Securing Lifelines: Safe Delivery of Critical Services in Areas with Volatile Security Situation via a Stackelberg Game Approach

Tien Mai[1], **Arunesh Sinha**[2]
[1]Singapore Management University
[2]Rutgers University
atmai@smu.edu.sg, arunesh.sinha@rutgers.edu

Abstract

Vaccine delivery in under-resourced locations with security risks is not just challenging but also life threatening. The COVID pandemic and the need to vaccinate added even more urgency to this issue. Motivated by this problem, we propose a general framework to set-up limited temporary (vaccination) centers that balance physical security and desired (vaccine) service coverage with limited resources. We set-up the problem as a Stackelberg game between the centers operator (defender) and an adversary, where the set of centers is not fixed a priori but is part of the decision output. This results in a mixed combinatorial and continuous optimization problem. As part of our scalable approximation solution, we provide a fundamental contribution by identifying general duality conditions of switching max and min when both discrete and continuous variables are involved. Via detailed experiments, we show that the solution proposed is scalable in practice.

1 Introduction

Vaccine delivery has always been a challenge in under-resourced parts of the world (Zaffran et al. 2013). The problem is further aggravated by the threat of violence against vaccine providers in areas where the security situation is volatile (Gannon, Meldrum, and Keath 2020). A safer vaccine delivery plan for such places can save lives, both via vaccination of the population and physical protection of the front line vaccine providers. Specifically, vaccination drives in underdeveloped areas are often performed by setting up a *limited* number of temporary vaccination centers. Motivated by this issue, we propose and study a *general* framework to set-up temporary centers that balance physical security needs of the centers and achieve desired service coverage.

Our *first contribution* is a flexible model that allows choice of a small subset of centers to operate, along with a consideration of how to allocate security resources to the operational centers. Further, the framework allows for fairness constraints that ensure fairness in center allocation for different geographical regions as well as fairness in security allocation to operating centers. The model is set-up as a Stackelberg game between the centers operator (defender) and a bounded rational (Quantal Response) adversary. The model takes inspiration from Stackelberg security games (Tambe 2011) in the manner in which security resources are allocated and the center selection aspect is inspired by product assortment and pricing problems (Wang 2012). While inspired from these models from disparate areas, we believe this is a *first* model that brings together security issues and subset of centers (targets) selection within one framework.

Our *second contribution* is a hybrid algorithm that combines the best of two different approaches. The first of these two approaches is a Mixed Integer Linear Program (MILP) with guaranteed approximation, and the second is a polynomial time heuristic. All algorithms work by a sequence of modifications to the original optimization problem, starting with a binary search on the objective value. The first algorithm exploits properties of the problem at hand to convert a bi-linear MIP formulation to a MILP with guaranteed approximation, but, MILPs are not scalable in practice. Hence, we design a heuristic where we use Lagrangian duality to reach a sub-problem which is a max-min-max tri-level optimization with discrete variables for the outer max and continuous variables for the two inner min-max problems.

Then, in a *fundamental technical contribution*, we identify general conditions for minimax equality when the variables involved are discrete and continuous. While the conditions do not hold in rare cases for our problem, we use this result as a heuristic to transform our sub-problem to a min-max-max problem and then we present a polytime approximation for the transformed problem. The polynomial time approach provides for immense scalability in practice, with solutions close to the MILP approach in almost all cases. Finally, we show that the MILP and heuristic can be made to work together in a scalable hybrid approach with approximation guarantees for the output solution.

We conduct thorough experiments to analyze various aspects of our three approaches. We show that our main hybrid algorithm is scalable in practice (solving for up to 5000 potential centers within 1.5 minutes) and also much better than competing baseline approaches. While inspired by vaccine centers operation, our model and ideas can be applied to critical services operation of temporary health camps, exam centers, etc. in underdeveloped and security risk-prone areas or even existing security games work where a subset of targets can be chosen to be made unavailable.

2 Model and Problem Formulation

We model the stated problem as a general sum Stackelberg game between a defender and a QR adversary. The defender operates a subset of centers (which changes at set frequency, e.g., weekly) and allocates security resources to operational centers. The adversary's attacks an operational center.

Action spaces. The defender has *a candidate set of potential centers*, denoted by $\mathcal{K}$. The variable S, where $S \subseteq \mathcal{K}$, denotes the defender's choice of *centers to operate*. However, not all subsets are feasible; $F(\mathcal{K}) \subset 2^{\mathcal{K}}$ is the feasible set of sets of centers. Suppose every center can vaccinate at least $P_{\min}$ people. Two natural restrictions are that for every $S \in F(\mathcal{K})$ to have $|S| \leqslant C$ and $P_{\min}|S| \geqslant N_P P_{\min}$ for some integer constants C, N_P, capturing a budget constraint and a minimum number of $N_P P_{\min}$ people to be vaccinated every round. In addition, we consider another natural constraint that the candidate center set $\mathcal{K}$ is partitioned into $\mathcal{K}_1, \ldots, \mathcal{K}_L$ such that any feasible choice $S \in F(\mathcal{K})$ contains at least one location from each partition $l \in \{1, \ldots, L\}$ ($L < C$), that is, $|\mathcal{K}_l \cap S| \geqslant 1$ for all l. The set $\mathcal{K}_l$ contains the possible operable center locations in a contiguous geographic region l; hence, this constraint captures fairness in the allocation of vaccine centers across different regions. We call this *fairness in vaccine center allocation* (FVCA).

Continuing with the defender's action description, the defender also allocates security resources to the centers $S \in F(\mathcal{K})$ that are chosen to be operational. The number of security resources is fixed and denoted by m. The defender's pure strategy is to allocate m security resources to the $|S|$ operational centers. The mixed strategy is represented succinctly by $x_S = \langle x_j \rangle_{j \in S}$ ($x_S \in \mathcal{D}_S^x = [0,1]^{|S|}$), which denote the marginal probability of defending the $|S|$ centers that are chosen to operate. Further, in order to provide a *fairness in security allocation* (FSA), we impose the constraints: $\sum_{j \in \mathcal{K}_l \cap S} x_j \leqslant \beta_l$ for all l and some given constants β_l. In words, these constraints imposes an upper bound β_l on the chance of protecting facilities that operate in region l thereby ensuring that no geographic region is given unusual preference in terms of protection. E.g., we could have $\beta_l \propto \frac{|\mathcal{K}_l|}{|\mathcal{K}|} m$ with the proportionality constant greater than 1.

Overall, the defender's action is (S, x_S) with constraints as stated above. The adversary's pure strategy is to choose one among the $|S|$ operational centers to attack; locations in $\mathcal{K} \backslash S$ have no operational center and cannot be attacked.

Utilities. For every potential center $j \in \mathcal{K}$, if the center is operating and the adversary attacks j and the center is protected then the defender obtains reward r_j^d and the adversary obtains l_j^a. Conversely, if the defender is not protecting the operating center j, then the defender obtains l_j^d ($r_j^d > l_j^d$) and the adversary gets r_j^a ($r_j^a > l_j^a$). Given x_j, the expected utility of the defender and attacker for an attack on an operational center j is as follows: $U_j^d(x_j) = x_j r_j^d + (1 - x_j) l_j^d$ and $U_j^a(x_j) = x_j l_j^a + (1 - x_j) r_j^a$. For ease of notation, we note that these utilities are linear in x_j and rewrite these as

$$U_j^d(x_j) = w_j^d x_j + l_j^d \quad \text{where } w_j^d = r_j^d - l_j^d \geqslant 0$$
$$U_j^a(x_j) = -w_j^a x_j + r_j^a \quad \text{where } w_j^a = r_j^a - l_j^a \geqslant 0$$

These utilities are valid for location $j \in \mathcal{K}$ only when j has an operational center, that is, $j \in S$. Note that S is a decision variable and not fixed apriori.

We use the QR model for the adversary's response. QR is a well-known model (McFadden 1976; McKelvey and Palfrey 1995), and used extensively in Stackelberg security games. Specifically, QR posits that the adversary will attack an operational center $j \in S$ with probability:

$$q_j(x_S; \lambda) = \frac{e^{\lambda(-w_j^a x_j + r_j^a)}}{\sum_{i \in S} e^{\lambda(-w_i^a x_i + r_i^a)}} \tag{1}$$

Parameter $\lambda \geqslant 0$ governs rationality. $\lambda = 0$ means least rational, as the adversary chooses its attack uniformly at random and $\lambda = \infty$ means fully rational (i.e., attacks a center with highest utility). Thus, QR has the flexibility to model a range of behavior. Then, defender's expected utility for (S, x_S) is

$$\mathcal{F}(S, x_S) = \sum_{j \in S} q_j(x_S; \lambda)(w_j^d x_j + l_j^d) \ .$$

Stackelberg equilibrium. The Stackelberg equilibrium (also called Quantal Stackelberg Equilibrium in (Cerny et al. 2020)) can be computed using the following optimization:

$$\max_{S \in F(\mathcal{K}), x_S \in \mathcal{D}_S^x} \mathcal{F}(S, x_S) \tag{EqOPT}$$
$$\text{subject to} \sum_{j \in S} x_j \leqslant m \ , \tag{2}$$
$$\sum_{j \in \mathcal{K}_l \cap S} x_j \leqslant \beta_l \quad \forall l \ , \tag{3}$$

Constraint (2) states that there can be at most m security resources. Constraint (3) captures the FSA fairness criteria. The $S \in F(\mathcal{K})$ in the subscript of $\max$ captures the three constraints $|S| \leqslant C$, $|S| \geqslant N_P$, and FVCA fairness criteria; objective $\mathcal{F}(S, x_S)$ is the expected utility of the defender.

Related Work

Product Assortment and Pricing. Our problem is most closely related to the problem that combines product assortment and price optimization problem in which a subset of products is chosen from an assortment of products and simultaneously the prices of products is chosen with the goal of maximizing profit in a market where the buyers choose a product following the QR (Wang 2012) model. The relation to our problem stems from the analogy of set of products to set of centers and of continuous prices to security allocation. However, prior work (Wang 2012) solves an unconstrained optimization (no constraints on prices), whereas we deal with a constrained optimization.

Other works in this area optimize just the product assortment with fixed prices (Davis, Gallego, and Topaloglu 2014; Désir, Goyal, and Zhang 2020). In other works (still fixed prices), different models of buyers have been considered such as rational best responding buyers (Immorlica et al. 2018) and other discrete choice models (Davis, Gallego, and Topaloglu 2014; Gallego, Ratliff, and Shebalov 2015). Further away, there is work on an online version of the product assortment problem using a multi-armed bandits formulation (Agrawal et al. 2016; Cheung and Simchi-Levi 2017).

Stackelberg Security Games. There is a large body of work in game theoretic models and algorithms for physical security (Tambe 2011; Xu 2016; Fang et al. 2017; Sinha et al. 2018; Cerny et al. 2020; Yang, Ordonez, and Tambe 2012; Haghtalab et al. 2016) of a given fixed set of vulnerable targets, which have also been deployed in real world (Tambe 2011). The *difference from the standard security game* with QR adversary (Fang et al. 2017) is in the aspect that we allow the defender to choose a subset of targets that then are part of the game and other targets are not part of the game; this yields a novel mixed combinatorial and continuous optimization problem. A naive exploration of all subsets is computationally infeasible, and we present novel approaches to address this problem.

Recent work explores the complexity of quantal response players in more generality for Stackelberg games (Cerny et al. 2020; Milec et al. 2021) calling the equilibrium as Quantal Stackelberg Equilbrium (QSE). Our problem differs because of an additional combinatorial dimension (operate a subset of centers) of the defender's action space. Moreover, in contrast to the hardness results in these work, we obtain arbitrarily precise approximation for our problem in polynomial time. Thus, we identify a sub-class of games where the QSE is efficiently approximable to arbitrary precision, even with a complex defender action space. Our technique can be viewed as one in robust optimization (Bertsimas, Brown, and Caramanis 2011) or a Stackelberg game with constrained follower strategy space (Goktas and Greenwald 2021), but a primary difference is that some of our variables are discrete and hence (sub) gradient or other continuous optimization methods do not apply directly.

Facility Location. Our problem might seem like the maximum capture problem in competitive facility location (Benati and Hansen 2002; Freire, Moreno, and Yushimito 2016; Mai and Lodi 2020; Golowich, Narasimhan, and Parkes 2018; Aziz et al. 2020). However, our problem is very different, as we consider security for centers and have no consideration of capturing market share and abstract away from logistic modelling. The interplay between discrete and continuous optimization is an important aspect of our work.

3 Guaranteed Approximate Equilibrium

As stated in the introduction, we present two approaches, one MILP based with solution quality guarantees (in this section) and another based on a polynomial time heuristic in Section 4. We further combine these two approaches into one hybrid approach in Section 5 that provides approximation guarantees for the solution, as well as similar scalability as the heuristic. *All missing proofs are in the appendix*.

Common Binary Search Transformation

We start with a transformation that is used in all our approaches. We use the Dinkelbach transform (Dinkelbach 1967) to convert the fractional objective of EqOPT to a non-fractional one. We use the shorthand notation $q_j(x_S;\lambda) = N(x_j)/D(x_S)$ to express Eq. 1 succinctly (as λ is obvious, it is not stated explicitly). By definition, $D(x_S) = \sum_{j\in S} N(x_j)$. Hence the objective of EqOPT can be written as

$$\sum_{j\in S} \frac{N(x_j)}{D(x_S)}(w_j^d x_j + l_j^d).$$

The Dinkelbach transform works by seeking to find the highest value of a threshold δ such that there exists some feasible S, x_S such that $\sum_{j\in S} \frac{N(x_j)}{D(x_S)}(w_j^d x_j + l_j^d) \geqslant \delta$. The highest possible δ can be computed by a binary search where in each round of the search the feasibility problem stated above is solved for a particular δ_0. For given δ_0, the feasibility problem is easily solved by checking if the maximum of the following optimization is $\geqslant 0$ or not

$$\max_{S\in F(\mathcal{K})} \max_{x_S\in\mathcal{D}_S^x} \sum_{j\in S} N(x_j)(w_j^d x_j + l_j^d) - \delta_0 D(x_S) \quad \text{(BOPT)}$$

subject to Constraints (2-3)

While the above is a common technique, for completeness, the binary search template over δ_0 is shown in Algorithm 1, where BOPT is solved repeatedly (line 4) till convergence. We present different ways of solving BOPT in the sequel for our different approaches. However, all our approach of solving BOPT are approximate, making Algorithm 1 a binary search with inexact function evaluation, which necessitates the following result:

Algorithm 1: *Binary Search Template*

1 $U = \max_{j\in\mathcal{K}} w_j^d + l_j^d, L = \min_{j\in\mathcal{K}} l_j^d, S = \phi$
2 **while** $U - L \geqslant \epsilon$ **do**
3 $\quad \delta_0 = (U + L)/2$
4 $\quad obj =$ Solve BOPT(δ_0) and get obj. value
5 $\quad$ **if** $obj \geqslant 0$ **then** $L = \delta_0$ **else** $U = \delta_0$
6 return δ_0, S, x_S from the last BOPT solution

Theorem 1. *Suppose* BOPT *is computed with additive error of ξ and runtime T in line 4 of Algorithm 1. Let $f(\delta)$ be the optimal value of* BOPT *for δ. We can show that $d|\delta - \delta'| \leqslant |f(\delta) - f(\delta')| \leqslant D|\delta - \delta'|$ for some constants $d, D > 0$ under the assumption that the absolute value of the utility parameters are bounded by constants. Moreover, Algorithm 1 runs in $O(\log(1/\epsilon) \times T)$ time with and additive error of $O(\xi + \epsilon)$.*

In the rest of this paper, all our approaches will focus on approximately solving BOPT (line 4) in Algorithm 1. The additive approximation guarantee of each approach can directly be used as ξ in the above theorem to obtain the overall approximation guarantee of that approach.

MILP Approach via Compact Linearization

The BOPT optimization objective is separable in x_j's. We exploit this and use a piecewise linear approximation (PWLA) to obtain a non-linear integer program. We follow a recipe similar to prior work (Yang, Ordonez, and Tambe 2012), and divide the range of x_j, i.e., $[0, 1]$, into K equal intervals and represent each $x_j = \sum_{k\in[K-1]} r_{jk}$, where $r_{jk} = 1/K$ if $k \leqslant \lfloor Kx_j \rfloor$ and $r_{jk} = x_j - \lfloor Kx_j \rfloor/K$ if $k = \lfloor Kx_j \rfloor + 1$ and $r_{jk} = 0$ otherwise. However, in contrast

to (Yang, Ordonez, and Tambe 2012) we need additional binary variables θ_j ($\theta_j = 1$ if $j \in S$, $\theta_j = 0$ otherwise) to represent center selection, which results in *bi-linear* terms of the form $\theta_j r_{jk}$ in the resultant MIP. This bilinear MIP is shown in the appendix. Bi-linear terms can be linearized using the well-known big-M approach (Wu 1997). But, this naive linearization results in $K|\mathcal{K}|$ additional variables and $3|\mathcal{K}|K$ additional constraints, making such approach not scalable at all. Next, we show that the naive bi-linear MIP can be formulated as a MILP with no additional variables and only $|\mathcal{K}|$ additional constraints, with upper bounds for the approximation error given in the result below.

Theorem 2. BOPT *is approximated by the following MILP*

$$\max_{\theta,r,z} \quad K \sum_{j\in\mathcal{K}} \theta_j (g_j(0) - \delta_0 g_j(0)) \qquad \text{(ApxOPTL)}$$
$$+ \sum_{j\in\mathcal{K}} \sum_{k\in[K]} \left(\gamma^g_{jk} - \delta_0 \gamma^N_{jk}\right) r_{jk}$$

subject to
$$\sum_{j\in\mathcal{K}} \sum_{k\in[K]} r_{jk} \leqslant Km$$
$$\sum_{j\in\mathcal{K}_l} \sum_{k\in[K]} r_{jk} \leqslant K\beta_l, \forall l \in [L]$$
$$\theta_j \geqslant z_{j1}, \ \forall j \in \mathcal{K}$$
$$z_{jk} \geqslant z_{j,k+1}, \ k \in [K-1], j \in \mathcal{K}$$
$$z_{jk}/K \leqslant r_{jk} \leqslant 1/K, \ k \in [K], j \in \mathcal{K}$$
$$r_{jk} \leqslant z_{j,k+1}/K, \ k \in [K-1], j \in \mathcal{K}$$
$$N_P \leqslant \sum_{j\in\mathcal{K}} \theta_j \leqslant C, \ \sum_{j\in\mathcal{K}_l} \theta_j \geqslant 1, \ l \in [L]$$
$$z_{jk}, \theta_j \in \{0,1\}, \ \forall j \in \mathcal{K}, k \in [K].$$

Let $B(S, x_S)$ be the objective function of (BOPT) *and S^*, x^*_S be an optimal solution of* (BOPT). *Let (θ', z', r') be an optimal solution of* (ApxOPTL), *which provides solution S', x'_S such that $S' = \{j|\theta'_j = 1\}$ and $(x'_S)_j = \sum_{k\in[K]} r'_{jk}$. Then, $|B(S', x'_S) - B(S^*, x^*_S)| \leqslant O(\frac{1}{K})$.*

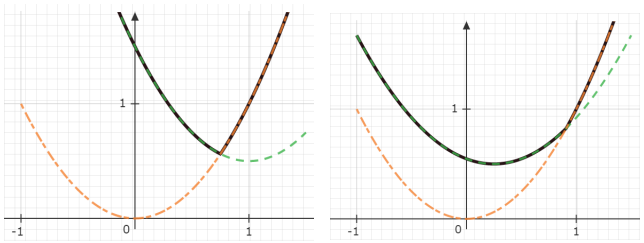
Using Theorem 1 and the above result, it can be shown that when solution of ApxOPTL is used in line 4 of the binary search of Algorithm 1, we attain an approximation of $O(\frac{1}{K} + \epsilon)$. However, the runtime of a MILP in the worst case is not polytime. Next, we propose a heuristic that is polytime but an approximation guarantee does not hold in rare cases.

4 Polynomial Time Heuristic

Our starting point for the heuristic is problem BOPT. We aim to transform the problem further to remove constraints but the inner max problem in x_S is not concave. However, a simple variable transform $y_j = e^{-\lambda w^a_j x_j}$ makes the inner problem, now in y_S, concave with the same optimal objective value. Then, we form the Lagrangian dual of the inner max problem (in y_S) with *guaranteed same solution* due to concavity and consequent strong duality. The optimization *after Lagrangian dualizing* is

$$\max_{S\in F(\mathcal{K})} \min_{\nu,\mu\geqslant 0} \max_{y_S\in\mathcal{D}^y_S} \phi(S, \nu, \mu, y_S, \delta_0) \qquad \text{(DualOPT)}$$

where $\phi(S, \mu, \nu, y_S, \delta_0) =$



(a) $f_1, f_2, \max(f_1, f_2)$ (b) $f'_1, f'_2, \max(f'_1, f'_2)$

Figure 1: Simple illustration of Thm. 3. Horizontal axis is the y-axis. Max functions are in bold. f_1, f'_1 touch the origin.

$$\sum_{j\in S} N(y_j)\Big(\frac{-w^d_j \log y_j}{\lambda w^a_j} + l^d_j\Big) - \delta_0 D(y_S)$$
$$- \nu\Big(\sum_{j\in S} \frac{-\log y_j}{\lambda w^a_j} - m\Big) - \sum_l \mu_l \Big(\sum_{j\in\mathcal{K}_l\cap S} \frac{-\log y_j}{\lambda w^a_j} - \beta_l\Big)$$

and $\mathcal{D}^y_S$ is the Cartesian product $\times_{j\in S}[e^{-\lambda w^a_j}, 1]$, $N(y_j) = y_j e^{\lambda r^a_j}$ and $D(y_S) = \sum_{j\in S} N(y_j)$ and $\nu, \langle\mu_l\rangle_{l\in\{1,\ldots,L\}}$ are the dual variables. We aim to switch the outer max and min in DualOPT; first, we present a general result when such switching is possible.

A General Minimax Equality

We present a general result about conditions for minimax equality when discrete variables are involved. We note that, as far as we know, there is no such result in literature.

Theorem 3. *Let $f : X \times Y \to \mathbb{R}$ be a function such that the following are true:*

- *X is a set with finitely many points.*
- *Y is a convex set in a Euclidean space.*
- *$f(x, \cdot)$ is continuous and convex on Y.*
- *For all $x \in X$ and all b, the sub-level sets of $f(x, \cdot)$ given by $\{y \mid f(x, y) \leqslant b\}$ are compact and convex (if Y is compact, this condition is implied and can be removed)*
- *For any $y^* \in \operatorname{argmin}_{y\in Y} \max_{x\in X} f(x, y)$ there is a unique x^* such that $x^* = \operatorname{argmax}_{x\in X} f(x, y^*)$.*

Then $\min_{y\in Y} \max_{x\in X} f(x, y) = \max_{x\in X} \min_{y\in Y} f(x, y)$

We provide a simple illustration with 1d quadratic functions in Fig. 1 for better intuition of the above result. The exact functions used are stated in the appendix. Fig. 1a shows a situation where there is a non-unique maximizer (both $1, 2$) of $\min_y \max(f_1(y), f_2(y))$ for functions f_1, f_2 (occurring at $y = 0.75$ with value $f_1(0.5) = f_2(0.5) = 0.5625$). Whereas, $\max_{1,2} \min_y(f_1(y), f_2(y))$, which is min of each function separately and then the max from among those, occurs at $y = 1$ with value $0.5 \neq 0.5625$. Fig. 1b shows a situation where $\min_y \max(f'_1(y), f'_2(y))$ is uniquely determined by f'_2 and hence is same as $\max_{1,2} \min_y(f'_1, f'_2)$.

A Polytime Solution for Switched Problem

Continuing with our aim of solving DualOPT, we switch the outer max and min in

DualOPT to obtain SwitchedDualOPT: $\min_{\nu,\mu\geqslant 0}\max_{S\in F(\mathcal{K})}\max_{y_S\in\mathcal{D}}\phi(S,\nu,\mu,y_S,\delta_0)$. This switching is a heuristic as the uniqueness condition in Theorem 3 may not hold for every possible parameter values of the problem at hand (other conditions hold); see also discussion after Lemma 1. In any case, SwitchedDualOPT provides tight bounds, which we use for a hybrid guaranteed approximate solution in the next section.

Solving for fixed duals. We first show a polytime arbitrary close approximation for the inner problem in SwitchedDualOPT. For any given *fixed* $\nu,\mu\geqslant 0$, consider:

$$\max_{S\in F(\mathcal{K})}\ \max_{y_S\in\mathcal{D}_S^y}\ \phi(S,\nu,\mu,y_S,\delta_0) \qquad \text{(FixedDuals)}$$

Observe that the objective $\phi(\cdot)$ of FixedDuals is additively separable into terms $g_j(\nu,\mu,y_j,\delta_0)$ that depend only on the single scalar variables y_j as follows:

$$\phi(S,\nu,\mu,y_S,\delta_0)=\sum_{j\in S}g_j(\nu,\mu,y_j,\delta_0)+\nu m+\sum_l\mu_l\beta_l$$

$$\text{where } g_j(\nu,\mu,y_j,\delta_0)=N(y_j)\big(\frac{-w_j^d\log y_j}{\lambda w_j^a}+l_j^d-\delta_0\big)$$

$$-(\nu+\mu_l)\frac{-\log y_j}{\lambda w_j^a}\text{ for }l\text{ s.t. }j\in\mathcal{K}_l$$

We use this separability to solve FixedDuals to any precision in polytime in Algorithm 2. First, in lines (2-4) we maximize g_j over y_j for each $j\in\mathcal{K}$ to obtain $h_j(\nu,\mu,\delta_0)$. With h_j's, the objective of FixedDuals takes the form $\max_{S\in F(\mathcal{K})}\sum_{j\in S}h_j$ plus constants $\nu m+\sum_l\mu_l\beta_l$. Recalling the definition of $F(\mathcal{K})$, we solve this by sorting the h_j's (line 5) and extracting the best h_j values from each of the L partitions of locations (lines 6-8). This ensures at least one center in each partition is selected. Then, among remaining locations, we choose the best (as measured by h_j) centers (lines 9-12) ensuring at least N_P are chosen. The final choice of centers is S_o (line 13).

Next, we show that there is a closed form formula for h_j in line 4 of Algorithm 2. This is a concave max optimization as g_j is concave in y_j. The closed form formula enables a much faster solving of prior security game models with QR adversary (no subset selection). The main result in Yang, Ordonez, and Tambe (2012) was a piecewise linear approximation based optimization solution of the standard security games with QR adversary (no subset selction of targets), which we vastly improve upon by showing that the same can be computed to optimality using the closed form formula in the lemma below. In more details, we show that optimization in Yang, Ordonez, and Tambe (2012) has a closed form *convex* dual (using the lemma below) with much fewer variables (only μ,ν) than the primal; hence the whole expensive piecewise linear approximation approach in (Yang, Ordonez, and Tambe 2012) can just be replaced by an easy gradient descent on the dual program. This result provides scalability in all other works in security games, including deployed wildlife (An et al. 2013) and coast guard (Fang et al. 2017) applications, that use the result from Yang, Ordonez, and Tambe (2012).

Algorithm 2: *FixedDualsSolver* (ν,μ,δ_0)

1 $S_o=\phi, k=0$
2 **for** $j\in\mathcal{K}$ **do**
3 Let l be that index such that $j\in\mathcal{K}_l$
4 compute $h_j(\nu,\mu,\delta_0)=\max_{y_j\in[e^{-\lambda w_j^a},1]}g_j(\nu,\mu,y_j,\delta_0)$
5 In list H, store the indexes j sorted by $h_j(\nu,\mu,\delta_0)$ in descending order.
6 Partition H into $H_1,\ldots,H_L$, such that any $j\in H_l$ satisfies $j\in\mathcal{K}_l$ and each H_l is still sorted by h_j.
7 **for** $i\in\{1,\ldots,L\}$ **do**
8 $S_o=S_o\cup\{H_l(0)\}$ `// best` h_j `per` $\mathcal{K}_l$
9 **while** $|S_o|<C$ **do**
10 $j=H(k), k=k+1$
11 **if** $h_j(\nu,\mu,\delta_0)>0$ *or* $|S_o|<N_P$ **then**
12 $S_o=S_o\cup\{j\}$
13 return $S^*=S_o$ and the y_j^*'s that maximize h_j

Lemma 1. *The solution of the problem in line 4 of Algorithm 2 is*

$$y_j^*=e^{-\lambda w_j^a\max(0,\min(\beta_j,1))}\ \textit{where}\ \beta_j\ \textit{is}$$

$$\frac{1}{\lambda w_j^a}\Big[1-\frac{\lambda w_j^a}{w_j^d}(l_j^d-\delta_0)-W\big(\frac{\nu+\mu_l}{w_j^d}e^{1-\lambda r_j^a-\frac{\lambda w_j^a}{w_j^d}(l_j^d-\delta_0)}\big)\Big]$$

and W is the Lambert W function (Corless et al. 1996).

For further notational ease, let $\Phi(S,\nu,\mu,\delta_0)=\max_{y_S\in\mathcal{D}_S^y}\phi(S,\nu,\mu,y_S,\delta_0)$. Using the solution y_S^* from the above lemma, we get that $\Phi(S,\nu,\mu,\delta_0)=\phi(S,\nu,\mu,y_S^*,\delta_0)$. Note that y_S^* is still a function of ν,μ,δ_0. It can be readily checked that Φ satisfies the conditions of Theorem 3 with Φ as f, and ν,μ as y, and S as x (δ_0 is a constant), *except* for the uniqueness of S^* (x^* in Theorem 3) for optimal ν^*,μ^* (y^* in Theorem 3). However, observe that in Algorithm 2, the solution S^* (for given ν^*,μ^*) may be non-unique only if $h_j=h_i$ for some $i\neq j$; the stringent equality needed makes non-uniqueness highly unlikely; indeed, we encounter non-uniqueness in only 1.1% cases in experiments. We prove the following:

Lemma 2. *1. Alg. 2 runs in poly time $O(|\mathcal{K}|\log|\mathcal{K}|)$.*
2. Alg. 2 solves FixedDuals *to any arbitrary fixed precision.*

Gradient descent on duals. Next, in order to solve SwitchedDualOPT, we perform projected gradient descent (PGD) on dual variables ν,μ. Recall that the inner problem in SwitchedDualOPT is $\max_{S\in F(\mathcal{K})}\max_{y_S\in\mathcal{D}_S^y}\phi(S,\nu,\mu,y_S,\delta_0)$, which, with slight abuse of notation, we refer to as FixedDuals(ν,μ,δ_0), which we already solved in Algorithm 2 for given ν,μ,δ_0. The use of PGD is justified by:

Proposition 1. FixedDuals(ν,μ,δ_0) *is convex in ν,μ.*

Proof. We again skip writing δ_0 for ease of notation. Using the notation introduced and Lagrangian duality,

Algorithm 3: *SwitchedDualOptSolver*(ξ, δ_0)

1 $(\nu^0, \mu^0) = 0$
2 **repeat**
3 $\quad (\nu^t, \mu^t) = \mathbb{P}_{\nu\mu \geqslant 0}\big((\nu^{t-1}, \mu^{t-1}) - \eta_t \nabla_{\nu,\mu} \mathsf{FixedDuals}(\nu^{t-1}, \mu^{t-1}, \delta_0)\big)$
4 $\quad S, y_S = FixedDualsSolver(\nu^t, \mu^t, \delta_0)$
5 **until** *objective changes by less than* ξ
6 return S, y_S

the inner $\Phi(S, \nu, \mu) = \max_{y_S \in D} \phi(S, \nu, \mu, y_S)$ problem is convex in ν, μ. The function in the theorem is $\max_{S \in F(\mathcal{K})} \Phi(S, \nu, \mu)$. As this is a max over multiple convex functions $\Phi(S, \nu, \mu)_{S \in F(\mathcal{K})}$, this function is convex. □

Combining all sub-results for the polynomial heuristic, in Algorithm 3 PGD is used in lines 3-4, where $\mathbb{P}_{\nu,\mu \geqslant 0}$ denotes projection to the space $\nu, \mu \geqslant 0$. The gradient of FixedDuals w.r.t. ν, μ can be computed using Danskin's Theorem (Bertsekas, Hager, and Mangasarian 1998) (details in appendix). Alg. 3, when plugged in as solver for BOPT in line 4 of the binary search Alg. 1 is the full polynomial time heuristic.

Using Theorem 1 and Lemma 2, it can be readily seen that the heuristic has a runtime of $O(\log(\frac{1}{\epsilon}) \frac{|\mathcal{K}| \log |\mathcal{K}|}{\xi})$ (gradient descent takes $O(\frac{1}{\xi})$ iterations under mild smoothness condition). Also, if the uniqueness condition of Theorem 3 holds for any problem instance then we would obtain $O(\xi + \epsilon)$ approximation. However, as stated earlier, the uniqueness might not hold in rare cases and hence we next propose a hybrid approach combining the advantages of this heuristic and the earlier MILP approach.

5 Hybrid Approach

This hybrid approach is inspired by the observation that the heuristic is time efficient and if the minimax equality hold, the resulting solution is optimal for EqOPT. If the minimax eaulity does not, we can still use solution of SwitchedDualOPT to construct tight lower and upper bounds for the MILP approach with guaranteed solution quality. Towards that end, we present the following result:

Theorem 4. *If we run the heuristic (Algorithm 3 plugged in line 4 of Algorithm 1) and obtain solution* $(\overline{\delta}_0, \overline{S}, \overline{x}_S)$*, then with* S^*, x^*_S *optimal for* EqOPT *there exists a small enough* ξ *(in Algorithm 3) such that (1)* $|\mathcal{F}(S^*, x^*_S) - \mathcal{F}(\overline{S}, \overline{x}_S)| \leqslant |\overline{\delta}_0 - \mathcal{F}(\overline{S}, \overline{x}_S)| + 2\epsilon$ *and (2)* $\mathcal{F}(\overline{S}, \overline{x}_S)$ *and* $\overline{\delta}_0 + 2\epsilon$ *can be used as a lower and upper bounds for the MILP approach (*ApxOPTL *plugged in line 4 of Algorithm 1).*

Theorem 4 implies that if $\overline{\delta}_0$ is sufficiently close to $\mathcal{F}(\overline{S}, \overline{x}_S)$, then $(\overline{S}, \overline{x}_S)$ is a near-optimal solution to EqOPT. Using the above result, we design a hybrid approach combining the MILP and the heuristic to efficiently find a near-optimal solution in Algorithm 4. It can be seen from the above result that if the algorithm stops at line 3, then the returned solution is additive 3ϵ-optimal to EqOPT,

Algorithm 4: *Hybrid algorithm*

1 run heuristic (Algorithm 3 used in Algorithm 1) to get $(\overline{\delta}_0, \overline{S}, \overline{x}_S)$
2 **if** $|\overline{\delta}_0 - \mathcal{F}(\overline{S}, \overline{x}_S)| \leqslant \epsilon$ **then**
3 $\quad$ return $(\overline{S}, \overline{x}_S)$
4 **else**
5 $\quad$ run MILP approach (Algorithm 1 with ApxOPTL), with $L = \mathcal{F}(\overline{S}, \overline{x}_S)$, $U = \overline{\delta}_0 + 2\epsilon$, obtain (S^*, x^*_S)
6 $\quad$ run heuristic with fixed $S = S^*$ to improve x^*_S
7 $\quad$ return the last solution obtained.

i.e., $\mathcal{F}(S^*, x^*_S) - \mathcal{F}(\overline{S}, \overline{x}_S) \leqslant 3\epsilon$. Otherwise, if the algorithm stops at line 7, then a guarantee is already established via Theorem 2, i.e., $\mathcal{F}(S^*, x^*_S) - \mathcal{F}(\overline{S}, \overline{x}_S) \leqslant O(1/K + \epsilon)$. So, it is guaranteed that a solution returned by Algorithm 4 is additive $O(1/K + \epsilon)$-optimal to EqOPT. In line 6, we run the heuristic again to further improve the solution of the MILP. In experiments we show that in most of the cases, the hybrid algorithm stops at line 3, making it much faster than the MILP approach.

6 Experiments

Our experiments are simulated over 10 random instances for each measurement that we report. In each game instance, payoffs are chosen uniformly randomly, from 1 to 10 for r^d_j and r^a_j and from -10 to -1 for l^d_j and l^a_j. Following past user studies (Yang et al. 2011), the parameter λ of the QR model is chosen as 0.76. We select $C = \lfloor 2\mathcal{K}/3 \rfloor$, $N_P = \lfloor \mathcal{K}/2 \rfloor$, $m = |\mathcal{K}|/10$, and split the set of centers into $L = 5$ disjoint partitions of equal size. For each partition l, we choose $\beta_l = 2m/L$. All experiments were conducted using Matlab on a Windows 10 PC with Intel i7-9700 CPUs (3.00GHz).

We compare our algorithm (Alg. 4, denoted as *Hybrid*) with two baseline methods. One is based on the convex optimization approach (denoted as *ConvexOpt*) proposed in prior work (Yang, Ordonez, and Tambe 2012); this method is not capable of choosing which center to operate hence we run it with all the centers chosen as operational. The other method is based on a two-steps procedure (denoted as *TwoSteps*); for this we first solve EqOPT with fixed $S = \mathcal{K}$ (using *ConvexOpt*) to find a strategy x^*. We then use this strategy to select at least N_P and at most C centers from $\mathcal{K}$ by sorting the individual rewards $\{w^d_j x^*_j + l^d_j, j \in \mathcal{K}\}$ and selecting N_P centers with highest rewards and then, from the remaining centers, selecting no more than $C - N_P$ centers with highest and positive rewards. Then, a subset S^* is selected and we solve EqOPT with fixed $S = S^*$ to re-optimize the strategy.

We first determine the number of pieces K needed in our PWLA approach for a good approximation, noting that, as is typical for approximation algorithms, the bound in Theorem 2 can be too conservative for average case problems. We vary K in $\{5, 10, \ldots, 30\}$ and compute the percentage gap between the objective values given by PWLA with K

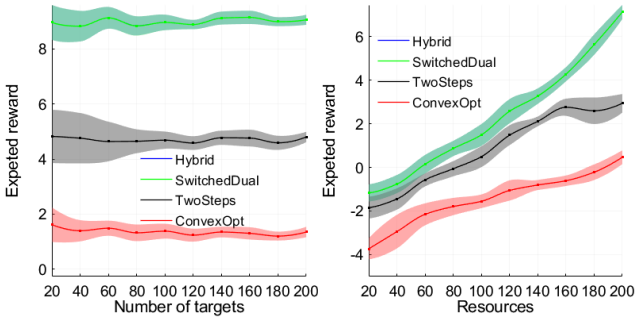


Figure 2: Expected reward comparison, the curves given by *Hybrid* and *SwitchedDual* are almost identical.

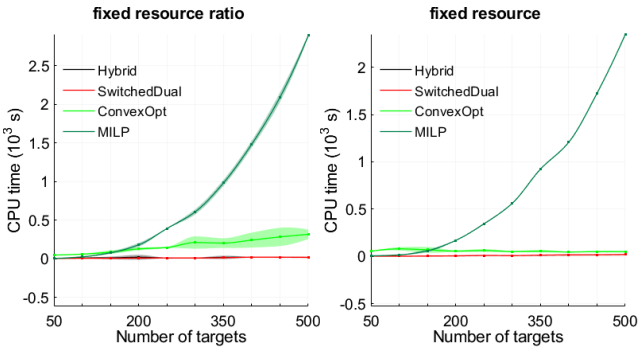


Figure 3: Computational scalability comparison, the curves given by *Hybrid* and *SwitchedDual* are almost identical.

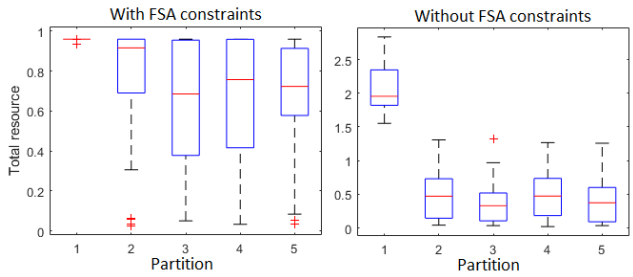


Figure 4: Fairness in security allocation over vaccine centers

pieces and with a large $\overline{K} = 200$ pieces. We observe that the percentage gaps become relatively small (less than 1.4%) if $K \geqslant 20$ (figure in Appendix). Thus, we fix $K = 20$ for the rest of the experiments.

Expected rewards comparison. The means and standard errors of the expected rewards of different approaches are plotted in Fig. 2, where in the left figure we vary the number of centers from 20 to 200 and set resource budget as $m = |\mathcal{K}|/10$, and in the right figure we fix $|\mathcal{K}| = 50$ and vary the resource budget m from 2 to 20. The expected rewards of *SwitchedDual* are equal to those of *Hybrid* for 178/180 test instances, and only slightly smaller for 2/180 instances. This shows that the *minimax equality* holds with high probability.

Also, *Hybrid* and *SwitchedDual* consistently outperform other methods. In particular, for varying number of centers, *Hybrid* provides 85%-96% larger rewards than *TwoSteps*. For varying resource budget, the improvements are up to 142%. *ConvexOpt* returns very low rewards, revealing clear benefit of selecting centers instead of operating all centers.

Computational scalability. In this experiment, we run *Hybrid*, *SwitchedDual*, *ConvexOpt* and *MILP*, where MILP refers to using ApxOPTL. We vary the number of potential centers from 50 to 500 with two settings, one with a fixed resource ratio as $m = 0.1|\mathcal{K}|$ and one with a fixed resource budget $m = 20$. Fig. 3 shows the means and standard errors of the CPU time of the four approaches over 10 repetitions. We see that the curves given by the *Hybrid* and *SwitchedDual* are almost identical (a log scale figure is in appendix), indicating that Alg. 4 mostly stops at line 3, i.e., the minimax equality holds. There are only a few instances of $|\mathcal{K}| = 200$ or $|\mathcal{K}| = 350$ that Alg. 4 stopped at line 7, noting that even in these cases, the total time is still about 5-10 times less than the time required by the *MILP*, due to the tightness of the bounds provided by the *SwitchedDual* in line 5. While *ConvexOpt* solves an easier problem with no selection of centers, still *Hybrid* runs faster than *ConvexOpt* on average.

The time taken by *MILP* grows very fast as the number of centers increases whereas the time required by *Hybrid* and *SwitchedDual* is stable and small. More precisely, for fixed resource ratio, with $|\mathcal{K}| = 500$, the times required by *Hybrid, SwitchedDual*, *ConvexOpt* and *MILP* are about 4, 4, 400 and 4700 (secs), respectively. To further demonstrate scalability, we increased the number of centers to 5000 and *Hybrid* and *SwitchedDual* finished in 70 seconds.

We evaluate the impact of the FSA constraints on the fairness of security resource distribution. To this end, we select $|\mathcal{K}| = 20$ and also divide all the centers into $L = 5$ partitions of equal size. To better illustrate the fairness, we add 5 units to r_j^a, $\forall j \in \mathcal{K}_1$. This implies that the adversary will get more rewards if attacking a target in Partition #1. We also tighten the selection of parameters β_l by choosing $\beta_l = 1.2m/L$, for all $l = 1, \ldots, L$. The box plot in Fig. 4 reports the distributions of the total resources assigned to the 5 partitions, with and without the FSA constraints. Without the FSA constraints, the first partition gets a high chance of being protected (allocating an expected number of more than two resources to the four centers in Partition #1), thus lowering the protection of other partitions. On the other hand, the FSA constraints maintains a fairness of security allocation between partitions.

7 Conclusion

We proposed a model for security of vaccine delivery in underdeveloped and security risk-prone areas and presented an efficient solver using a novel strong duality result with discrete variables. As such, a number of other schemes require security (setting up voting centers, medical camps, etc.). We believe our model can serve as a basis for formally modeling such problems and its variants. Our technical approach can also inform other techniques where a mix of discrete and continuous variables are used.

Acknowledgments

This research/project is supported by the National Research Foundation Singapore and DSO National Laboratories under the AI Singapore Programme (AISG Award No: AISG2-RP-2020-017).

References

Agrawal, S.; Avadhanula, V.; Goyal, V.; and Zeevi, A. 2016. A near-optimal exploration-exploitation approach for assortment selection. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, 599–600.

An, B.; Ordóñez, F.; Tambe, M.; Shieh, E.; Yang, R.; Baldwin, C.; DiRenzo III, J.; Moretti, K.; Maule, B.; and Meyer, G. 2013. A deployed quantal response-based patrol planning system for the US Coast Guard. *Interfaces*, 43(5): 400–420.

Aziz, H.; Chan, H.; Lee, B.; Li, B.; and Walsh, T. 2020. Facility location problem with capacity constraints: Algorithmic and mechanism design perspectives. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 34, 1806–1813.

Benati, S.; and Hansen, P. 2002. The maximum capture problem with random utilities: Problem formulation and algorithms. *European Journal of Operational Research*, 143(3).

Bertsekas, D. P.; Hager, W.; and Mangasarian, O. 1998. *Nonlinear programming*. Athena Scientific Belmont, MA.

Bertsimas, D.; Brown, D. B.; and Caramanis, C. 2011. Theory and applications of robust optimization. *SIAM review*, 53(3): 464–501.

Cerny, J.; Lisý, V.; Bošanský, B.; and An, B. 2020. Dinkelbach-Type Algorithm for Computing Quantal Stackelberg Equilibrium. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 246–253.

Cheung, W. C.; and Simchi-Levi, D. 2017. Thompson sampling for online personalized assortment optimization problems with multinomial logit choice models. *Available at SSRN 3075658*.

Corless, R. M.; Gonnet, G. H.; Hare, D. E.; Jeffrey, D. J.; and Knuth, D. E. 1996. On the Lambert W function. *Advances in Computational mathematics*, 5(1): 329–359.

Davis, J. M.; Gallego, G.; and Topaloglu, H. 2014. Assortment optimization under variants of the nested logit model. *Operations Research*, 62(2): 250–273.

Désir, A.; Goyal, V.; and Zhang, J. 2020. Capacitated Assortment Optimization: Hardness and Approximation. *Available at SSRN 2543309*.

Dinkelbach, W. 1967. On nonlinear fractional programming. *Management science*, 13(7): 492–498.

Fang, F.; Nguyen, T. H.; Pickles, R.; Lam, W. Y.; Clements, G. R.; An, B.; Singh, A.; Schwedock, B. C.; Tambe, M.; and Lemieux, A. 2017. PAWS—A deployed game-theoretic application to combat poaching. *AI Magazine*, 38(1).

Freire, A. S.; Moreno, E.; and Yushimito, W. F. 2016. A branch-and-bound algorithm for the maximum capture problem with random utilities. *European journal of operational research*, 252(1).

Gallego, G.; Ratliff, R.; and Shebalov, S. 2015. A general attraction model and sales-based linear program for network revenue management under customer choice. *Operations Research*, 63(1): 212–232.

Gannon, K.; Meldrum, A.; and Keath, L. 2020. Wars, instability pose vaccine challenges in poor nations. https://www.startribune.com/wars-instability-pose-vaccine-challenges-in-poor-nations/600004408/. Online; accessed 01 January 2021.

Goktas, D.; and Greenwald, A. 2021. Convex-Concave Min-Max Stackelberg Games. *Advances in Neural Information Processing Systems*, 34: 2991–3003.

Golowich, N.; Narasimhan, H.; and Parkes, D. C. 2018. Deep Learning for Multi-Facility Location Mechanism Design. In *27th International Joint Conference on Artificial Intelligence International Joint Conference on Artificial Intelligence (IJCAI)*, 261–267.

Haghtalab, N.; Fang, F.; Nguyen, T. H.; Sinha, A.; Procaccia, A. D.; and Tambe, M. 2016. Three Strategies to Success: Learning Adversary Models in Security Games. In *25th International Joint Conference on Artificial Intelligence (IJCAI)*.

Immorlica, N.; Lucier, B.; Mao, J.; Syrgkanis, V.; and Tzamos, C. 2018. Combinatorial Assortment Optimization. In *Web and Internet Economics*, 218–231.

Mai, T.; and Lodi, A. 2020. A multicut outer-approximation approach for competitive facility location under random utilities. *European Journal of Operational Research*, 284(3).

McFadden, D. L. 1976. Quantal choice analaysis: A survey. In *Annals of Economic and Social Measurement, Volume 5, number 4*, 363–390. NBER.

McKelvey, R. D.; and Palfrey, T. R. 1995. Quantal response equilibria for normal form games. *Games and economic behavior*, 10(1): 6–38.

Milec, D.; Černỳ, J.; Lisỳ, V.; and An, B. 2021. Complexity and Algorithms for Exploiting Quantal Opponents in Large Two-Player Games. In *Proceedings of the AAAI Conference on Artificial Intelligence*.

Sinha, A.; Fang, F.; An, B.; Kiekintveld, C.; and Tambe, M. 2018. Stackelberg security games: Looking beyond a decade of success. In *27th International Joint Conference on Artificial Intelligence (IJCAI)*.

Tambe, M. 2011. *Security and game theory: algorithms, deployed systems, lessons learned*. Cambridge university press.

Wang, R. 2012. Capacitated assortment and price optimization under the multinomial logit model. *Operations Research Letters*, 40(6): 492–497.

Wu, T.-H. 1997. A note on a global approach for general 0–1 fractional programming. *European Journal of Operational Research*, 101(1): 220–223.

Xu, H. 2016. The mysteries of security games: Equilibrium computation becomes combinatorial algorithm design. In *Proceedings of the 2016 ACM Conference on Economics and Computation*, 497–514.

Yang, R.; Kiekintveld, C.; Ordonez, F.; Tambe, M.; and John, R. 2011. Improving resource allocation strategy against human adversaries in security games. In *International Joint Conference on Artificial Intelligence (IJCAI)*, 458.

Yang, R.; Ordonez, F.; and Tambe, M. 2012. Computing optimal strategy against quantal response in security games. In *11th International Conference on Autonomous Agents and Multiagent Systems (AAMAS)*.

Zaffran, M.; Vandelaer, J.; Kristensen, D.; Melgaard, B.; Yadav, P.; Antwi-Agyei, K.; and Lasher, H. 2013. The imperative for stronger vaccine supply and logistics systems. *Vaccine*, 31: B73–B80.