, the common objective function degenerates to the total number of available/usable/connected O-D pairs. As to the maximum flow-based model, we simplify the term $lc^w \cdot V^w - \underline{q}^w$ in Eq. (13) as V^w-p. When p=0, Eq. (13) means this O-D pair is connected as long as V^w-0>0; When p=1, Eq. (13) means this O-D pair w is considered as connected only when V^w-1>0. As to the shortest path-based model, we set all links to have the same travel time of 1 unit, which will be used in Eq. (14). One can readily consider link-specific travel time and O-D pair specific maximum flow constraint and travel demand in the proposed model.

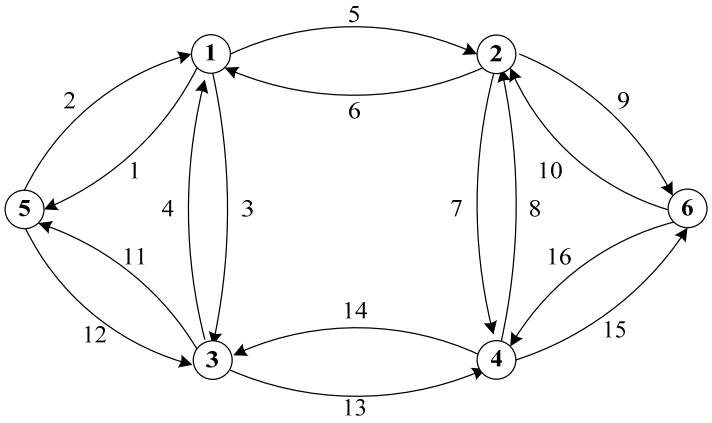


Figure 1 Example network

4.1 Example 1

First of all, we solve the proposed vulnerability bound model with both lower-level subprogram alternatives: the virtual link capacity-based maximum flow problem and the virtual link cost-based shortest path problem. Specifically, we set p=0 in the maximum flow-based model and θ=M (a very large positive constant) in the shortest path-based model. Figure 2 shows the optimal upper and lower bounds of the remaining throughput (i.e., objective value) as a function of the number of disrupted links (n). Both the maximum flow-based and the shortest path-based models generate exactly the same vulnerability bounds. This is consistent with the setting of p=0 and θ=M, which

means that: as long as an O-D pair is physically connected (i.e., $V^w>0$ or $u^w<M$), the travel throughput of this O-D pair is considered to be realized.

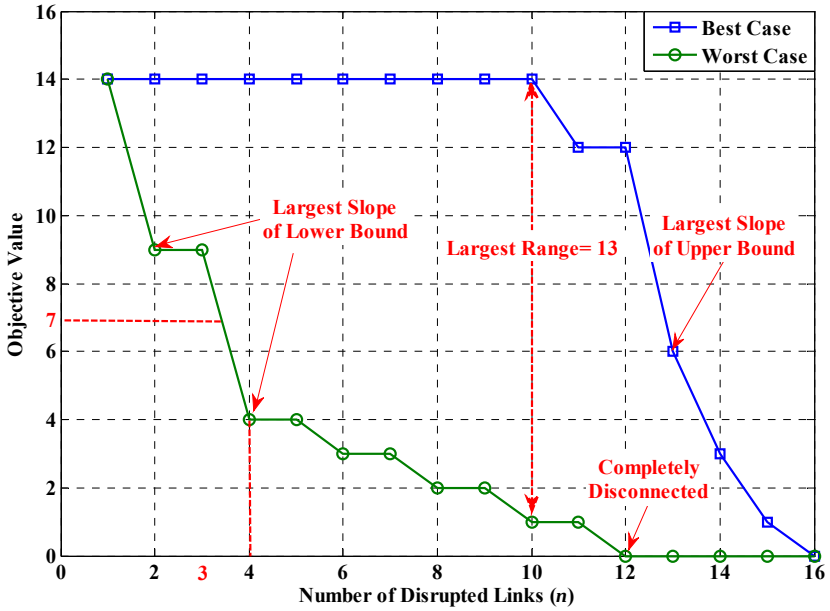


Figure 2 Upper and lower bounds (shortest path-based model, $\theta=M$)

To verify the correctness of the optimal upper and lower bounds, without loss of generality, we set n=4 and enumerate all possible combinatorial scenarios (C_{16}^4=1820). For each scenario, we use the virtual link cost-based shortest path problem to check the connectivity of all 14 O-D pairs. Figure 3 shows the boxplot of the objective value (i.e., the number of connected O-D pairs) from the complete enumeration approach. One can see that the maximum objective value is 14 and the minimum objective value is 4. These complete enumeration results verify the optimal upper and lower bounds obtained from the proposed approach shown in Figure 2. We should point out that the complete enumeration approach has a combinatorial complexity, and hence is computationally burdensome. For this simple example, we need to solve 25,480 (i.e., 1820 scenarios × 14 O-D pairs) linear programs of the virtual link cost-based shortest path problem. The entire objective value distribution is unknown without a complete enumeration of all possible combinations. However, the proposed optimization approach is able to derive the exact upper and lower bounds (i.e., range) without the need to conduct a complete network scan.

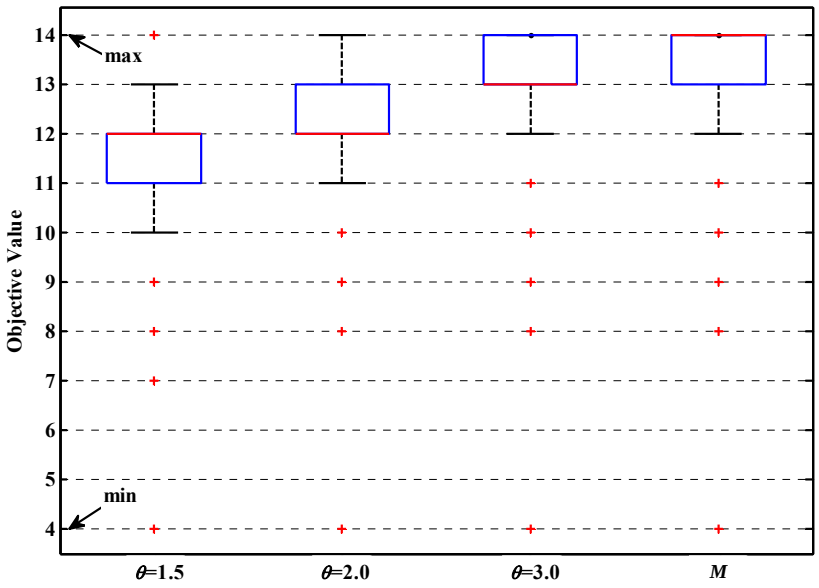


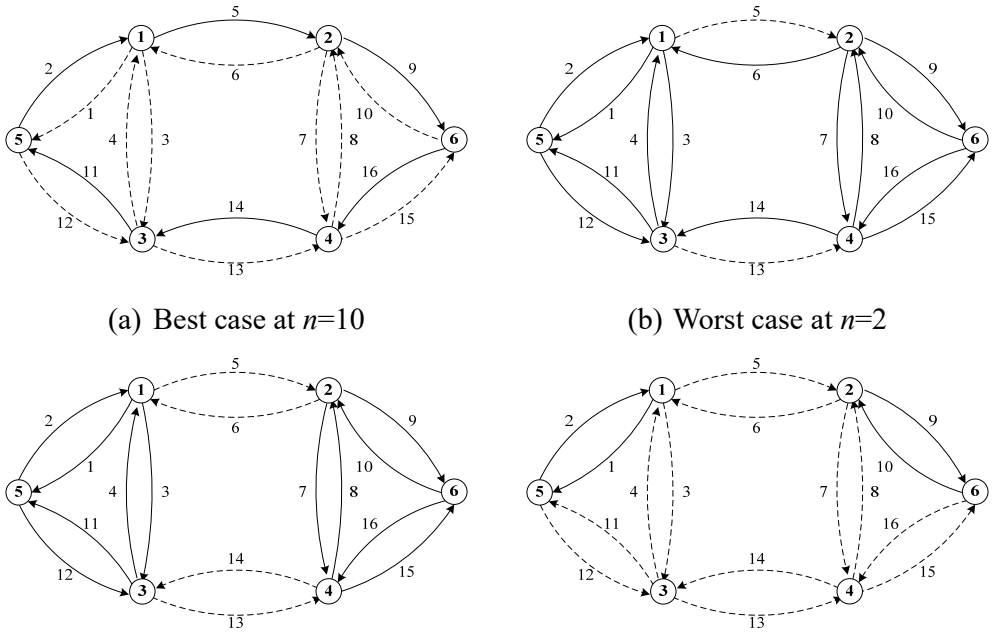
Figure 3 Boxplot of objective values obtained from complete enumeration approach

Below we explore the details of the optimal vulnerability bounds shown in Figure 2. The following observations can be drawn:

- For each number of disrupted links, the upper and lower bounds capture all possible disruption scenarios provided by the complete enumeration approach. For all numbers of disrupted links, these two bounds form a vulnerability envelope. Both the upper and lower bounds of the remaining throughput are not strictly decreasing with the number of disrupted links, nor a convex/concave curve with respect to n. Overall, the consequence of more disrupted links becomes more severe regardless of the optimistic or pessimistic evaluation.
- The reformulations presented in Section 3 are equivalent without approximation. Due to the LP structure of the proposed reformulation model, the derived upper and lower bounds are exact and globally optimal. However, the LP structure also leads to multiplicity of optimal solutions (i.e., non-unique optimal solutions in terms of disruption location $\mathbf{x}$ that can achieve the same objective value).
- The best and worst case curves intersect at n=1 (no connectivity loss) and n=16 (completely disconnected). The largest vulnerability range is 13 occurred at n=10. Among the 16 values of n, 12 scenarios have a range of vulnerability greater than 5, and 9 scenarios have a range of vulnerability greater than 10. These vulnerability ranges indeed highlight the importance of systematically considering all possible simultaneous disruptions as well as the efficacy of the proposed optimization approach. A complete network scan is too computational burdensome, while a partial scan may miss some critical scenarios that may lead to a biased vulnerability

assessment. These vulnerability ranges could be used to assist in developing possible system improvements in the contingency planning. The disrupted links that result in the largest system loss should be protected as much as possible, while the disrupted links that result in the least system loss could be ignored especially with limited resources in planning contingencies.

- As to the best case curve, due to the relatively slack setting of p=0 and θ=M, the O-D connectivity does not decrease until n=11. From the optimistic perspective, this network allows at most 10 disrupted links, and the remaining 6 links constitute a closed loop for guaranteeing a full connectivity of all 14 O-D pairs as shown in Figure 4(a). When n=11, the maximum objective value reduces to 12, due to the connectivity loss of O-D pairs (5, 6) and (6, 5) as shown in Figure 5. The largest slope occurs at n=13, which reduces the upper bound from 12 (85% connectivity) at n=12 to 6 (42% connectivity) at n=13.



(a) Best case at n=10 (b) Worst case at n=2

(c) Worst case at n=4 (d) Worst case at n=12

Figure 4 Network configuration after certain disruptions (solid line: connected link; dashed line: disconnected link)

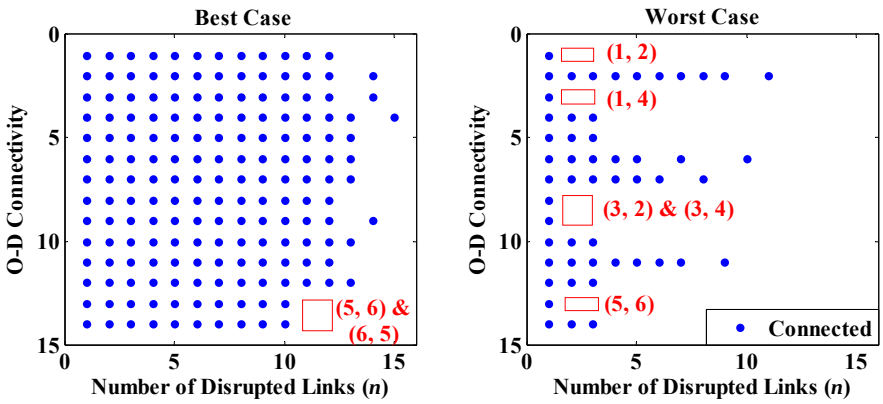


Figure 5 Change of O-D connectivity under various numbers of disrupted links

- As to the worst case curve, any single link disruption (n=1) does not degrade the full connectivity of the 14 O-D pairs, while n=2 reduces the lower bound to 9. O-D pairs (1, 2), (1, 4), (3, 2), (3, 4) and (5, 6) become disconnected due to the disruptions on links 5 and 13 as shown in Figure 4(b) and Figure 5. These five O-D pairs are sensitive or vulnerable O-D pairs for this particular network. The largest slope of lower bound occurs at n=2 and n=4, which reduce the objective value by 5 units. Also, n=4 is a critical point, starting from which the worst network connectivity is less than 50% (i.e., 7) and half of the 14 O-D pairs are disconnected. From Figure 4(c), the disruptions on links 5, 6, 13, and 14 lead to the disconnection of 10 O-D pairs, while only zones 1 and 3 and zones 2 and 4 are still connected. The disrupted links corresponding to the largest slope of lower bound (e.g., at n=2 and n=4) should be strategically protected to maintain network resiliency against disruptions, whose failures could create a dramatic degradation of the vulnerability bounds. The complete network disconnection happens at n=12, where only links 1, 2, 9 and 10 shown in Figure 4(d) are available. The earlier the completely disconnected point occurs, the less survivable is the network to disruptions.

4.2 Example 2

This section extends Example 1 to demonstrate the flexibility of the proposed framework in terms of allowing different lower-level subprograms, as well as implicitly defining route availability without path enumeration. First, we examine the impact of p on the maximum flow-based vulnerability bounds. Figure 6 shows the optimal upper and lower bounds obtained from the maximum flow-based model with p=0 and p=1. The upper (and lower) bound curve with p=0 is above that with p=1. The reason is that p=1 represents a tighter requirement of O-D connectivity (V^w>1). An O-D pair is considered to be connected only when it has at least two units of virtual maximum flow

or two distinct routes. The lower bound with p=1 is equal to 9 at n=1, rather than 14. In this scenario, the disruption of link 5 reduces the virtual maximum flow of O-D pairs (1, 2), (1, 4), (3, 2), (3, 4) and (5, 6) to a single unit, which is considered as disconnected with p=1. Again, these five sensitive O-D pairs are consistent with p=0. From this perspective, the upper and lower bounds may not necessarily constitute a closed vulnerability envelope.

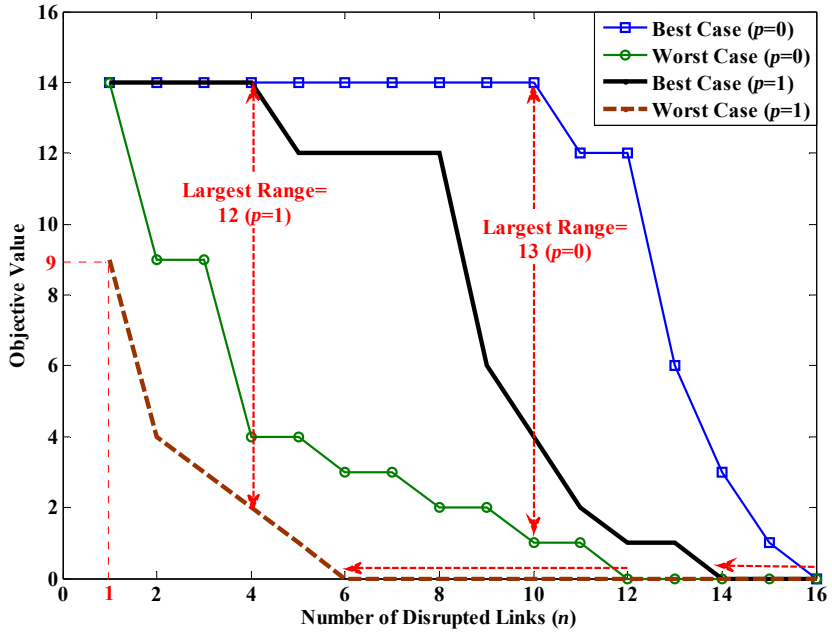


Figure 6 Upper and lower bounds (maximum flow-based model, p=0 vs. p=1)

The range between upper and lower bounds informs how resilient the network is against disruptions. The largest range with p=1 appears early (at n=4) than that with p=0 (at n=10). The best case starts to lose its O-D connectivity starting from n=4 (versus n=10 with p=0). At n=5, failures on links 7, 10, 11, 12, and 15 reduce the virtual maximum flow of O-D pairs (5, 6) and (6, 5) to a single unit. Similarly, the best case gets completely disconnected at n=14 (versus n=16 with p=0); and the worst case gets completely disconnected even much earlier at n=6 (versus n=12 with p=0). The early occurrence of the completely disconnected points indicates that the network is less survivable to disruptions. Disruptions on links 3, 4, 8, 9, 13, and 14 (i.e., n=6) make all 14 O-D pairs with V^w=1. Also, the cumulative range between upper and lower bounds is 122 at p=0 and 99 at p=1.

Similar to Figure 6, Figure 7 shows the impact of θ on the shortest path-based vulnerability bounds. Specifically, we examine the upper and lower bounds under

θ=1.5, 2.0, and 3.0. Recall that θ (>1) denotes the travelers' allowable elongation ratio of accepting a detoured route relative to the normal situation. A smaller value of θ indicates that the travelers are only willing to accept a small route elongation relative to the normal situation. As shown in Figure 7, with the increase of θ, both the upper and lower bounds shift to the upper right corner. The increase of θ from 1.5 to 3.0 right shifts the starting point of upper bound reduction from 4 to 6, and then to 8, and the completely disconnected point of lower bound from 8 to 12. To give an example on the distinguishing capability of θ, we look at the upper bound with θ=3 at n=8, where at most eight links (i.e., 1, 3, 6, 8, 10, 12, 13, and 15) can be disrupted without impacting the complete network connectivity. This is different from Figure 2 that at most ten links can be disrupted for complete network connectivity when θ=M. The closed loop constituted by the six links shown in Figure 4(a) makes some O-D pairs unacceptably elongated, e.g., u^w=4 for O-D pairs (1, 3) and (4, 2) and u^w=5 for O-D pairs (2, 1) and (3, 4). The shortest path-based model with a reasonable route elongation ratio is able to avoid this unreasonable result. In summary, the above results on the impacts of p and θ indicate that ignoring the behavioral requirement of route usability (i.e., simply use physically connected p=0 or θ=M as behaviorally usable) overestimates both the optimistic and pessimistic estimates of network vulnerability under disruptions.

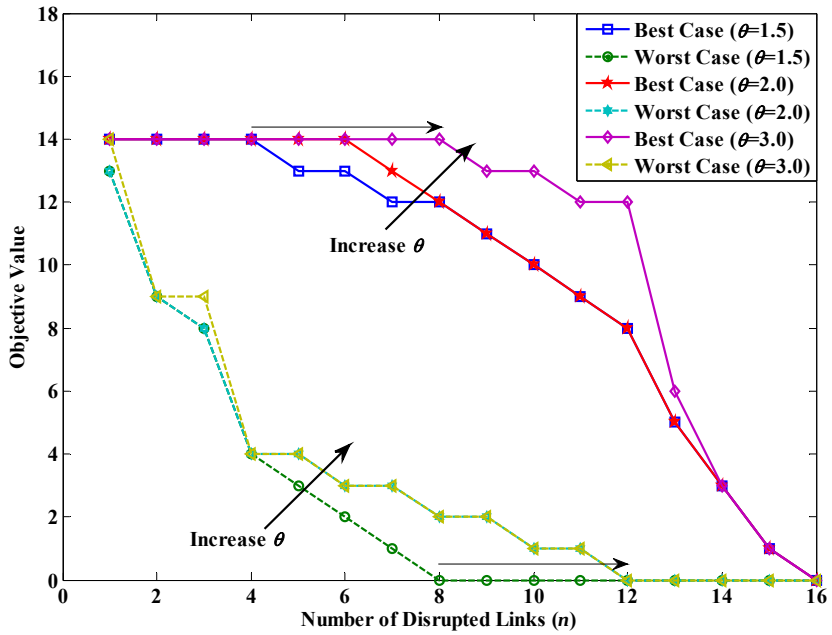


Figure 7 Upper and lower bounds (shortest path-based model, θ=1.5, 2.0, and 3.0)

Both lower-level subprogram alternatives have their own advantages. On the one hand, the maximum flow-based lower-level subprogram has fewer decision variables than the shortest path-based lower-level subprogram, rendering more computational efficiency.

The reason is that the shortest path-based lower-level subprogram has a bilinear objective function, which is reformulated as four linear inequalities associated with an additional variable p_a^w as shown in Eq. (29). On the other hand, the shortest path-based lower-level subprogram has a more intuitive way of modeling route availability without path enumeration (via the travelers' allowable elongation ratio for taking alternative routes that deviate from their normal best routes) in Eqs. (17)-(18).

4.3 Example 3

Figure 8 compares the vulnerability envelope between individual link disruption and pairwise link disruption. The individual link disruption means that there is no restriction on the disrupted link location, while the pairwise link disruption means that the two-way links are disrupted simultaneously. The pairwise link disruption may occur at two-way two-lane roads or during large-scale infrastructure failure due to earthquake or flood, etc. Herein we consider all 30 (i.e., 6×5) O-D pairs. One can see that Figure 8 with 30 O-D pairs is significantly different from Figure 2 with 14 O-D pairs. As to the individual link disruption, both the upper and lower bounds degrade to zero at n=16. The largest vulnerability range is 25 (83%) also occurred at n=10. The vulnerability envelope of the pairwise link disruption is within that of the individual link disruption. This is consistent with optimization theory. More constraints in the pairwise link disruption decrease the maximum objective value and increase the minimum objective value. The upper bound starts to degrade at n=8 due to the additional disruptions on links 9 and 10, and the lower bound has a significant reduction at n=4 due to the additional disruptions on links 5 and 6 or 13 and 14.

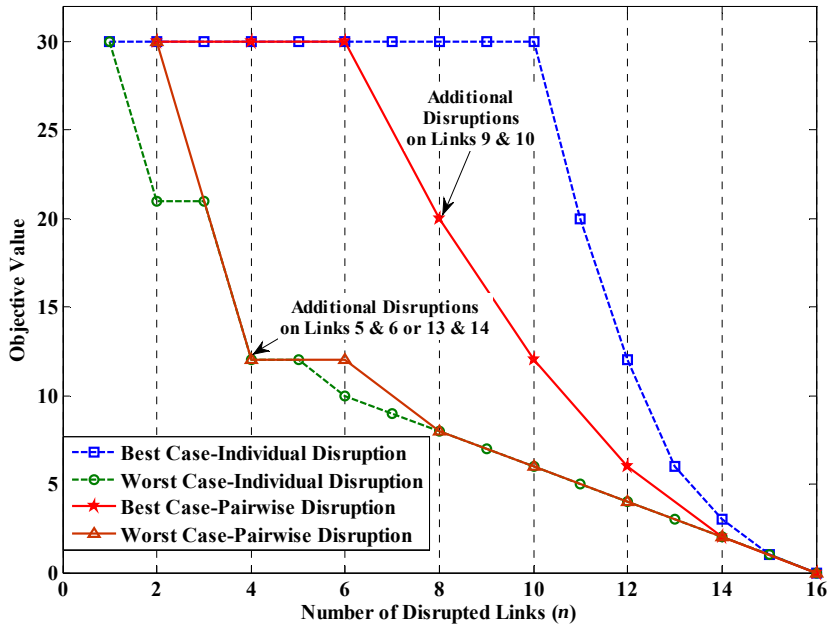


Figure 8 Vulnerability bound comparison between individual and pairwise disruptions

5 CONCLUDING REMARKS

In this paper, we developed an optimization framework for deriving both the upper and lower bounds of network vulnerability envelope under multiple simultaneous disruptions, while circumventing the need of enumerating all possible disruption scenarios. The vulnerability bound envelope constructed by the upper and lower bounds is capable of considering all possible combinations while avoiding the combinatorial complexity of enumerating multi-disruption scenarios. With the vulnerability envelope as a network performance assessment tool, planners and managers could more cost-effectively plan for system protection against disruptions, and prioritize system improvements to minimize disruption risks with limited resources.

Mathematically, we formulated the upper and lower bounds of network vulnerability as a BLP. The upper-level subprogram maximizes or minimizes the remaining network travel throughput under a given number of disrupted links, which corresponds to the upper and lower vulnerability bounds. The lower-level subprogram checks the connectivity or usability of each O-D pair under a network disruption scenario. The virtual link capacity-based maximum flow problem formulation and the virtual link cost-based shortest path problem formulation were developed as two alternative modeling approaches for the lower-level subprogram. To solve the proposed BLP model, we reformulated it as a single-level MILP, such that it could be solvable by

existing solvers for its globally optimal solution without path enumeration. Numerical examples were also provided to systematically demonstrate the validity, capability, and flexibility of the proposed optimization model. The proposed modeling framework allowed flexible specifications of the common objective function for different modeling purposes, and flexible modeling approaches to check the O-D connectivity and to implicitly define route usability without path enumeration.

A few directions are worthy of further investigations: (1) To further consider traveler's rerouting effect on the remaining network travel throughput, we will include a suitable traffic assignment model as the lower-level subprogram. The O-D specific binary auxiliary variable in the upper-level objective function could be replaced by a continuous variable to represent the realized proportion of travel demand between this O-D pair. (2) We will explore other specifications of vulnerability measures, e.g., route diversity as a measure of network redundancy (Xu *et al*., 2015). (3) The current paper focused on how to model the network vulnerability envelope in an optimization framework and how to reformulate it into a more solvable form. Computational efforts required for solving MILPs are complex and difficult (i.e., solving large-scale MILPs is computationally expensive). More efficient algorithms with a better utilization of the problem/model structure (e.g., use parallel computing due to the O-D pair specific lower-level subprogram) should be developed for enhancing its practical application in large-scale realistic networks. (4) It is also of interest to extend to consider node-based disruptions and interdependent infrastructures.

ACKNOWLEDGEMENTS

The authors are grateful to three anonymous referees for their constructive comments and suggestions to improve the quality and clarity of the paper. The work described in this paper was jointly supported by research grants from the National Natural Science Foundation of China (51408433), the Research Grants Council of the Hong Kong Special Administrative Region (Project No. PolyU 15267116), the Research Committee of the Hong Kong Polytechnic University (Project No. 1-ZE5T), the Chang Jiang Chair Professorship Program of the Ministry of Education in China, Fundamental Research Funds for the Central Universities of China, and Chenguang Program sponsored by Shanghai Education Development Foundation and Shanghai Municipal Education Commission. These supports are gratefully acknowledged.

REFERENCES

Ahuja, R.K., Magnanti, T.L., Orlin, J.B., 1993. *Network flows: theory, algorithms and*

applications. Prentice Hall.

Bell, M.G.H., 2000. A game theory approach to measuring the performance reliability of transport networks. *Transportation Research Part B* 34(6), 533–545.

Bell, M.G.H., Cassir, C., 2002. Risk-averse user equilibrium traffic assignment: an application of game theory. *Transportation Research Part B* 36, 671-681.

Bell, M.G.H., Kuaruchi, F., Perera, S., Wong, W., 2017. Investigating transport network vulnerability by capacity weighted spectral analysis. *Transportation Research Part B* 99, 251–266.

Berdica, K., 2002. An introduction to road vulnerability: what has been done, is done and should be done. *Transport Policy* 9(2), 117-127.

Chen, A., Kasikitwiwat, P., 2011. Modeling network capacity flexibility of transportation networks. *Transportation Research Part A* 45(2), 105-117.

Chen, A., Kongsomsaksakul, S., Zhou, Z., Lee, M., Recker, W., 2007a. Assessing network vulnerability of degradable transportation systems: An accessibility-based approach. *In Proceedings of the 17th International Symposium of Transportation and Traffic Theory*, Elsevier. Edited by M.G.H. Bell, B. Heydecker, and R. Allsop, pp. 236-262.

Chen, A., Pravinvongvuth, S., Chootinan, P., Lee, M., Recker, W., 2007b. Strategies for selecting additional traffic counts for improving O-D trip table estimation. *Transportmetrica* 3(3), 191-211.

Chen, A., Yang, C., Kongsomsaksakul, S., Lee, M., 2007c. Network-based accessibility measures for vulnerability analysis of degradable transportation networks. *Network and Spatial Economics* 7(3), 241-256.

Chen, A., Yang, H., Lo, H.K., Tang, W., 2002. Capacity reliability of a road network: an assessment methodology and numerical results. *Transportation Research Part B* 36, 225-252.

Chen, B., Lam, W.H.K., Sumalee, A., Li, Q., Li, Z., 2012. Vulnerability analysis for large scale and congested road networks with demand uncertainty. *Transportation Research Part A* 46, 501-516.

Chootinan, P., Chen, A., Yang, H., 2005. A bi-objective traffic counting location model for origin-destination trip table estimation. *Transportmetrica* 1(1), 65-80.

Faturechi, R., Miller-Hooks, E., 2014. Travel time resilience of roadway networks under disaster. *Transportation Research Part B* 70, 47-64.

Ford, L.R., Fulkerson, D.R., 1956. Maximal flow through a network. *Canadian Journal of Mathematics* 8, 399-404.

Gao, Z.Y., Song, Y.F., 2002. A reserve capacity model of optimal signal control with user-equilibrium route choice. *Transportation Research Part B* 36(4), 313-323.

Hanson, W., Martin, R.K., 1990. Optimal bundle pricing. *Management Science* 36(2), 155-174.

He, X., Zheng, H., Peeta, S., 2015. Model and a solution algorithm for the dynamic resource allocation problem for large-scale transportation network evacuation. *Transportation Research Part C* 59, 233–247.

Ho, H.W., Sumalee, A., Lam, W.H.K., Szeto, W.Y., 2013. A continuum modeling approach for network vulnerability analysis at regional scale. *Procedia – Social*

and Behavioral Sciences (Proceedings of the 20th International Symposium on Transportation and Traffic Theory) 80, 846-859.

Hoffman, A.J., Kruskal, J.B., 1956. Integral boundary points of convex polyhedra. In: *Linear Inequalities and Related Systems* (editted by Kuhn, H.W., Tucker, A.W.), Princeton University Press, pp. 223–246.

Jenelius, E., 2010. Redundancy importance: links as rerouting alternatives during road network disruptions. *Procedia Engineering* 3, 129-137.

Jenelius, E., Mattsson, L.G., 2015. Road network vulnerability analysis: Conceptualization, implementation and application. *Computers, Environment and Urban Systems* 49, 136-147.

Luathep, P., Sumalee A., Ho.H.W., Kurauchi, F., 2011. Large-scale road network vulnerability analysis: a sensitivity analysis based approach. *Transportation* 38(5) 799-817.

Luenberger, D.G., Ye, Y., 2008. *Linear and Nonlinear Programming* (3rd edition), Springer.

Matisziw, T.C., Murray, A.T., 2009. Modeling s-t path availability to support disaster vulnerability assessment of network infrastructure. *Computers & Operations Research* 36, 16-26.

Matisziw, T.C., Murray, A.T., Grubesic, T.H, 2007. Bounding network interdiction vulnerability through cutset identification. In: *Critical infrastructure: reliability and vulnerability. Advances in Spatial Science*, Springer-Verlag. Edited by A.T. Murray, and T.H. Grubesic. pp. 243-256.

Mattsson, L.G, Jenelius, E., 2015. Vulnerability and resilience of transport systems: A discussion of recent research. *Transportation Research Part A* 81, 16-34.

Murray, A.T., Grubesic, T.H., 2007. *Critical Infrastructure: Reliability and Vulnerability*. Springer-Verlag, Berlin, Heidelberg.

Murray-Tuite, P., Wolshon, B., 2013. Evacuation transportation modeling: An overview of research, development, and practice. *Transportation Research Part C* 27, 25–45.

Nagurney, A., Qiang, Q., 2010. *Fragile Networks: Identifying Vulnerabilities and Synergies in an Uncertain World*. John Wiley & Sons, Inc.

Nicholson, A., Du, Z.P., 1997. Degradable transportation systems: an integrated equilibrium model. *Transportation Research Part B* 31(3), 209-223.

Sumalee, A., Watling, D.P., 2003. Travel time reliability in a network with dependent link modes and partial driver response. *Journal of the Eastern Asia Society for Transportation Studies* 5, 1687-1701.

Szeto, W.Y., O'Brien, L., O'Mahony, M, 2007. Generalisation of risk averse traffic assignment. In *Transportation and Traffic Theory*, edited by R. E. Allsop, M. G. H. Bell, and B. G. Heydecker, 127-153. Amsterdam: Elsevier Science.

Wang, D.Z.W., Liu, H., Szeto, W.Y., Chow, A.H.F., 2016. Identification of critical combination of vulnerable links in transportation networks – a global optimisation approach. *Transportmetrica A: Transport Science* 12(4), 346-365.

Wang, D.Z.W., Lo, H.K., 2010. Global optimum of the linearized network design problem with equilibrium flows. *Transportation Research Part B* 44(4), 482-492.

Wang, L., Yang, L., Gao, Z., Li, S., Zhou, X., 2016. Evacuation planning for disaster responses: A stochastic programming framework. *Transportation Research Part C* 69, 150–172.

Watling, D.P., Rasmussen, T.K., Prato, C.G., Nielsen, O.A., 2015. Stochastic user equilibrium with equilibrated choice sets: Part I-Model formulations under alternative distributions and restrictions. *Transportation Research Part B* 77, 166–181.

Wong, S.C., Yang, H., 1997. Reserve capacity of a signal-controlled road network. *Transportation Research Part B* 31(5), 397-402.

Xu, X., Chen, A., Jansuwan, S., Heaslip, K, Yang, C., 2015. Modeling transportation network redundancy. *Transportation Research Procedia* (21st International Symposium on Transportation and Traffic Theory) 9, 283-302.

Yang, C., Chen, A., Xu, X., Wong, S.C., 2013. Sensitivity-based uncertainty analysis of a combined travel demand model. *Transportation Research Part B* 57, 225-244.

Yang, H., Bell, M.G.H., Meng, Q., 2000. Modeling the capacity and level of service of urban transportation networks. *Transportation Research Part B* 34(4), 255-275.

Yang, H., Gan, L., Tang, W.H., 2001. Determining cordons and screen lines for origin-destination trip studies. *Proceedings of the 3rd Eastern Asia Society for Transportation Studies* 3(2), 85-99.

Yang, H., Yang, C., Gan, L., 2006. Models and algorithms for the screen line-based traffic-counting location problems. *Computers & Operations Research* 33, 836–858.

Zhao, W., Alam, S., Abbass, H.A., 2013. Evaluating ground–air network vulnerabilities in an integrated terminal maneuvering area using co-evolutionary computational red teaming. *Transportation Research Part C* 29, 32-54.