A budget-constrained partial protection planning of a rail intermodal terminal network

Hassan Sarhadi¹ · Satyaveer S. Chauhan² · Manish Verma³

Received: 21 March 2020 / Accepted: 7 April 2021

Published online: 01 May 2021

© The Author(s) 2021 [OPEN](#)

Abstract

Rail-truck intermodal transportation is an essential component of freight transportation in North America, and thus, its associated infrastructure is deemed crucial for the wellbeing of the society. In this paper, a budget-constrained partial protection planning model is proposed that addresses the fortification of a rail intermodal network such that the effects of an intentional disruption are minimized. Novel solution methodologies that make use of metaheuristic approaches and a decomposition approach are proposed to solve the challenging tri-level mixed-integer mathematical model. The proposed analytical approaches are then used to solve and analyze problem instances generated using the realistic infrastructure of a major railroad operator. Finally, the computational efficiency and effectiveness of the proposed approaches over the existing solution techniques in the literature are discussed and future research directions are outlined.

Article Highlights

- The paper suggests partial protection of facilities together with two metaheuristic approaches to solve the model.
- The proposed metaheuristic approaches are efficient in solving the model compared to the existing exact approach.
- Having the option to partially protect facilities provides effective use of scarce defensive resources.

Keywords Partial protection planning · Intermodal transportation · Tabu search metaheuristic · Genetic algorithm · Decomposition · Fortification

Mathematics Subject Classification 90-XX (Operations Research, Mathematical Programming)

JEL Classification C61 (Optimization Techniques)

✉ Hassan Sarhadi, hassan.sarhadi@acadiau.ca; Satyaveer S. Chauhan, satyaveer.chauhan@concordia.ca; Manish Verma, mverma@mcmaster.ca | ¹F.C. Manning School of Business, Acadia University, Wolfville, Canada. ²John Molson School of Business, Concordia University, Montreal, Canada. ³DeGroote School of Business, McMaster University, Hamilton, Canada.

1 Introduction

Rail-truck intermodal transportation is currently the largest revenue segment for the railroad industry in North America and is expected to continue its growth into the future (ref. [1]). Such a popularity for rail-truck intermodal transportation services stems from both commercial and environmental reasons. From the commercial perspective, it combines the best attributes of the two distinct modes of transportation which in turn eliminates the inefficiency of long-haul trucking (ref. [2, 3]) at enjoys from the schedule-based operations of intermodal trains to reduce uncertainty associated with delivery and lead-time (ref. [4]). Furthermore, from the environmental preprojective, it shows a better track record and thus has been favored by public and private sectors to mitigate highway congestion (ref. [5]) and to reduce carbon emissions (ref. [6]).

It is evident that intermodal transportation plays a crucial role in the economic viability of North America, and hence, any disruption in its services would entail a crippling effect on security, economy, public health, and safety of the society (ref. [7]). There are two ways to mitigate the adverse impact of disruptions: *first*, to design infrastructure so that it operates efficiently both in normal conditions and when a disruption occurs (ref. [8]); and *second*, to assess the vulnerability to disruptions and then develop strategies to preclude it. This paper proposes a novel methodology to determine the best fortification strategy in the wake of intentional disruption of transportation systems and thus falls under the second category. It should be added that intentional disruption of transportation infrastructure is a major concern for transportation companies as every year they witness several disruptions on their railroad and highway networks (ref. [9, 10]). For example, major Canadian railroad companies faced several blockades on their networks in February 2020 which eventually led to shutting down many railyards across the country and caused financial damages to railroads and other related businesses (ref. [11]). In this paper, a detailed review of the relevant literature of intentional disruption domain is provided in the next section which has been a mainstream area of research since the 9/11 terrorist attacks.

The first attempt to study intentional disruption of rail intermodal terminals (that is fixed facilities) within a transportation context is done by ref. [12] wherein a tri-level optimization program was proposed, where the outer problem belongs to the network operator with a limited budget to completely fortify some of the terminals, the middle level to the attacker with enough resources to fully interdict some of the unprotected terminals, and the

inner level to the intermodal operator who attempts to meet the demand on a reduced network. Although the existing optimization packages could solve small problem instances, the realistic size problem challenged their capability. Hence, the authors proposed a two-stage heuristic solution methodology to solve larger problem instances wherein the *first* stage employed an implicit enumeration technique proposed in ref. [13] to break the tri-level problem into a set of bi-level subproblems, and the *second* stage solved the resulting bi-level subproblems using a decomposition-based solution technique to determine the lowest cost subsequent attack, which was also the best fortification plan.

It is pertinent to add that given the exploratory nature of the above research efforts, the authors had adopted the traditional approach to fortification, viz. fortified assets are completely immune to attacks, while unfortified ones are completely vulnerable. In the proposed study, the assumption about complete fortification is relaxed and the network operator can partially protect rail terminals. More precisely, it is assumed that the acquired immunity is proportional to the amount of defensive resources invested in each asset, i.e., while spending up to a certain level will ensure complete immunity, anything less would leave the asset vulnerable to attacks. Doing so, indeed, provides more flexibility to the network operator who can judiciously distribute limited defensive resources to fully protect important assets and leave others partially protected or even unprotected. To incorporate the partial protection feature, the tri-level optimization program proposed in ref. [12] is adapted, which is then solved via two proposed metaheuristic-decomposition-based solution methodologies. The proposed framework is then tested on problem instances generated using the realistic infrastructure of a Class I railroad operator and the proposed solution methodology returns optimal solutions much faster than the existing techniques. The results show that the partial protection approach presents larger number of possible strategies for the decision-makers, which in turn increases flexibility and the performance of the intermodal transportation system.

The rest of the paper is organized as follows: Sect. 2 reviews the relevant literature, followed by the problem description in Sect. 3; mathematical models including assumptions are outlined in Sect. 4, followed by a discussion about the solution methodology in Sect. 5; comprehensive outline of the case study setting including parameters is contained in Sect. 6, while the solution and relevant analyses are discussed in Sect. 7, and, finally, conclusions, contributions, and directions of future research are outlined in Sect. 8.

2 Literature review

The relevant literature can be organized under three streams: protection and fortification planning; metaheuristic applications in fortification; and disruption in rail-truck intermodal transportation system.

Protection and fortification planning The intricacies of a typical intermodal infrastructure, the interdependencies among different components, and the prohibitively high costs make this a challenging undertaking (ref. [14]). Fortunately, given the significance, this area has received increased attention from researchers over the past two decades, and the related works can be reviewed under three threads. The *first* thread includes studies focused on complete redesign of the network so that the system is robust to attacks and includes the effort of ref. [15] who augmented the classical p -median and the un-capacitated facility location problems to account for failures of the facilities. In subsequent works, ref. [16] sought to maximize the combination of initial demand coverage and the minimum post-disruption coverage, whereas ref. [17] proposed a mathematical program to design a network capable of performing well in both pre- and post-disruption conditions.

The *second* thread focuses on protecting the existing infrastructure (i.e., fortification), and has witnessed most of the academic effort. The relevant works under this thread can be organized under two sub-threads: *ascertaining criticality*; and *fortification planning*. While the pioneering effort of ref. [18] to ascertain criticality was rooted in military planning, the subsequent works have been more diverse that have investigated: the impact of arc interdiction in a network (ref. [19]); ways to maximize the shortest path between a given O/D pair (ref. [20]); the impact on revenue (ref. [21]); the use of a bi-level approach to identify the critical components of an electrical supply system (ref. [22]); and the impact of interdicting supply and emergency facilities (ref. [23]). Under the second sub-thread focused on fortification planning, ref. [24] made a noticeable effort by extending their median-based interdiction model and then by proposing an implicit enumeration technique to solve the resulting bi-level programs (ref. [13, 25]). More recently, researchers have considered fortification within both a capacitated (ref. [26, 27]), and an un-capacitated setting (ref. [28]). The concept of fortification against worst-case losses was conceptually introduced in ref. [29, 30] that makes use of a bi-level model to represent fortification and interdiction decisions (i.e., defender-attacker framework), and tri-level models to also include system operation as the third decision (i.e.,

defender-attacker-defender framework). Subsequently, these two frameworks have been used in several applications including power grid (ref. [31]), water supply (ref. [32]), railway infrastructure (ref. [33]), impact of flooding on a road network (ref. [34]), allocation of limited resources in decentralized supply systems (ref. [35]), and the intermodal transportation domain (ref. [12]).

Under the *third* thread, probabilities were used to account for uncertainties associated with attacks on facilities (ref. [24]), a probability distribution to estimate the number of facilities attacked (ref. [36]), and a probability of success measure to investigate the vulnerability of a protected facility (ref. [37]). Ref. [38] advocated for an all-hazards approach to simultaneously incorporate the possibility of worst-case and random attacks, while ref. [39] proposed formulations for both random and mixed types of attacks where targets cannot be predicted, i.e., imperfect information between the defender and the attacker. Finally, it is pertinent to mention that some recent studies have attempted to combine the design, protection, and interdiction decisions (ref. [40]).

Metaheuristic applications in fortification Computational complexity of the resulting mathematical programs and the consequent intractability have motivated researchers to develop metaheuristic-based solution methods. So far, there are four peer-reviewed efforts in this direction. Ref. [41] suggested three solution methods to solve the bi-level program designed to study the p -median problem for the planning and protection of critical facilities. Ref. [42] applied a tabu search-based solution methodology to analyze the design of a public service facility network in the presence of possible terrorist attacks. Ref. [43] proposed simulated annealing, variable depth neighborhood search, and the combination of the two to study protection planning within a hierarchical facility location setting. Finally, ref. [44] examined tabu search, genetic algorithm, and simulated annealing and then advocated the use of a hybrid technique combining the strengths of both metaheuristics and exact solution methods to efficiently tackle the problem.

Rail-truck intermodal transportation system Although rail-truck intermodal transportation has been an active research area over the last two decades (ref. [45]), research about disruption is still in its infancy. Within the rail-truck intermodal transportation domain, few attempts are made to address this issue (ref. [12]).

To sum, in this paper, a tri-level defender-attacker-defender (D-A-D) framework will be adopted to incorporate partial protection/ interdiction of intermodal terminals. Then, a metaheuristic-based solution methodology will be proposed.

3 Problem description

In this section, the process by which a typical rail-truck intermodal transportation system works will be briefly delineated which would entail specifying the relevant players in this domain. A typical rail-truck intermodal chain would consist of a highway component performed by trucks (called drayage); and the rail-haul portion connecting intermodal terminals. Within a given network, a rail-track segment between consecutive terminals is called a service leg. The service legs in the network are then connected by one or more train services. Rail-truck intermodal fortification problem in the literature is modeled as an interaction between three players: the network operator; the interdictor; and the intermodal operator (ref. [12, 29, 30]). Figure 1 depicts the hierarchical and sequential decisions among the three players. The network operator, at the highest level, strives to minimize the cost of the worst-case disruption by fortifying a number of intermodal terminals. This in turn is only possible if the network operator knows the worst-case disruption caused by the interdictor. It should be added that, in this paper, unlike the prevalent

approach wherein protected assets are either completely immune to attacks or entirely vulnerable (ref. [12, 13]), the immunity level is proportional to the number of defensive resources invested in each asset. In the middle level, the interdictor attempts to maximize the cost of using the system by attacking/disrupting a limited number of (unfortified) terminals, which is achieved through complete information about the intermodal operator's problem. Finally, in the lowest level and following the interdiction, the intermodal operator makes use of the available resources on the reduced intermodal network (and possibly fewer train services) to meet the customer demand at the minimum cost. This means that, in the network that some terminals are not operational due to the interdictor's attacks, the intermodal operator takes the best combination of available terminals and train services to minimize the cost of meeting the demands.

Hence, the managerial problem is to determine the optimal distribution of finite defensive resources to completely fortify a certain number of most important rail intermodal terminals, and/or partially fortify some of them and leave others vulnerable such that the functionality of the network

Fig. 1 Hierarchical structure for fortification planning

is preserved as much as possible in the event of the worst-case disruption.

4 Optimization program

In this section, first, the modeling assumptions will be outlined, and then, the optimization program will be presented. The following modeling assumptions are also made: *first*, each terminal can be either completely or partially fortified, depending on how much resources are being invested; *second*, a partially fortified terminal can be interdicted, and the resulting loss in capacity is inversely proportional to the amount of defensive resources invested; *third*, the network operator has finite budget for fortification; *fourth*, each terminal has finite traffic handling capacity; *fifth*, delivery dates are specified when placing the order, and a penalty cost per container per hour is incurred for late deliveries; *sixth*, there is no congestion at the terminals; *seventh*, if an intermediate terminal associated with an intermodal train service is interdicted, the train can still serve the remaining terminals on its route; *eight*, an interdicted terminal cannot be used as either origin or destination for any shipment; *ninth*, in an effort to ensure feasible solutions (i.e., demand is satisfied), direct trucking is permitted between each shipper–receiver pair; and *tenth*, the interdictor intends to impose the worst-case attack on the system. The next section provides the notations used to develop the model.

Sets

- I Set of shippers, indexed by i .
- J Set of receivers, indexed by j .
- P_{ij} Set of intermodal paths between shipper i and receiver j , indexed by p .
- K Set of intermodal terminals in the network, indexed by k .
- P_{ij}^k Set of intermodal paths between shipper i and receiver j which uses intermodal terminal k as either origin or destination.
- V Set of intermodal train services defined on the network, indexed by v .
- L^v Set of service legs for train service v , indexed by l .
- $S^{l,v}$ Set of intermodal paths using service leg l of train service v .

Variables

- X_{ij}^p Number of containers using intermodal path p between shipper i and receiver j .

- XT_{ij} Number of containers using direct trucking service between shipper i and receiver j .
- F_v Number of train service of type v .
- Z_k The protection level of terminal k .

$$Y_k = \begin{cases} 1 & \text{if terminal } k \text{ is interdicted} \\ 0 & \text{otherwise} \end{cases}$$

Parameters

- R Maximum number of terminals that the interdictor can disrupt.
- C_{ij}^p Cost of transporting a container from shipper i to receiver j on intermodal path p .
- CT_{ij} Cost of sending a container using trucks on the shortest path from shipper i to receiver j . Cost of sending a container using trucks on the shortest path from shipper i to receiver j .
- T_{ij}^p Expected travel time from shipper i to receiver j on intermodal path p .
- T_{ij} Delivery time using trucks on the shortest path from shipper i to receiver j .
- $\bar{T}_{ij}$ Delivery due date promised by shipper i to receiver j .
- D_{ij} Number of containers demanded by receiver j from shipper i .
- PC_{ij} Penalty cost per container per hour between shipper i and receiver j .
- α_v Capacity of train service v .
- CO_v Fixed cost of operating train service v .
- U_k Capacity of intermodal terminal k .
- B_k The maximum protection level in terminal k .
- B The maximum protection budget in the network.

(P)

$$\min_Z [C(z)] \quad (1)$$

subject to:

$$\sum_{k \in K} Z_k \leq B \quad \forall k \in K \quad (2)$$

$$0 \leq Z_k \leq B_k \quad \forall k \in K \quad (3)$$

where,

$$C(z) = \max_Y [C(Y)] \quad (4)$$

subject to:

$$\sum_{k \in K} Y_k \leq R \quad (5)$$

$$Y_k + Z_k \leq B_k \quad \forall k \in K \quad (6)$$

$$Y_k \in \{0, 1\} \quad \forall k \in K \quad (7)$$

where,

$$C(y) = \min \left[\sum_{i \in I} \sum_{j \in J} \sum_{\substack{p \in P_{ij} \\ T_{ij}^p > \bar{T}_{ij}}} (T_{ij}^p - \bar{T}_{ij}) PC_{ij} X_{ij}^p + \sum_{i \in I} \sum_{\substack{j \in J \\ T_{ij} > \bar{T}_{ij}}} (T_{ij} - \bar{T}_{ij}) PC_{ij} XT_{ij} + \sum_v CO_v F_v \right] \quad (8)$$

subject to:

$$\sum_{p \in P_{ij}} X_{ij}^p + XT_{ij} \geq D_{ij} \quad \forall i \in I, \forall j \in J \quad (9)$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{p \in P_{ij} \cap P_{ij}^k} X_{ij}^p \leq U_k Y_k \left(\frac{Z_k}{B_k} \right) + (1 - Y_k) U_k \quad \forall k \in K \quad (10)$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{p \in P_{ij} \cap S^{l,v}} X_{ij}^p \leq \alpha_v F_v \quad \forall v \in V, l \in L^v \quad (11)$$

$$N^v \geq 0, \in Z_+ \quad \forall v \in V \quad (12)$$

$$X_{ij}^p \geq 0, \in Z_+ \quad \forall i \in I, \forall j \in J, \forall p \in P_{ij} \quad (13)$$

$$XT_{ij} \geq 0, \in Z_+ \quad \forall i \in I, \forall j \in J \quad (14)$$

(**P**) represents the tri-level model that could be used to make protection planning decisions. The *outer* level problem belongs to the network operator whose objective is to minimize the total cost by (completely or partially) fortifying a certain number of intermodal terminals. Constraint set Eq. (2) describes the total budget available for fortification, while Eq. (3) ensures that defensive budget at each intermodal terminal is nonnegative and that it does not exceed the maximum protection level denoted by B_k . The *middle* level problem belongs to the interdicator who intends to maximize the total cost of using the system. Constraint set Eq. (5) expresses the finite resources available for interdiction of intermodal terminals, whereas Eq. (7) represents the binary nature of the interdiction decisions. Constraint set Eq. (6) ensures that only the terminals that are not fortified to the maximum level could be disrupted. Finally, the *inner* level problem belongs to the intermodal operator who intends to minimize the total cost of using the system. Note that the

objective function includes the overall cost of moving shipments using the rail-truck intermodal system, any direct trucking service, the penalty cost for late deliveries, and the fixed cost of running different intermodal trains in the network. Constraint set Eq. (9) ensures that

demand is satisfied either using the intermodal option or through the direct trucking service. Constraint set Eq. (10) enforces the capacity restrictions at various terminals in the network such that, and for the interdicted terminals, the maximum available capacity depends on the ratio of the defensive resources invested to the amount needed to attain maximum level of fortification (i.e., Z_k/B_k). Constraint set Eq. (11) determines the number of intermodal trains of a specific type needed in the network. Finally, the sign and integrality restrictions are imposed through constraint sets according to Eqs. (12)–(14).

5 Solution methodology

The model (**P**) has a tri-level formulation, which not only renders it very complex but also difficult to solve. It is important to note that the optimal solution to the network operator's problem (i.e., *outer* level) is combinatorial in nature since all possible fortification combinations, for a given budget, needs to be evaluated, which is computationally intensive (ref. [12]). The extant literature suggests employing an implicit enumeration scheme where the network operator makes a binary decision about either completely fortifying a terminal or leaving it totally vulnerable (ref. [13]), and this information is used to solve the interdicator's problem, i.e., *middle* and *lower* level (ref. [12]). However, partial protection of terminals in Model (**P**) necessitated using general integer variables and not binary variables to model the fortification, which in turn also made the implicit enumeration technique inapplicable. Hence, a methodology is proposed to solve Model (**P**) efficiently. More specifically, a tabu search-based algorithm is proposed to find the best way to distribute finite defensive resources among terminals so that the cost imposed by the subsequent worst-case disruption is minimized, i.e., *outer* level problem. Note that, for a specific allocation of

defensive resources, the cost associated with the worst-case disruption is determined by solving the interdicator's problem, i.e., *middle* and *lower* level.

approximation of the bi-level interdicator's problem. This approximation of the actual interdicator's problem is called MP. The objective function of MP could be obtained as follows:

$$\max_Y \left[\max_{\gamma, \lambda, \theta, e, d} \left[\sum_{i \in I} \sum_{j \in J} (D_{ij} * \gamma_{ij}) + \sum_{k \in K} (U_k * \lambda_k) + \sum_{v \in V, l \in L^v} (\theta_l * 0) + \sum_{v \in V} (\bar{F}_v * e_v) + \sum_{k \in K} (\bar{Y}_k * d_k) \right] \right] \rightarrow \text{yields}$$

$$\max_{\gamma, \lambda, \theta, e, d} \left[\left(\sum_{i \in I} \sum_{j \in J} (D_{ij} * \gamma_{ij}) \right) + \sum_{k \in K} (U_k * \lambda_k) + \sum_{v \in V} (\bar{F}_v * e_v) + \sum_{k \in K} (Y_k * d_k) \right]$$

5.1 Solving interdicator's problem

The interdicator's problem is inherently a bi-level problem (ref. [46]) which is known to be computationally challenging. Thus, a Bender's decomposition methodology (ref. [47]) is proposed to efficiently tackle the interdicator's problem by breaking it into a master problem and a sub-problem. The sub-problem (SP), i.e., the intermodal operator's problem, seeks to find the least costly routing plan for shipments. The master problem (MP), on the other hand, is a single-level approximation of the interdicator's problem and is built based on the unimodularity of the intermodal operator's problem. The following steps are required to build MP: *first*, SP is solved, and the optimum values of train frequency variables are obtained; *second*, the optimal train frequency variables enter the middle level problem as constraints; *third*, the interdiction variables are treated as fixed values in the intermodal operator's problem; and *fourth*, three sets of dual variables corresponding to constraints Eqs. (15)–(17) are introduced. Constraints Eqs. (18) and (19) pertain to the train frequency and interdiction variables; respectively, and thus, their dual variables are also needed.

$$\sum_{p \in P_{ij}} X_{ij}^p + XT_{ij} \geq D_{ij} \forall i \in I, \forall j \in J \quad (\gamma_{ij} \geq 0) \quad (15)$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{p \in P_{ij} \cap P_{ij}^k} X_{ij}^p \leq U_k Y_k \left(\frac{Z_k}{B_k} \right) + (1 - Y_k) U_k \forall k \in K \quad (\lambda_k \leq 0) \quad (16)$$

$$\sum_{i \in I} \sum_{j \in J} \sum_{p \in P_{ij} \cap S^{l,v}} X_{ij}^p - \alpha_v F_v \leq 0 \forall v \in V, l \in L^v \quad (\theta_l \leq 0) \quad (17)$$

$$F_v \geq \bar{F}_v \forall v \in V \quad (e_v \geq 0) \quad (18)$$

$$Y_k \geq \bar{Y}_k \forall k \in K \quad (d_k \geq 0) \quad (19)$$

Having defined the above dual variables, the dual of the inner level model will result in a single-level

Since the above objective function contains nonlinear terms, a suitable linearization scheme is needed to resolve the nonlinearity.

Linearization scheme In the above objective function, there are nonlinear terms in the following form that must be linearized:

$$Y_k * d_k = g_k$$

In this nonlinear term, when Y_k is 0, g_k is 0 and when Y_k is 1, it becomes d_k . To resolve the nonlinearity, g_k should replace the nonlinear term $Y_k * d_k$ in the objective function and the following three constraints should be added to make sure that g_k always obtains the right value.

$$g_k \geq 0$$

$$g_k \leq M Y_k$$

$$g_k \leq d_k + M(1 - Y_k)$$

The following notation and Fig. 2, help describe how the decomposition scheme utilizes SP and MP to solve the attacker's problem.

Notation for the decomposition scheme

SP : Sub-problem

MS : Master problem

UB : Upper bound of the decomposition algorithm

LB : Lower bound of the decomposition algorithm

l : Index of the current iteration

Y_l : The solution of the master problem (i.e., list of interdicted terminals) at iteration l

X_l : The solution of the sub-problem at iteration l

The procedure starts by setting up lower and upper bounds (i.e., the lower bound is set to $-\infty$ and the upper bound is set to $+\infty$), the index of iteration is set at 1 and the initial list of interdictions is empty ($Y_1 = \emptyset$). Then, SP is solved, and its optimal solution is found and the upper bound is updated ($UB \leftarrow OFV(X^*)$). If the termination

Fig. 2 Flowchart of the decomposition algorithm

condition is not met, then MP is solved and its optimal solution (Y_1^*) is found and the lower bound is updated (i.e., $LB \leftarrow OFV(Y_1^*)$). Then, the index of iteration is incremented and if the termination condition is not met, the next iteration will be started. The algorithm stops if the iteration number reaches its maximum or the upper bound and the lower bound become close enough: i.e., $UB - LB \leq \epsilon LB$.

5.2 Genetic algorithm for the outer level problem

Once the interdicator's problem is solved and the cost associated with each defense strategy is known, two metaheuristic approaches will be developed to guide the search and to find the best way to distribute the defensive resources. As alluded, finding the best distribution of the defensive resources among terminals is a computationally challenging task, and thus, two metaheuristic approaches are proposed to tackle the problem. First, in this section a genetic algorithm adapted to solve model P will be elaborated and in the following section (Sect. 5.3), a tabu search-based algorithm will be presented.

Introduced by Holland (ref. [48]), genetic algorithm as inspired by the theory of evolution from Charles Darwin uses general ideas like breeding, mutation, and selection over several generations to produce a population with better fitness over long time. The algorithm tries to mimic the natural phenomenon called survival of the fittest where individuals that carry good genes will have a higher chance to survive. In this section, first, notation and different components of the genetic algorithm will

be explained, and then, the related details of implementation will be presented in Fig. 3.

Notation and explanation for genetic algorithm scheme

Φ : Initial population of random solutions

Φ^+ : Extended population of random solutions

m : Number of chromosomes created through the crossover operation

m' : Number of chromosomes mutated through the mutation operation

Figure 3 provides the details of the proposed genetic algorithm. The GA starts with a finite population of random individuals (Φ) represented by their chromosomes. Chromosomes represent the traits of each individual and a fitness function can be used to evaluate the goodness of each individual. Then, through crossover procedure, m random pairs of individuals within the set Φ are selected, and for each pair, parts of the chromosomes are exchanged to create a new individual. The m new individuals are added to the set of individuals to create an extended set of individuals (Φ^+). Then, through mutation procedure, m' individuals are selected to suffer random changes which would add to the diversity of the individuals in the set Φ^+ . At the end, through gene selection procedure, all individuals in the extended set Φ^+ will be evaluated and the best ones will be selected to form the initial population Φ for the next generation of the algorithm. It should be added that, when implementing the genetic algorithm, each feasible solution of the optimization problem is represented by a chromosome of an individual within the population

Generate n random chromosomes to form the population set Φ

While (termination condition is not met) **do**

Create m new chromosomes by applying the crossover operation to randomly selected individuals and amend them to set Φ to form set Φ^+

Mutate m' randomly selected chromosomes within set Φ^+ by applying the mutation procedure

Apply biased roulette wheel procedure to all chromosomes within set Φ^+ to select n chromosomes and set it as set Φ

End (While)

Return the best chromosome within set Φ as the final output of the genetic algorithm

Fig. 3 Pseudocode of the genetic algorithm

of individuals. The search process continues until the termination condition (such as the number of interactions or the time spent) is met. Then, the best solution found in the population will be reported as the output of the algorithm.

Encoding Each chromosome represents a feasible solution as a vector of length n , where n is the total number of terminals in the network. A given cell k of the vector contains up to B_k defensive units such that the total defensive units in the vector adds up to the total defensive budget B . For instance, in a network with 7 terminals, total defensive budget of 16 units, and where complete fortification is possible with 4 defensive units, the following vector shows a feasible solution. Thus, the fourth terminal is completely vulnerable, whereas second and fifth are completely fortified.

3	4	2	0	4	1	2
---	---	---	---	---	---	---

Crossover Procedure The crossover procedure creates new chromosomes using parts of the two other chromosomes (parents) in the population. In this research, the crossover procedure is applied such that the new chromosome retains feasibility. Assuming that parents have length of n each, two parent chromosomes will be selected randomly. Then, in the new chromosome, the first $n/2$ cells are copied from one of the selected chromosomes and the remaining cells are copied from the other parent chromosome. If in the new chromosome, the total defense resource used adds up to more than the total available

defense budget, then a corrective procedure is called that randomly selects cells in the chromosome and reduces the defensive resource in each cell such that the total defensive resource used equals the available defense budget. On the other hand, if the total defensive resources applied adds up to less than the available defense resources, then several cells will be selected randomly to receive additional resources such that the total available defensive resources will be utilized.

Mutation Procedure As in real life, the chromosomes in each population could experience a random mutation. The parameters that are normally used to control the mutation procedure are: (1) population mutation rate which specifies the percentage of the population that will go through the mutation, and (2) chromosome mutation rate which clarifies the percentage of a change in each selected chromosome. The mutation procedure begins by randomly selecting chromosomes from the population according to the population mutation rate. Then, within the selected chromosomes and based on the chromosome mutation rate, a few genes will be selected randomly to experience the change. The mutated chromosomes will then join the population such the population size does not change. It should be noted that the mutation procedure is performed such that it creates feasible chromosomes.

Gene Selection Procedure To determine which chromosomes will survive a generation, a biased roulette wheel selection procedure (as described in ref. [49]) is used. A chromosome will have a chance to survive proportional

Step 1: Evaluate the fitness of each chromosome i as $f_i = \frac{1}{OFV_i}$ where OFV_i refers to the objective function value associated with chromosome i

Step 2: Compute the probability, p_i , of selecting each one of the chromosomes in the population: $p_i = \frac{f_i}{\sum_{j=1}^n f_j}$

Step 3: Compute the cumulative probability, P_i , for each chromosome as: $P_i = \sum_{j=1}^i p_j$

Step 4: Generate a random number $r \in (0,1]$

Step 5:

If $(r < P_1)$ then, select chromosome 1

Else

If $(P_{i-1} < r < P_i)$, select chromosome i

End

Step 6: Repeat steps 4 and 5 until required number of chromosomes are selected

Fig. 4 Pseudocode of the biased roulette wheel approach

to its fitness value. In this problem, since having a lower objective function value is favorable, the fitness of a given chromosome is equivalent to the inverse of its objective function value. In the biased roulette wheel, chromosomes with higher fitness values will have a bigger slot size than the ones with lower values of fitness. By randomly spinning the roulette wheel, the chromosomes will be selected one-by-one to contribute to the new population in the end of each iteration of the genetic algorithm. The details of the biased roulette wheel approach are described in Fig. 4.

5.3 Tabu search for the outer level problem

Tabu search algorithm is believed to expedite the search process due to its demonstrated success on problems involving dispatching n objects into k sets within a constrained setting, and some recent applications in interdiction-fortification. The reader is invited to refer to ref. [50–53] for successful applications. Different steps of tabu search adapted for Model **P** are outlined below together with the details of the execution of the algorithm.

Encoding When implementing tabu search algorithm, a given feasible solution is represented in the same way that it can be represented for the genetic algorithm. **Neighborhood**: In an effort to thoroughly explore the neighborhood of a feasible solution, the neighboring solutions are generated by subtracting a unit of defensive budget from a given cell and adding it to another cell in the vector. Therefore, two neighboring solutions have identical vectors, except for two cells. For instance, the following solutions are considered neighbors since the second vector is obtained by subtracting one unit of defensive resource from the second cell of the first vector and adding it to the third cell of the second vector.

3	4	2	0	4	1	2
3	3	3	0	4	1	2

Tabu tenure The complete exploration of the neighborhood of the current solution constitutes an iteration. The iteration at which a move is made is noted, and a condition is added to forbid the reverse move for four iterations of the algorithm. It should be clear that when a swap is happening, the most recent swap displaces the oldest swap on the tabu tenure list, and that the latter becomes available for possible transformation. Hence, one could think of the proposed tabu tenure technique as a first-in-first-out queue with a capacity to hold four recent moves.

Initial Values

$$S_c \leftarrow S_0$$

$$S_b \leftarrow S_0$$

$$L \leftarrow 0$$

$$T \leftarrow 0$$

Explore the neighborhood of S_c and **Find** S_{bn}

IF $OFV(S_{bn}) < OFV(S_b)$

$$S_b \leftarrow S_{bn}$$

$$L \leftarrow L + 1$$

$$T \leftarrow 0$$

Else

$$T \leftarrow T + 1$$

END (IF)

IF $(T \geq T_{\text{Max}}) \text{ OR } (L \geq L_{\text{Max}})$

Report S_b and $OFV(S_b)$

Exit

Else

$$S_c \leftarrow S_{bn}$$

Go to Explore

END (IF)

Fig. 5 Pseudocode of the tabu search algorithm

Stopping condition The algorithm stops either when the number of iterations reaches 10, or if there is no improvement in the solution in 4 consecutive iterations.

Two distinct initial solutions are experimented with in this study: *first*, the available defensive resources are distributed randomly among assets. *Second*, the bi-level problem of the interdictor is solved, and then, the defensive resources are distributed so that the critical assets are defended as much as possible.

Notation and explanation for tabu search scheme

S_0 : Initial solution

S_c : The current solution

S_{bn} : The best solution found in the neighborhood of the current solution

S_b : The best solution found so far

L : Index of the number of iterations

T : Index of the number of iterations without improvement

L_{Max} : Maximum number of iterations

T_{Max} : Maximum number of iterations without improvement

Figure 5 provides a scheme of the proposed tabu search algorithm that starts with an initial solution (S_0). The current solution is set at the initial solution, and then, the neighborhood of the current solution is explored. The best solution found in the neighborhood of the current solution (i.e., S_{bn}) is then compared with the best solution found so far (S_b). If it is better than the best solution found so far then: the best solution will be updated; the index for the

Fig. 6 **a** Shipper/receiver locations. **b** Intermodal terminals and trains

number of iterations will be incremented; and the index of the number of iterations without improvement will be set at zero. If the best solution found in the neighborhood of the current solution is worse than the best solution found so far, then the index of the number of iterations without improvement will be incremented. Then, the termination condition is checked and if it is not met, the best solution found in the neighborhood will be the current solution of the next iteration. The search procedure stops when either the number of iterations or the number of iterations without improvement reach to their maximum levels. When this happens, the best solution found so far (S_b) and its objective function value will be reported.

6 Numerical analysis

In this section, the details of the case study including estimation of relevant parameters will be provided. Then, to highlight the efficiency of the proposed solution techniques, solutions obtained by using the existing techniques are also reported.

6.1 Case study setting

Figure 6 depicts the intermodal infrastructure of a Class I railroad operator in the United States and represents the area of interest for the managerial problem in this paper. This figure was created via a geographical information system (GIS) using ArcView (ref. [54]) and represents a portion of the network developed in ref. [55]. A total of 37 shipper/receiver locations have access to 18 intermodal terminals, which are connected by 62 types of intermodal train services. These train services are defined by routes they serve and intermediate stops. In the network under study, 31 trains of *regular* type, and another 31 of *express* type (that is 25% faster) are defined. Finally, for expositional reasons and also to facilitate discussions in the sections to follow, a legend for the intermodal terminals is provided in Table 1.

Table 1 Terminal legends

Terminal	Legend	Terminal	Legend	Terminal	Legend
Chicago	<i>Chi</i>	Columbus	<i>Col</i>	Norfolk	<i>Nor</i>
Fort Wayne	<i>FoW</i>	Pittsburg	<i>Pit</i>	Knoxville	<i>Kno</i>
Detroit	<i>Det</i>	Philadel- phia	<i>Phi</i>	Charlotte	<i>Cha</i>
Cleveland	<i>Cle</i>	Cincinnati	<i>Cin</i>	Atlanta	<i>Atl</i>
New York	<i>NY</i>	Roanoke	<i>Roa</i>	Macon	<i>Mac</i>
Indianapo- lis	<i>Ind</i>	Richmond	<i>Ric</i>	Jackson- ville	<i>Jac</i>

Cost: The cost of the drayage operation is proportional to the amount of time the truck driver is involved, and an conservative estimate of \$300/hour (including the estimated hourly fuel cost) is used with a penalty cost of \$40 per hour per container. It should be noted that in the United States the maximum travel speed of trucks is 50 miles/hour, but an average speed of 40 miles/hour is assumed in this study to account for the impact of lights and the traffic (ref. [55]). The average intermodal train speeds for regular and express services were estimated to be 27.7 miles/hour for regular service, and 36.8 miles/hour for express service (according to the Railroad Performance Measure website (ref. [56])) with an estimated variable cost of \$0.875/mile for regular and \$1.164/mile for express service in line with the published works. The fixed cost of running a regular intermodal train is estimated \$500/hour, which takes into consideration the hourly rates for a driver, an engineer, a brakeman, and an engine, which are \$100, \$100, \$100, and \$200, respectively. Finally, it is assumed that the express service comes with an extra 50% charge at \$750/hour (ref. [55]).

Due Dates: The travel times for each shipper–receiver pair (measured in hours) in this study was computed by dividing their distance by 40 (the average speed of trucks), while the distance (measured in miles) was estimated using ArcView GIS (ref. [54]). A constant of 15 was then added to the estimated travel time to obtain a more realistic delivery due date for each shipper–receiver pair.

Demand Level and Terminal Capacity: The intermodal operator's problem was solved in CPLEX 12.6.0 (ref. [57]) on the dataset introduced in ref. [55] to estimate the traffic volume passing through each intermodal terminal. Then, it was assumed that the terminal utilization rate was set at 80%, and thus, the terminal capacity becomes 1.25 times (i.e., 1 divided by 0.8) the traffic volume in each terminal. The travel demand for each shipper–receiver pair is based on the dataset used in ref. [55]. Finally, the optimality gap, ϵ , within the decomposition algorithm is set to 1%, and the maximum number of iterations is fixed at 10.

7 Solution and analyses

In this subsection, first, the tri-level problem (**P**) is solved using genetic algorithm and tabu search algorithm, both implemented in C# and solved using CPLEX 12.6.0 concert technology on a PC with Core 2 Quad, 2.4 GHz processor with 4 GB of RAM. For the tabu search algorithm, two variants are considered as follows:

- **Variant 1** Under this scheme, tabu search algorithm starts from a randomly generated solution where defensive resources are distributed randomly among

Table 2 Comparing the two variants of tabu search algorithm

Defense/attack budget		Best OFV (\$ mns)			Std. Dev. (\$ mns)		
		3	5	7	3	5	7
Genetic algorithm	10	14.03	14.30	14.89	0.332	0.754	0.760
	30	12.01	12.68	12.22	0.642	0.349	0.475
	50	11.33	11.55	11.61	0.454	0.178	0.334
	70	10.99	10.99	11.00	0.023	0.011	0.007
Tabu variant 1	10	12.78	13.63	13.89	0.217	0.148	0.101
	30	12.12	12.05	12.45	0.520	0.306	0.299
	50	11.31	11.29	11.39	0.078	0.041	0.039
	70	10.99	10.99	10.99	0.000	0.000	0.000
Tabu variant 2	10	12.60	13.13	13.59	0.122	0.088	0.077
	30	11.84	11.85	12.21	0.220	0.125	0.118
	50	11.19	11.22	11.38	0.061	0.035	0.039
	70	10.99	10.99	10.99	0.000	0.000	0.000

the intermodal terminals. In an effort to intensively search the neighborhood of any given solution, a 1-swap neighborhood scheme is adopted as delineated in Sect. 5.3.

- *Variant 2* In here, tabu search intelligently exploits the solution of the interdicator's problem. More specifically, the interdicator's problem in the absence of any fortification is solved, which will yield the worst-case attack and facilitate finding the critical terminals. Subsequently, the defensive resources are distributed among the critical terminals to hinder the worst-case attack as much as possible. In case of a tie, it will be broken by random assignment of resources. Finally, a 1-swap neighborhood search is implemented to generate neighboring solutions.

Table 2 summarizes the snapshot of the best solution and the standard deviation of all the solutions encountered within 1 h of computation time for each of the genetic and tabu search algorithms. It should be added that standard deviation is used to account for the impact of randomness (existed in obtaining the initial solution, or when breaking ties) in obtaining the optimal solutions. Each problem instance can be identified by a combination of defense budget and attack budget. For expositional reasons, and without losing generality, three levels of attack budget (i.e., 3, 5, and 7 terminals), and four levels of defense budget (i.e., 10, 30, 50, and 70 units of resource) are considered.

It is evident from Table 2 that for fixed attack budget, for both tabu search and genetic algorithm, the objective function value (OFV) decreases with increase in the defense budget. This is rather intuitive because while the attack budget is fixed, when going down a column, more defensive resources are available to protect the network and thus more costly subsequent attacks will be hindered.

On the other hand, when the defense budget is fixed (i.e., any given row), the OFV increases across different columns since the interdicator has more resources to disrupt the system thereby increasing the resulting cost. In all combinations of attack and defense budget, tabu search variants outperform the genetic algorithm (they end up having lower OFVs for each and every combination of attack and defense budget). Also, tabu search variants have smaller values of standard deviation compared to the genetic algorithm. The fact that genetic algorithm underperforms tabu search variants and has higher standard deviations can be explained by noting that during the creation of new solutions (through crossover procedure) sudden and huge changes are introduced into the solutions. This will, in turn, slow down convergence to the optimal solution as sudden changes might end up deteriorating quality of solutions. This also is the reason that solutions obtained through the genetic algorithm show more volatility and are accompanied with higher standard deviations. On the other hand, tabu search solutions experience slow and smooth changes throughout the search process as they build up by gradually (through 1-swap or 2-swap moves) and thus have higher chances to retain the quality of good solutions found during the search process and use them to converge to the optimal solution. Note that, among the two variants of tabu search, *Variant 2* is outperforming *Variant 1* for all combinations, and this can be attributed to the more intelligent starting, i.e., it starts by hedging against the worst-case attack as the initial solution. Furthermore, the relative consistency of *Variant 2* is also evident in the much lower variability reflected in the standard deviations of the encountered solutions. Finally, it is interesting to note that for both variants, variability continues to increase until peaking at a defense budget of 30 and then decreases to zero for a defense budget of 70. This can

Table 3 Traditional v/s proposed approaches to fortification

	Defense/attack budget	Best OFV (\$ mns)			CPU time ('000 s)			Feasible region
		3	5	7	3	5	7	
Traditional approach	5	12.22	12.82	13.15	0.247	1.206	2.333	8.57×10^3
	6	11.88	12.35	12.66	0.816	2.311	5.919	1.86×10^4
	7	11.58	11.95	12.24	1.107	4.661	13.30	3.18×10^4
	8	11.26	11.58	11.79	1.415	8.499	29.73	4.38×10^3
	9	11.18	11.41	11.52	1.796	14.75	46.80	4.86×10^4
Proposed approach	20	12.08	12.25	12.27	3.6			6.99×10^9
	24	11.85	11.93	12.16				3.49×10^{10}
	28	11.51	11.70	11.64				1.06×10^{11}
	32	11.49	11.60	11.51				2.03×10^{11}
	36	11.22	11.33	11.34				2.51×10^{11}

Fig. 7 Performance of Variant 1 and Variant 2

be explained by considering the fact that the size of the feasible region is in its minimum with the defense budget of 70 (as the system is almost fully protected) and that the variability for both variants of the algorithm is conversely proportional to size of the feasible region (more discussion on the size of the feasible region this will be provided in Table 3). While tabu search solutions outperform solutions obtained from the genetic algorithm, in an effort to better highlight the quality of the solutions obtained from tabu search, the OFVs of tabu search variants are plotted against the OFV of the rail-truck intermodal transportation system running under normal condition (i.e., without any interdictions). Figure 7 could be used to make three observations. *First*, as reflected in Table 2, Variant 2 outperforms Variant 1 for all combinations. *Second*, the OFV of all the combinations converge to the same value (i.e., \$10.99 million) when the defense budget approaches the maximum level of 70. In other words, a defense budget of 70 is high enough to protect 17 intermodal terminals

completely, leaving most of the attack budget unused since there are not many choices left to interdict unfortified terminals. *Third*, given that the cost of a normally operating system is \$10,997,001 (i.e., the lower bound) and assuming the highest defense budget of 70, the cost imposed by the worst-case attack gets very close to the lower bound exhibiting a gap of $1.55 \times 10^{-5}\%$. This means that, with the abundance of defensive resources, no matter how big the attack budget is, or which tabu search variant is used, the cost imposed converges to the cost of running the system under normal conditions.

7.1 Fortification vis-à-vis existing techniques in literature

As mentioned in the introduction, one of the contributions of this paper is to introduce partial protection as a more effective fortification plan, and advocate against the all-or-nothing approach proposed in the extant literature. To

that end, the results obtained by the existing protection approach using binary defense variables are compared against those obtained through the proposed approach using integer defense variables and partial protection. Table 3 shows the OFV (in \$ millions), and the computation time (thousands of seconds) needed for the two approaches for a given number of attack and defense budget combinations. Note that under the traditional approach, the defense budget is enough to fully protect 5, 6, 7, 8, and 9 terminals, which is equivalent to having 20, 24, 28, 32, and 36 defensive resources (4 units for each terminal) under the proposed approach. For the proposed approach, *Variant 2* is selected as it outperforms the other solution approaches (according to Table 2). Thus, ten

independent runs of each combination were solved using *Variant 2* to account for the randomness in the optimal solutions, and the average OFVs and computation times are reported. For the traditional approach, the exact solution technique as developed in the literature (ref. [12, 13]) was implemented to solve each combination, and the corresponding OFV and computation time are reported. It is clear from Table 3 that, for each of 15 combinations, the partial protection approach results in better (i.e., less costly) protection plans than the traditional approaches, and consequently superior network performance in the wake of worst-case disruptions. Thus, it becomes clear that the results obtained by the proposed partial protection approach could serve as lower bounds for the associated results obtained by the traditional approach. Also, it is possible to conclude that *Variant 2* might yield even better solutions, if allowed to continue beyond one-hour cap. However, this superior performance comes with the cost of having a bigger feasible region. In other words, the proposed approach results in much larger feasible regions than those for the traditional all-or-nothing approach which could potentially hinder its application to large instances. The last column of Table 3 provides a snapshot of the size of feasible regions, while “Appendix” A describes our proposed methodology to calculate the size of the feasible region.

The proposed tabu search algorithm helps the proposed approach get solved in a reasonable computation time, even if it deals with much larger feasible region. To compare the computational performance, note that the tabu search implementation for the proposed approach

Fig. 8 Traditional v/s proposed approaches for attack budget of 7

Fig. 9 Comparing the impact of the two protection strategies

was capped at 3600 s (i.e., 1 h), while the traditional approach was uninterrupted until reaching the optimum solution. Based on Table 3, it is visible that the proposed approach significantly outperforms the traditional approach for larger defense-attack budget combinations since it is able to obtain much lower OFVs in much shorter computation time. However, the traditional approach does better for smaller problem instances.

Figure 8 compares OFVs obtained from the two approaches when the attack budget is fixed at 7 and the defense budget varies from 20 to 36 units. According to this figure, the gap between the OFVs obtained by the two approaches is gradually decreasing with the defense budget meaning that taking the proposed approach would add more values when the defense budget becomes scarce.

Finally, in an effort to attain a better sense of the optimal solutions of the two approaches at the network level, the optimal solutions of the proposed and traditional approaches are decoded when the defense budget is 36 (or equivalent to 9 for the traditional approach) and the attack budget is 3. The decoded solutions are then portrayed through Fig. 9. As indicated earlier, the traditional approach makes use of binary defense variables to either completely protect a terminal (i.e., allocate defense budget of 4) or leave one completely unprotected (i.e., allocate defense budget of 0). On the other hand, the proposed approach permits partial protection of terminals, and thus, each could receive anywhere from 0 to 4 resources. This means that the proposed approach provides a much larger number of defense strategies, which in turn provides higher flexibility for defending the system. Such a flexibility has a direct impact on developing much superior protection strategies, and on lowering the OFV associated with the subsequent attacks. For example, while the traditional approach leaves cities like New York, Cleveland, Detroit, and Knoxville totally vulnerable, the proposed approach is able to provide a minimum level of protection for these four cities. Also, both approaches are in total agreement to leave Richmond undefended. However, in Fort Wayne, Roanoke, Norfolk, and Jacksonville the two approaches have completely different views on how much resources should be spent. The proposed approach is able to distribute the defensive resources in a more effective way which will help reduce the cost imposed by the ensuing worst-case attacks while the computational burden of the proposed approach is remedied by the proposed tabu search.

8 Conclusion

In this research, a budget-constrained partial protection planning problem was proposed to hedge against disruptions in a rail-truck intermodal transportation system. The resulting tri-level mixed-integer program was computationally challenging and necessitated the development of customized solution methodology. To this end, two algorithms based on tabu search and genetic algorithms were proposed to efficiently solve the outer level problem, which was combined with a decomposition-based procedure to tackle the middle and inner level problems. The proposed methodology was used to solve and analyze problem instances based on the realistic infrastructure of a Class I railroad operator. In addition, the computational efficiency of the proposed methodology was highlighted in relation to the existing technique in the literature.

Through numerical analysis, two conclusions can be made. *First*, the proposed solution methodology returns optimal solutions much faster than the existing techniques, and especially for larger numerical instances. *Second*, the partial protection approach provides the decision maker with larger number of possible strategies, which in turn translates into increased flexibility and hence better system performance following possible disruptions.

A number directions could be pursued for future research to expand the current modeling framework. *First*, terminal capacities are assumed to be the input to this study while in fact they could be part of the model (as decision variables). *Second*, the model could be extended to include the design of intermodal train services in the wake of interdictions. *Third*, the current model can be enhanced by modeling the impacts of imperfect information into the decision-making process. For instance, the attacker may not have access to the complete information about the network, and thus, the attacker might resort to random disruption instead of optimal ones. Furthermore, the defender may not be fully aware of the attacker's capabilities (for instance in terms of the exact number of disruptions) in advance and, thus, may have to fortify the network against the expected number of disruptions.

Declarations

Conflict of interest The authors declare that they have no conflict of interest.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless

indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

Appendix A: Finding the size of the feasible regions of the protection planning problem

The feasible region of the proposed approach consists of all solutions for the following problem:

$$X_1 + X_2 + X_3 + \dots + X_m = N, 0 \leq X_i \leq k$$

A given solution is a distribution of N undistinguishable items into m distinguishable groups such that no group can have more than k items. Thus, the problem seeks a distribution that leads to the least costly attack.

When $k = 1$ (i.e. $0 \leq X_i \leq 1$), each variable X_i becomes a binary variable with only two possible values. This is what is happening in the traditional approach with binary defense variables.

The size of the feasible region (i.e. the number of feasible solutions) for general values of N , k , and m can be determined as follows:

Step 1 Find the number of solutions for the general problem (without the cap on the number of items in each group):

$$X_1 + X_2 + X_3 + \dots + X_m = N, 0 \leq X_i \leq N, i = 1, 2, \dots, m \quad (20)$$

The total number of solutions for the general problem will be:

$$\binom{N+m-1}{m-1} \quad (21)$$

In this problem, the restriction on the number of items on each group is disregarded and each group can contain all the items. Therefore, some of the solutions of the general problem are violating the limitation on the maximum number of items in each group and, thus, must be taken out.

Step 2 Find the number of solutions that violate the restrictions on the maximum number of items in each group.

At this step, we need to eliminate the solutions that violate the limitation on the maximum number of items in each

group. To start, we first find the number of solutions in which only one group has more than k items:

$$\begin{aligned} X_1 + X_2 + \dots + X_j + \dots + X_m &= N, \\ 0 \leq X_i &\leq N (i = 1, 2, \dots, m), \quad X_j \geq k + 1 \end{aligned} \quad (22)$$

The number of solutions to this problem is:

$$\binom{m}{1} \binom{N - (k+1) + m - 1}{m-1} \quad (23)$$

Then, we need to subtract the above figure from the total number of solutions:

$$\binom{N+m-1}{m-1} - \binom{m}{1} \binom{N - (k+1) + m - 1}{m-1} \quad (24)$$

The above formula subtracts more solutions that it should. Namely, solutions that violate the limit for two groups have been subtracted twice and therefore must be added back:

$$\begin{aligned} &\binom{N+m-1}{m-1} - \binom{m}{1} \binom{N - (k+1) + m - 1}{m-1} \\ &+ \binom{m}{2} \binom{N - 2 * (k+1) + m - 1}{m-1} \end{aligned} \quad (25)$$

At this stage, we added too much since the solutions that violate the limit for three groups have been added back more than once and must be subtracted:

$$\begin{aligned} &\binom{N+m-1}{m-1} - \binom{m}{1} \binom{N - (k+1) + m - 1}{m-1} \\ &+ \binom{m}{2} \binom{N - 2 * (k+1) + m - 1}{m-1} \\ &- \binom{m}{3} \binom{N - 3 * (k+1) + m - 1}{m-1} \end{aligned} \quad (26)$$

This adjustment (adding back solutions that have been subtracted more than once and subtracting extra solutions that have been added more than once) continues until all the possible solutions of the problem are found.

By generalizing the formula (26), the general formula to find the number of solutions for the Eq. (20) can be summarized as:

$$\sum_{i=0}^{\frac{N}{k+1}} (-1)^i \binom{m}{i} \binom{N - i * (k+1) + m - 1}{m-1} \quad (27)$$

References

1. AAR (2014) Rail Intermodal Keeps America Moving. Association of American Railroads- Policy and Economics Department. April 2014
2. AAR (2014) Rail time indicators. Association of American Railroads- Policy and Economics Department. June 2014
3. Jennings Barton E, Holcomb Mary C (2007) The role and impact of rail truck intermodalism on efficient and effective transportation, Part 1, *Logistics Quarterly*, pp. 26–27 (available at: <http://www.logisticsquarterly.com/issues/13-1/article2.html>) (accessed on 26 September 2014)
4. Nozick L, Morlok E (1997) A model for medium-term operations planning in an intermodal rail-truck service. *Transp Res Part A Policy Pract* 31(2):91–107
5. Bryan J, Weisbrod G, Martland CD (2007) Rail freight as a means of reducing roadway congestion: feasibility considerations for transportation planning. *Transp Res Rec J Transp Res Board* 2008(1):75–83
6. Kim NS, Van Wee B (2014) Toward a better methodology for assessing CO2 emissions for intermodal and truck-only freight systems: a European case study. *Int J Sustain Transp* 8(3):177–201
7. US DHS (2014) Department of Homeland Security, <<http://www.dhs.gov/what-critical-infrastructure>>. Accessed 26 August 2014
8. Snyder LV, Scaparra MP, Daskin MS, Church RL (2006) Planning for disruptions in supply chain networks. *Tutor Oper Res*, pp 234–257
9. START (2018) National Consortium for the Study of Terrorism and Responses to Terrorism. Global Terrorism Database. Retrieved from <<http://www.start.umd.edu/gtd>> Accessed on June 12
10. NCTC (2018) National Counterterrorism Center. <<http://www.nctc.gov>>. Accessed June 12
11. CBC (2020) Retrieved from<<https://www.cbc.ca/news/canada/with-rail-blockades-lifted-effort-begins-to-measure-economic-damage-1.5487874>> Accessed February 2020
12. Sarhadi H, Tulett DM, Verma M (2017) An analytical approach to the protection planning of a rail intermodal terminal network. *Eur J Oper Res* 257(2):511–525
13. Scaparra MP, Church RL (2008) A bilevel mixed-integer program for critical infrastructure protection planning. *Comput Oper Res* 35(6):1905–1923
14. Liberatore F, Scaparra MP, Daskin MS (2012) Hedging against disruptions with ripple effects in location analysis. *Omega* 40(1):21–30
15. Snyder LV, Daskin MS (2005) Reliability models for facility location: the expected failure cost case. *Transp Sci* 39(3):400–416
16. O'Hanley JR, Church RL (2011) Designing robust coverage networks to hedge against worst-case facility losses. *Eur J Oper Res* 209(1):23–36
17. Peng P, Snyder LV, Lim A, Liu Z (2011) Reliable logistics networks design with facility disruptions. *Transp Res Part B Methodol* 45(8):1190–1211
18. Wollmer R (1964) Removing arcs from a network. *Oper Res* 12(6):934–940
19. Wood RK (1993) Deterministic network interdiction. *Math Comput Model* 17(2):1–18
20. Israeli E, Wood RK (2002) Shortest-path network interdiction. *Networks* 40(2):97–111
21. Lim C, Smith JC (2007) Algorithms for discrete and continuous multicommodity flow network interdiction problems. *IIE Trans* 39(1):15–26
22. Salmeron J, Wood K, Baldick R (2004) Analysis of electric grid security under terrorist threat. *IEEE Trans Power Syst* 19:905–912
23. Church RL, Scaparra MP, Middleton RS (2004) Identifying Critical Infrastructure: The Median and Covering Facility Interdiction Problems. *Ann Assoc Am Geogr* 94(3):491–502
24. Church RL, Scaparra MP (2007) Protecting Critical Assets: The r-Interdiction Median Problem with Fortification. *Geogr Anal* 39(2):129–146
25. Scaparra MP, Church RL (2008) An exact solution approach for the interdiction median problem with fortification. *Eur J Oper Res* 189(1):76–92
26. Scaparra MP, Church R (2012) Protecting supply systems to mitigate potential disaster a model to fortify capacitated facilities. *Int Reg Sci Rev* 35(2):188–210
27. Aksen D, Piyade N, Aras N (2010) The budget constrained r-interdiction median problem with capacity expansion. *CEJOR* 18(3):269–291
28. Li Q, Savachkin A (2013) A heuristic approach to the design of fortified distribution networks. *Transp Res Part E Logist Transp Res* 50:138–148
29. Brown G, Carlyle M, Salmeron J, Wood K (2005) Analyzing the vulnerability of critical infrastructure to attack and planning defenses. *Tutor Oper Res Emerg Theory Methods Appl*, pp 102–123
30. Brown G, Carlyle M, Salmeron J, Wood K (2006) Defending critical infrastructure. *Interfaces* 36(6):530–544
31. Alguacil N, Delgadillo A, Arroyo JM (2014) A trilevel programming approach for electric grid defense planning. *Comput Oper Res* 41:282–290
32. Qiao J, Jeong D, Lawley M, Richard JPP, Abraham DM, Yih Y (2007) Allocating security resources to a water supply network. *IIE Trans* 39(1):95–109
33. Starita S, Scaparra MP (2016) Optimizing dynamic investment decisions for railway systems protection. *Eur J Oper Res* 248(2):543–557
34. Starita S, Scaparra MP (2017) A dynamic model for road protection against flooding. *J Oper Res Soc* 68(1):74–88
35. Zhang X, Zheng Z, Cai KY (2018) A fortification model for decentralized supply systems and its solution algorithms. *IEEE Trans Reliab* 67(1):381–400
36. Liberatore F, Scaparra MP, Daskin MS (2011) Analysis of facility protection strategies against an uncertain number of attacks: The stochastic R-interdiction median problem with fortification. *Comput Oper Res* 38(1):357–366
37. Zhu Y, Zheng Z, Zhang X, Cai K (2013) The r-interdiction median problem with probabilistic protection and its solution algorithm. *Comput Oper Res* 40(1):451–462
38. Zhuang J, Bier VM (2007) Balancing terrorism and natural disasters-defensive strategy with endogenous attacker effort. *Oper Res* 55(5):976–991
39. Zhang X, Zheng Z, Zhu Y, Cai KY (2014) Protection issues for supply systems involving random attacks. *Comput Oper Res* 43:137–156
40. Aksen D, Aras N (2012) A bilevel fixed charge location model for facilities under imminent attack. *Comput Oper Res* 39(7):1364–1381
41. Aksen D, Aras N, Piyade N (2013) A bilevel p-median model for the planning and protection of critical facilities. *J Heuristics* 19(2):373–398
42. Kecici S, Aras N, Verter V (2012) Incorporating the threat of terrorist attacks in the design of public service facility networks. *Optim Lett* 6(6):1101–1121
43. Aliakbarian N, Dehghanian F, Salari M (2015) A bi-level programming model for protection of hierarchical facilities under imminent attacks. *Comput Oper Res* 64:210–224

44. Cheng CH, Lai TW, Yang DY, Zhu Y (2016) Metaheuristics for protecting critical components in a service system: A computational study. *Expert Syst Appl* 54:251–264
45. SteadieSeifi M, Dellaert NP, Nuijten W, Van Woensel T, Raoufi R (2014) Multimodal freight transportation planning: A literature review. *Eur J Oper Res* 233(1):1–15
46. Dempe S (2002) Foundations of bilevel programming. Kluwer Academic Publishers, Dordrecht
47. Benders JF (1962) Partitioning procedures for solving mixed-variables programming problems. *Numer Math* 4(1):238–252
48. Holland JH (1992) Adaptation in natural and artificial systems: an introductory analysis with applications to biology, control, and artificial intelligence. MIT press, Cambridge
49. Burke EK, Burke EK, Kendall G, Kendall G (2014) Search methodologies: introductory tutorials in optimization and decision support techniques. Springer, Berlin
50. Glover F (1989) Tabu search –Part I. *ORSA J Comput* 1:190–205
51. Galinier P, Hertz A (2006) A survey of local search methods for graph coloring. *Comput Oper Res* 33(9):2547–2562
52. Cordeau JF, Laporte G (2005) Tabu search heuristics for the vehicle routing problem. Metaheuristic optimization via memory and evolution. Springer, Boston, pp 145–163
53. Zufferey N, Amstutz P, Giaccari P (2008) Graph colouring approaches for a satellite range scheduling problem. *J Sched* 11(4):263–277
54. ESRI (2008) ArcView Geographical Information System: An ESRI product. <http://www.esri.com>
55. Verma M, Verter V, Zufferey N (2012) A bi-objective model for planning and managing rail-truck intermodal transportation of hazardous materials. *Transp Res Part E Logist Transp Rev* 48(1):132–149
56. RPM (2014) Railroad Performance Measures. <<http://www.railroadpm.org>>. Accessed May 20, 2014
57. IBM (2014) CPLEX, Version 12.6.0. <<http://www.ibm.org>>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.