An Online Learning Approach towards Far-sighted Emergency Relief Planning under Intentional Attacks in Conflict Areas

Haoyu Yang, **Kaiming Xiao***, **Lihua Liu**, **Hongbin Huang** and **Weiming Zhang**
Laboratory for Big Data and Decision,
National University of Defense Technology, Changsha, P.R.China 410073
{yanghaoyu20, kmxiao,hbhuang,wmzhang}@nudt.edu.cn

Abstract

A large number of emergency humanitarian rescue demands in conflict areas around the world are accompanied by intentional, persistent and unpredictable attacks on rescuers and supplies. Unfortunately, existing work on humanitarian relief planning mostly ignores this challenge in reality resulting a parlous and short-sighted relief distribution plan to a large extent. To address this, we first propose an offline multi-stage optimization problem of emergency relief planning under intentional attacks, in which all parameters in the game between the rescuer and attacker are supposed to be known or predictable. Then, an online version of this problem is introduced to meet the need of online and irrevocable decision making when those parameters are revealed in an online fashion. To achieve a far-sighted emergency relief planning under attacks, we design an online learning approach which is proven to obtain a near-optimal solution of the offline problem when those online revealed parameters are generated *i.i.d.* from unknown distributions. Finally, extensive experiments on a real anti-Ebola relief planning case based on the data of Ebola outbreak and armed attacks in DRC Congo show the scalability and effectiveness of our approach.

1 Introduction

Humanitarian relief operations have become much more significant because of the increasing number of disasters around the world [Besiou and van Wassenhove, 2020; De Vries and Van Wassenhove, 2020]. For instance, 2011 Japan earthquake, 2017 Hurricane Maria, 2019 Australia Wildfires, as well as the ongoing virus epidemics, i.e., the COVID-19 across the world. These disasters have taken millions of lives and resulted in huge economic losses. Fueled by the need of mitigating the miserable effects of these disasters, humanitarian relief operations and disaster management received an increasing attention from researchers, and different aspects of the operations management have been studied in the literature [Huang and Song, 2018; Ekici and Özener, 2020].

However, despite the uncertainty derived from the disaster itself, security issue and the derived uncertainty from malicious attackers are also vital factors affecting the actual humanitarian relief operations in conflict areas [Chaudhri *et al.*, 2019]. There has been a rise in the number of attacks on humanitarian aid workers (including UN staff, NGO staff, etc.), reporting 226 attacks on aid workers in 2018 with 131 deaths, which is nearly four times the 2004 number of 63 incidents (with 56 deaths) [AWSD, 2018]. Attacks against aid workers have become a big challenge facing humanitarian organizations to provide security in fragile and conflict-affected areas [Hoelscher *et al.*, 2017]. For example, the second largest DRC Congo Ebola outbreak in December 2019 has been a humanitarian crisis that originated in an active conflict area, which has severely affected the ability of relief efforts and vaccination [Schwartz, 2020]. The major epidemic area is the Eastern Congo including Nord-Kivu and Sud-Kivu province, in which more than 10 armed rebel groups exist [Usanov *et al.*, 2013]. These armed groups actively operate in the regions of Ebola virus relief, controlling the routes in and out of the regions; levying taxes on households and transport; attacking healthcare and relief workers [Schwartz, 2020]. Hence, humanitarian relief planning practice in conflict-affected areas facing huge risk and uncertainty from hostile rebel groups.

To the best of our knowledge, few studies has focused on the multi-stage online emergency relief planning problem in conflict areas (MonERP) which takes the uncertainty from both adversary and environment into account. In this paper, we address this challenging problem, in which the humanitarian organization aims at transporting urgent resources (e.g., living and medical supplies) from the resource base to cities applying for the relief in real time, while the hostile armed group operates to hinder the relief operations by controlling some routes and levying taxes on transport. Different from general extensive-form games, data and parameters of the game (e.g., the amount of demands, the cost of transportation, the distribution of hostile rebel groups) come in an online fashion instead of being revealed at the outset. On receiving a relief request, the rescuer with limited resources needs to make an irrevocable decision whether to accept or reject the demand, with the overall objective of maximizing the demand satisfaction while considering the security risks of

*Corresponding author

humanitarian operations. The attacker, on the other hand, has an opposite interest and can mobilize armed forces to control towns in their surrounding area.

In this work, we aim to develop the first salable and efficient online learning approach to MonERP, and the key contributions are: (1) We first model the omniscient version of MonERP as a multi-stage offline emergency relief planning problem (MoffERP), which is transformed as a binary integer linear programming with all data and parameters known in advance. The solution of MoffERP provides us with the upper bound of MonERP. (2) To meet the need of online and irrevocable decision making, MonERP is modeled as an online linear programming after integrating the result of Nash Equilibrium in each decision-making stage. (3) An online learning algorithm is proposed to solve MonERP, which achieves a $O(\frac{1}{\sqrt{n}})$ average regret when those online revealed parameters are generated *i.i.d.* from an unknown distribution; here, n is the number of decision-making stages. (4) Extensive experiments are conducted on a real anti-Ebola relief planning case based on the data of Ebola outbreak and armed attacks in DRC Congo. Evaluation results show the performance advantages of the far-sighted relief planning considering online features, and the proposed approach are applicable to large realistic-scale problems efficiently.

2 Related Works

2.1 Humanitarian Relief Planning

Humanitarian relief operations require efforts on many fronts, such as facility location and network design, vehicle routing, inventory management, network flow, and combination of them, etc. Typical works of relief routing problems include models and methods for achieving objectives like minimizing delivery delay in various situations [Yuan and Wang, 2009; Sabouhi *et al.*, 2019], maximizing the number of victims served [Ozdamar *et al.*, 2018], minimizing life losses and human sufferings [Huang and Song, 2018], reducing cost and guaranteeing fairness [Zhu *et al.*, 2019], etc. Some models focus on routing problems with determinate parameters, while problems with uncertainty attract more attention since they are closer to actual needs. For example, Allahviranloo et al. [2014] use robust optimization to address the demand uncertainty in vehicle routing problems. A robust optimization approach with a coaxial uncertainty set is developed by [Balcik and Yanıkoğlu, 2020] for humanitarian needs assessment routing under travel time uncertainty. Unfortunately, the risk and uncertainty derived from malicious attacks are not considered in existing research.

2.2 Network Interdiction Game

Network interdiction problem is a kind of Stackelberg game, which has been widely applied to security decision-making scenarios relating to operations on networks, involving military operations, humanitarian relief, fare evasion, telecommunication, transportation and logistics, and cyber-security [Smith and Song, 2020; Sinha *et al.*, 2018]. Besides of classic models [Fulkerson and Harding, 1977; Israeli and Wood, 2002], different variants of SPNI models with new features have received considerable attentions recently, including incomplete information [Borrero *et al.*, 2019], goal recognition assistance [Xu *et al.*, 2017], dynamic or adaptive interdiction [Correa *et al.*, 2017; Xiao *et al.*, 2018; Zhang *et al.*, 2019] etc. However, few studies focus on the interdiction game with online data inputs.

2.3 Online Linear Programming

Sequential decision making has been an increasingly attractive research topic especially in scenarios with dynamic environment and long-term consideration. For instance, the scenario of online packing problem[Buchbinder and Naor, 2009], auction problem [Balseiro and Gur, 2019], and resource allocation[Balseiro *et al.*, 2020]. These problems can be viewed as online linear programming (OLP), i.e., the problem of solving linear programs in a sequential setting. A simple and fast online algorithm for solving a general class of binary integer linear programs is proposed in [Li *et al.*, 2020] achieving $O(m\sqrt{n})$ expected regret under the stochastic input online data, where n is the number of decision variables and m is the number of constraints. These studies provide a theoretically guaranteed algorithm for long-term revenue in sequential decision making, which is instructive for achieving far-sighted emergency relief planning.

3 Offline Multi-stage Relief Planning

In this section, we first model the offline version of MonERP as a sequence of stage-based games between the rescuer and attacker, after which it is integrated and transformed to a linear programming. Then, based on the mixed Nash equilibrium of stage-based games, the optimal offline relief planning is given by solving the linear programming.

3.1 The Simultaneous Relief Routing Game at Each Stage

In each stage of the MonERP, there is a simultaneous relief routing game between the rescuer and attacker. Both of them play the game on a road network $G = (V, E)$ consisting a set of nodes V representing intersections (cities,towns, etc.) and a set of edges E representing roads. If the rescuer in stage t decides to meet the relief demand w_t from the resource base city o_t to the city of demand g_t, she needs to plan a road path through the network so as to minimize the cost of transporting. The attacker, on the other hand, has several armed forces distributed in different nodes of the network and aims to mobilize them to interdict the rescuer thereby levying taxes or directly robbing under the constraint of mobility. We then model this stage-based confrontation between the rescuer and attacker as a simultaneous relief routing game as follows:

The attacker problem is implementing his armed forces distributed in the network to control some selected nodes so as to interdicting the relief routing of the rescuer. Denote by $\boldsymbol{h}_t = (h_{jt})_{j \in V} \in \{0,1\}^{|V|}$ the distribution vector of the armed forces, and $h_{jt} = 1$ if there is an arm force stationed at node j. The strategy of the attacker is made up of $|I|$ vectors, denoted by $\boldsymbol{z}_{it} = (z_{ijt})_{j \in V} \in \{0,1\}^{|V|}$, and $z_{ijt} = 1$ if the armed force i is decided to reach and control node j at stage t. Also, the actions of attacker are restricted by movement constraints, i.e., the movement cost of each armed force $i \in I$ should not exceed the budget f_{it} at stage t (vector form $\boldsymbol{f}_t$). Then, the set of all feasible strategies for the attacker can be denoted by

$$\mathscr{X}_t = \left\{ Z_t = (\boldsymbol{z}_{it})_{i \in I} \,\middle|\, Z_t^\top \mathbf{1} \leq \mathbf{1},\ Z_t^\top L \boldsymbol{h}_t \leq \boldsymbol{f}_t, \right\}, \quad (1)$$

where $\mathbf{1}$ denotes the all 1 vectors, and the first constraint illustrates that one armed force can only be deployed once in each stage. Let $L = (l_{j'j})_{j',j \in V} \in \mathbb{R}_+^{|V| \times |V|}$ represent the cost of implementing the stationed force from node j to node j', which can be estimated by the attacker using road transport information. Hence, the second constraint makes the deployment of armed forces not exceed the attacker's budget.

The rescuer problem is planning a routing path through the network from o_t to g_t which can be modeled as a shortest-path problem. Denote by $\boldsymbol{y}_t = (y_{kt})_{k \in E} \in \{0,1\}^{|E|}$ the decision variables of the rescuer representing a path in the network, and $y_{kt} = 1$ if she selects the edge k to traverse. Let $FS(i)$ and $RS(i)$ represent the set of edges directed from/into node i. Then, the strategy space of the rescuer is determined by the following flow conversation constraints:

$$\sum_{k \in FS(i)} y_{kt} - \sum_{k \in RS(i)} y_{kt} = \begin{cases} 1 & \text{for} \quad i = o_t \\ 0 & \forall\, i \in V \backslash \{o_t, g_t\}, \\ -1 & \text{for} \quad i = g_t \end{cases} \quad (2)$$

$$y_{kt} \geq 0. \quad \forall\, k \in E \quad (3)$$

Then the set of all feasible strategies of the rescuer is

$$\mathscr{Y}_t = \left\{ \boldsymbol{y}_t \,\middle|\, \text{Constraints (2), (3)} \right\}, \quad (4)$$

The utility for both players is the cost of emergency relief transporting. The rescuer aims to minimize it by selecting the path traversed, while the attacker interdicts the relief routing. The utility function is bi-linear and made up of their decision variables as follows:

$$u_t = \sum_{k \in E} (c_k + z'_{kt} d_k) y_{kt}. \quad (5)$$

where c_k (vector form $\boldsymbol{c}$) denotes the cost of traversing edge k, and d_k (vector form $\boldsymbol{d}$) represents the additional cost of passing through edge k which is controlled by the attacker with an armed force. Let z'_{kt} be an intermediate decision variable, and $z'_{kt} = 1$ if the edge k is under control of the attacker at stage t. Specifically, we suppose that all edges directing out of a force garrisoned node are under control, i.e.,

$$z'_{kt} = z_{ijt}. \forall\, k \in FS(j), j \in V, i \in I \quad (6)$$

Denote by $\mathscr{P} = \{\text{attacker}, \text{rescuer}\}$ the set of players. Then, the simultaneous relief routing game at each stage can be expressed as a tuple $\mathcal{G}_t = (\mathscr{P}, \mathscr{X}_t, \mathscr{Y}_t, u_t)$. The final cost of transporting a_t (vector form $\boldsymbol{a}$) in this game is defined as the mixed Nash equilibrium of game $\mathcal{G}_t$.

Additionally, if the rescuer selects an attacker-controlled edge to traverse at one stage, the transported relief resources w_t will suffer a loss of being levied or robbed. The amount of loss at edge k depends on the levy rate $\boldsymbol{q} = (q_k)_{k \in E} \in (0,1)^{|E|}$ of armed forces the rescuer encountered. Denote by r_t (vector form $\boldsymbol{r}$) the left relief resources which can be successfully transported, we have

$$r_t = (1 - \sum_{k \in E} z'_{kt} q_k y_{kt}) w_t. \quad (7)$$

3.2 Multi-stage Offline Emergence Relief Planning

Based on results of game $\mathcal{G}_t$ at each stage, we construct a multi-stage offline emergence relief planning (MoffERP) by assuming that all data and parameters are known in advance. In this way, MoffERP is integrated as a linear programming.

Specifically, the rescuer has to make decisions about whether or not to meet the relief demand of a certain city at each stage. Denote by n the number of stages of MoffERP, and $N = \{1, 2 \cdots, n\}$ the set of stages. Let $\boldsymbol{x} = (x_t)_{t \in N} \in \{0,1\}^n$ represent the decision variables of the rescuer, and $x_t = 1$ if she decides to meet the demand w_t at stage t. To meet this demand, the rescuer has to pay a cost of a_t on transportation and can cover a certain amount of demand r_t as a kind of revenue at this stage. Hence, as a long-term revenue seeker, the rescuer aims to maximize the total amount of demand satisfaction under the limitation on total amount of relief resources b_1 and transportation budget b_2.

Supposing that all online data r_t and a_t are known in advance, we can formulate MoffERP to a binary integer linear programming as follows:

$$\text{[P-MoffERP]} \quad \max_{\boldsymbol{x}} \boldsymbol{r}^\top \boldsymbol{x}$$

$$\text{s.t.} \quad [\boldsymbol{w} \quad \boldsymbol{a}]^\top \boldsymbol{x} \leq \boldsymbol{b}, \quad (8)$$

where $\boldsymbol{x} \in \{0,1\}^n$, vectors $\boldsymbol{r}$, $\boldsymbol{w}$, $\boldsymbol{a} \in \mathbb{R}_+^n$, and denote by $\boldsymbol{b}$ the vector $[b_1, b_2]^\top$. In this way, MoffERP can be solved using common binary integer linear programming techniques. Denote by $\bar{\boldsymbol{x}}^*$ the optimal solution of P-MoffERP.

4 Online Multi-stage Relief Planning

In this section, we extend the MoffERP to a realistic online version (MonERP) where key parameters are revealed in an online manner. An online learning algorithm is then designed to solve it with theoretical analysis on average expected regret.

4.1 Online Problem Modeling

In the online version of problem, both players have to make irrevocable decisions without observing the future inputs. At each stage t, the parameter r_t and a_t are revealed to the rescuer based on the Nash equilibrium results of the stage game $\mathcal{G}_t$. Simultaneously, the rescuer needs to decide the value of x_t in real time. Unlike the setting in MoffERP, rescuer only knows the history information $\mathcal{H}_t = \{r_j, w_j, a_j, x_j\}_{j=1}^{t-1}$. Hence, the online decision strategy of rescuer can be presented as a function φ of the history and the observed parameters at the current stage t:

$$\text{[P-MonERP]} \quad x_t = \varphi(r_t, w_t, a_t, \mathcal{H}_t). \quad (9)$$

To design the function φ, we first analyze the distribution features of online parameters. It is assumed that the demand

$(g_t, w_t, \boldsymbol{f}_t, \boldsymbol{h}_t)$ in game $\mathcal{G}_t$ are generated *i.i.d.* from an unknown distribution. As mentioned in Section 3, online parameters (r_t, a_t) are generated in the process of each stage-based game $\mathcal{G}_t$. Denote by $\mathscr{N}$ the Nash equilibrium mapping from the parameters $(g_t, w_t, \boldsymbol{f}_t, \boldsymbol{h}_t)$ in game $\mathcal{G}_t$ to parameters r_t and a_t, i.e.,

$$(r_t, a_t) = \mathscr{N}(g_t, w_t, \boldsymbol{f}_t, \boldsymbol{h}_t). \tag{10}$$

In Theorem 1, (r_t, a_t) are proven to be *i.i.d.* sampled from unknown distribution $\mathcal{P}$.

Theorem 1. *The coefficient pair (r_t, a_t) are generated i.i.d. from unknown distribution, if parameters $(g_t, w_t, \boldsymbol{f}_t, \boldsymbol{h}_t)$ are generated i.i.d. from unknown distribution.*

Proof. From the definition of *i.i.d.*, we can prove that r_t and a_t satisfy the independent condition first. As shown in Equation (7) and (10), the parameters $g_t, w_t, \boldsymbol{f}_t, \boldsymbol{h}_t$ used to calculate r_t are independent at different stages, so that r_t satisfies the independent condition. Similarly, it can be proved that a_t is independent. Hence we have

$$\begin{cases} P(r_1 \cdots r_n) = P(r_1) \cdots P(r_n), \\ P(a_1 \cdots a_n) = P(a_1) \cdots P(a_n). \end{cases} \tag{11}$$

As for proving (r_t, a_t) are generated from identical distribution, both r_t and a_t are computed in a fixed pattern at each stage and the parameters $g_t, w_t, \boldsymbol{f}_t, \boldsymbol{h}_t$ are subject to identical distribution, which we could prove that (r_t, a_t) are subject to identical distribution too. Hence we have

$$\begin{cases} P(r_i = r) = P(r_j = r), \forall\ r \in R \\ P(a_i = a) = P(a_j = a). \forall\ a \in A \end{cases} \tag{12}$$

Therefore we conclude the proof. □

4.2 Online Learning Approach

Before given the online learning algorithm to MonERP, we present the linear programming relaxation of P-MoffERP as follows:

$$[\text{LP-MoffERP}] \quad \max_{\boldsymbol{x}} \boldsymbol{r}^\mathsf{T} \boldsymbol{x}$$

$$\text{s.t.} \quad [\boldsymbol{w} \quad \boldsymbol{a}]^\mathsf{T} \boldsymbol{x} \le \boldsymbol{b}, \tag{13}$$

$$\boldsymbol{0} \le \boldsymbol{x} \le \boldsymbol{1}. \tag{14}$$

Then, the dual problem of LP-MoffERP can be formulated as

$$[\text{DP-MoffERP}] \quad \min_{\boldsymbol{p}, \boldsymbol{s}} \boldsymbol{b}^\mathsf{T} \boldsymbol{p} + \boldsymbol{1}^\mathsf{T} \boldsymbol{s}$$

$$\text{s.t.} \quad [\boldsymbol{w} \quad \boldsymbol{a}]\, \boldsymbol{p} + \boldsymbol{s} \ge \boldsymbol{r}, \tag{15}$$

$$\boldsymbol{p} \ge \boldsymbol{0}, \boldsymbol{s} \ge \boldsymbol{0}, \tag{16}$$

where the dual decision variables are $\boldsymbol{p} \in \mathbb{R}^n$ and $\boldsymbol{s} \in \mathbb{R}^n$. Denote by $\boldsymbol{x}^*$, $\boldsymbol{p}_n^*$, and $\boldsymbol{s}^*$ the optimal solutions of problem LP-MoffERP and DP-MoffERP. From the complementary condition, we have

$$x_j^* = \begin{cases} 1, & r_j > [w_j \quad a_j]\, \boldsymbol{p}_n^* \\ 0, & r_j < [w_j \quad a_j]\, \boldsymbol{p}_n^* \end{cases} \quad \forall j \in 1, 2, \cdots, n \tag{17}$$

Algorithm 1 Online learning algorithm for MonERP

Input: n, online revealed coefficient pair $(\boldsymbol{r}, \boldsymbol{w}, \boldsymbol{a})$
Parameter: learning rate $\gamma_t = \frac{1}{\sqrt{n}}$
Output: a sequence of online decisions $\boldsymbol{x}$

1: Let $\boldsymbol{e} = \frac{\boldsymbol{b}}{n}$
2: Initialize $\boldsymbol{p}_1 = \boldsymbol{0}$
3: **for** $t = 1, 2, \cdots, n$ **do**
4: Set
$$x_t = \begin{cases} 1, & r_t > [w_t \quad a_t]\, \boldsymbol{p}_t \\ 0, & r_t \le [w_t \quad a_t]\, \boldsymbol{p}_t \end{cases}$$
5: Compute
$$\boldsymbol{p}_{t+1} = \max\{\boldsymbol{p}_t + \gamma_t(\begin{bmatrix} w_t \\ a_t \end{bmatrix} x_t - \boldsymbol{e}), \boldsymbol{0}\}$$
6: **end for**
7: **return** $\boldsymbol{x} = (x_t)_{t \in N}$

and when $r_j = [w_j \quad a_j]\, \boldsymbol{p}_n^*$, the value of x_j^* may be non-integer.

Note that parameters $\boldsymbol{r}$, $\boldsymbol{w}$ and $\boldsymbol{a}$ are *i.i.d.* sampled from unknown distribution. Based on Equation (17), we can obtain an online algorithm using the theoretical results of online linear programming technique [Li *et al.*, 2020] as shown in Algorithm 1. The step learning rate γ_t is set as $\frac{1}{\sqrt{n}}$ at each stage t.

To evaluate the performance of the online learning algorithm, we introduce the performance measure – *regret*– which is a common metric of online learning approach. Denote by $R_n^* = \boldsymbol{r}^\mathsf{T} \boldsymbol{x}^*$ the optimal objective value of the online problem P-MonERP, and $R_n = \boldsymbol{r}^\mathsf{T} \boldsymbol{x}$ the actual objective value under the online learning strategy $\boldsymbol{x}$. The expected optimality gap between them is

$$\Delta_n^{\mathcal{P}} = \mathbb{E}[R_n^* - R_n]. \tag{18}$$

Denote by Ξ the family of distribution $\mathcal{P}$, then the definition of *regret* is formally given as

$$\Delta_n = \sup_{\mathcal{P} \in \Xi} \Delta_n^{\mathcal{P}}. \tag{19}$$

Corollary 1. *If the step learning rate $\gamma_t = \frac{1}{\sqrt{n}}$ for $t \in N$, then Algorithm 1 achieves $O(\frac{1}{\sqrt{n}})$ average regret of problem MonERP.*

Proof. According to the state-of-the-art general theoretical results of online linear programming [Li *et al.*, 2020], if the step learning rate is set as $\gamma_t = \frac{1}{\sqrt{n}}$ then the regret satisfies $\Delta_n^{\mathcal{P}} \le m(\bar{a} + \bar{e})^2 \sqrt{n}$. We then apply this result to MonERP, where $m = 2$ representing the rows number of $[\boldsymbol{w} \quad \boldsymbol{a}]$. Hence, we have

$$\frac{\Delta_n^{\mathcal{P}}}{n} \le \frac{2(\bar{a} + \bar{e})^2}{\sqrt{n}}. \tag{20}$$

where

$$\bar{a} = \max_{j \in N} \max\{w_j, a_j\}, \tag{21}$$

$$\bar{e} = \max\{\frac{b_1}{n}, \frac{b_2}{n}\}. \tag{22}$$

It is clear that $\bar{a}$ and $\bar{e}$ are both finite constants in each instance of problem MonERP. Therefore, Algorithm 1 achieves $O(\frac{1}{\sqrt{n}})$ average regret as required. $\square$

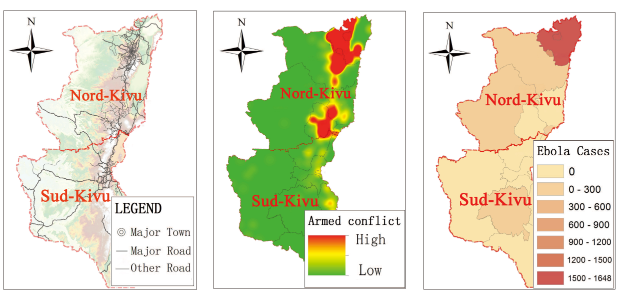
According to Corollary 1, when the number of stages $n \rightarrow \infty$, we have the average regret $\frac{\Delta_n}{n} \rightarrow 0$, which theoretically illustrates the performance advantage of the proposed online learning algorithm.

5 Experimental Evaluation

In this section, we evaluate the proposed online learning approach for emergency relief planning under intentional attacks using a real case in the conflict area of DRC Congo. The convergence and near-optimality of the online learning algorithm are first analyzed under the real case scenario and data. Then, the influence of attacker's deployment budget on the humanitarian relief operations.

5.1 Test Case and Environment

A case study of MonERP is conducted on the base of statistical data of DRC Congo including road network data, Ebola outbreak data, and conflict statistics in Nord-Kivu and Sud-Kivu province. The rescuer plans to pass through the conflict area in order to transport supplies to the city affected severely by the epidemic, while the rebel group as an attacker aims to gain payoff by controlling some routes.



(a) Road Network (b) Conflict Map (c) Ebola Outbreak

Figure 1: The Situation of Relief Routing, Ebola Epidemic and Rebel Conflicts in Nord-Kivu and Sud-Kivu Province of DRC Congo

The road network of Nord-Kivu and Sud-Kivu province is shown in Figure 1(a), which includes 810 vertices (i.e., cities and towns) and 2,188 arcs (i.e., major and other roads). The conflict data from 31 December 2018 to 5 January 2020 in the same area is shown in Figure 1(b), which is offered by the Armed Conflict Location & Event Data Project [ACLED, 2020]. The Ebola outbreak data in the same period reported by the World Health Organization is visualized as the map in Figure 1(c) [WHO, 2020]. Supposing that the emergency relief demand of a city is positively correlated with the number of Ebola cases outbreak in the city, we obtain the statistics on urban relief demand over time as shown in Figure 2. It is clear that the worst Ebola-affected areas such as Beni in Nord-Kivu have urgent and huge needs of humanitarian relief; however, the road network from non-infected areas such as Bukavu in Sud-Kivu province to these demand cities is in the control of rebel groups to a large extent.

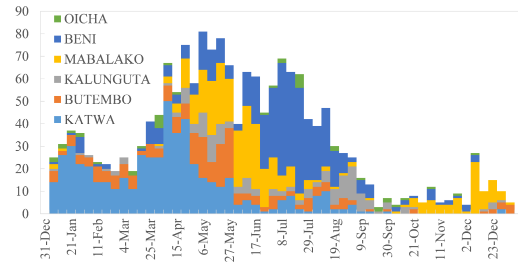


Figure 2: Statistics on Urban Relief Demand over Time in Nord-Kivu and Sud-Kivu province of DR Congo

Specifically, the traversing cost $\boldsymbol{c}$ is set as the length of the actual road between cities. The added cost of transporting $\boldsymbol{d}$ is assumed to be uniformly distributed on $[1, \bar{\boldsymbol{c}}]$, where $\bar{\boldsymbol{c}}$ denotes the average value of $\boldsymbol{c}$. The movement cost $\boldsymbol{f}_t$ is supposed to be uniformly distributed on $[1, v\bar{\boldsymbol{c}}]$, where v represents the maximum cost coefficient. Denote by $\mathcal{W} = \sum_{t=1}^{n} w_t$ the total demand during this period, and $\mathcal{C}$ the total transporting cost of satisfying all demands over the period. In this case, the city Bukavu is set as the base of relief resources, and the demand cities include Beni, Butembo, Katwa, Kalunguta, Mabalako and Oicha.

The proposed models and algorithms are solved using the Gurobi 9.0.1 solvers, and tested on a Windows10 (64) computer with Intel Core-i7 CPU and 16.0GB RAM.

5.2 Results on Convergence and Optimality

The first experiment presents the performance of Algorithm 1 on convergence and optimality. We run 100 instances of MonERP on DRC Congo scenario with different $\boldsymbol{d}$. Let the relief resources budget $b_1 = 0.5\mathcal{W}$, the transportation budget $b_2 = 0.4\mathcal{C}$, and the attacker's movement cost $\boldsymbol{f}_t \in [1, 3\bar{\boldsymbol{c}}]$. As shown in Figure 3, the results of average regret over stages illustrate that the expected optimality gap between the offline and online revenue converges to near zero as the number of stage n increase to $+\infty$. These experimental results numerically validates Corollary 1.

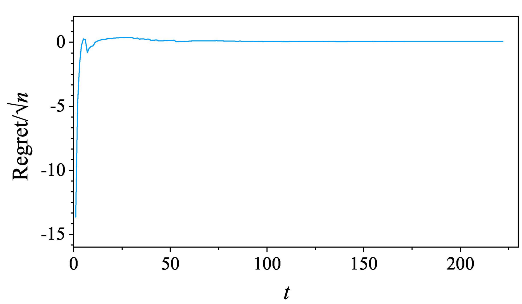


Figure 3: Experimental Results on Regret over Stages

In the second experiment (cf. Figure 4), we compare the time consumption T of obtaining the optimal offline and near-optimal online solutions as the increase of stage number. Apparently, the proposed online learning algorithm is more time-efficient than the offline algorithm, which shows the advantage of this work tackling the far-sighted emergency relief planning under intentional attacks. Note that, the optimal offline results cannot be obtained by the rescuer since the data and parameters are revealed in an online manner.

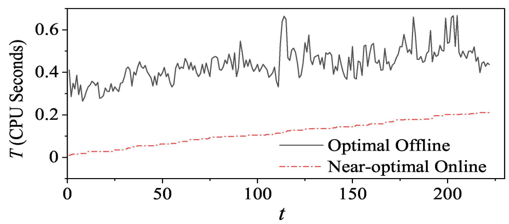


Figure 4: Comparison of Time Consumption as the Increase of n

5.3 Analysis of Impact Factors on Humanitarian Relief Planning

The third experiment focuses on the impact of relief resources and transportation budget on the amount of demand satisfaction as shown in Figure 5. We can observe a saturation critical boundary of their impacts. For instance, if we fix the ratio of relief resources budget at 50%, the increase of the transportation budget over 40% will not bring any growth on the amount of demand satisfaction.

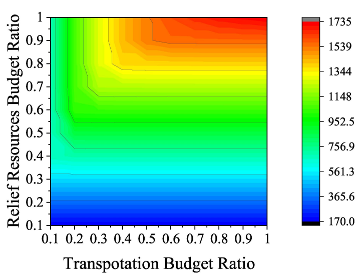


Figure 5: The Impact of Relief Resources and Transportation Budget on Optimal Demand Satisfaction

The final experiment explores the impact of relief resources and transportation budget on the amount of robbed relief resources by the attacker as shown in Figure 6. It is clear that the increase of transportation budget has almost no impact on the amount of robbed relief resources by the attacker when the ratio of relief resources budget is below than 30%. However, when the relief resources budget is relatively sufficient (with a ratio more than 60% in this case), the increase of transportation budget first raises the amount of robbed relief resources to a peak value and then decreases it. This illustrates excessive rescue operations including inappropriate relief resources providing and transport capacity using in conflict areas may breed attackers.

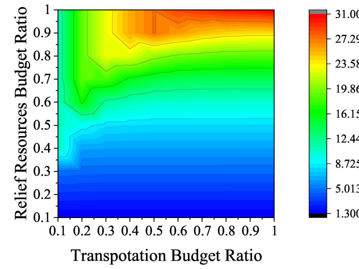


Figure 6: The Impact of Relief Resources and Transportation Budget on the Amount of Robbed Resources by Attacker

6 Conclusion

To meet the realistic need of far-sighted emergency humanitarian relief planning in conflict areas where intentional attacks pose a serious risk on the security and efficiency of rescuing operations, we propose the problem of multi-stage online emergency relief planning and design a efficient online learning approach for the first time. Specifically, we first propose an offline multi-stage optimization problem of emergency relief planning under intentional attacks, in which all parameters in the game between the rescuer and attacker are supposed to be known or predictable. Then, an online version of this problem is introduced to meet the need of online and irrevocable decision making when those parameters are revealed in an online fashion. The modeling process is integrated with the stage-based zero-sum simultaneous game between the rescuer and attacker. To achieve a far-sighted emergency relief planning under attacks, we design an online learning approach which achieves a $O(\frac{1}{\sqrt{n}})$ average regret when those online revealed parameters are generated *i.i.d.* from an unknown distribution. Finally, extensive experiments on a real anti-Ebola relief planning case based on the data of Ebola outbreak and armed attacks in DRC Congo show the scalability and effectiveness of our approach. This work might inspire more efforts in the field of AI on realistic data-driven humanitarian operations research for the good of those citizens suffering from both disasters and conflicts.

References

[ACLED, 2020] ACLED. Conflict data in nord-kivu and sud-kivu province of dr congo, 2020.

[Allahviranloo *et al.*, 2014] Mahdieh Allahviranloo, Joseph Y.J. Chow, and Will W. Recker. Selective vehicle routing problems under uncertainty without recourse. *Transportation Research Part E: Logistics and Transportation Review*, 62:68–88, 2014.

[AWSD, 2018] AWSD. Major attacks on aid workers: Summary statistics (2008-2018), 2018.

[Balcik and Yanıkoğlu, 2020] Burcu Balcik and İhsan Yanıkoğlu. A robust optimization approach for humanitarian needs assessment planning under travel time uncertainty. *European Journal of Operational Research*, 282(1):40–57, apr 2020.

[Balseiro and Gur, 2019] Santiago R Balseiro and Yonatan Gur. Learning in repeated auctions with budgets: Regret minimization and equilibrium. *Management Science*, 65(9):3952–3968, 2019.

[Balseiro *et al.*, 2020] Santiago Balseiro, Haihao Lu, and Vahab Mirrokni. Dual mirror descent for online allocation problems. In *International Conference on Machine Learning*, pages 613–628. PMLR, 2020.

[Besiou and van Wassenhove, 2020] Maria Besiou and Luk N. van Wassenhove. Humanitarian operations: A world of opportunity for relevant and impactful research. *Manufacturing and Service Operations Management*, 22(1):135–145, 2020.

[Borrero *et al.*, 2019] Juan Sebastian Borrero, Oleg A Prokopyev, and Denis Sauré. Sequential interdiction with incomplete information and learning. *Operations Research*, 67(1):72–89, 2019.

[Buchbinder and Naor, 2009] Niv Buchbinder and Joseph Naor. Online primal-dual algorithms for covering and packing. *Mathematics of Operations Research*, 34(2):270–286, 2009.

[Chaudhri *et al.*, 2019] Simran Chaudhri, Kristina Cordes, and Nathan Miller. Humanitarian health programming and monitoring in inaccessible conflict settings: a literature review. *Journal of International Humanitarian Action*, 4(1), dec 2019.

[Correa *et al.*, 2017] José Correa, Tobias Harks, Vincent JC Kreuzen, and Jannik Matuschke. Fare evasion in transit networks. *Operations research*, 65(1):165–183, 2017.

[De Vries and Van Wassenhove, 2020] Harwin De Vries and Luk N. Van Wassenhove. Do Optimization Models for Humanitarian Operations Need a Paradigm Shift? *Production and Operations Management*, 29(1):55–61, 2020.

[Ekici and Özener, 2020] Ali Ekici and Okan Örsan Özener. *Inventory routing for the last mile delivery of humanitarian relief supplies*. Springer Berlin Heidelberg, 2020.

[Fulkerson and Harding, 1977] D Ray Fulkerson and Gary C Harding. Maximizing the minimum source-sink path subject to a budget constraint. *Mathematical Programming*, 13(1):116–118, 1977.

[Hoelscher *et al.*, 2017] Kristian Hoelscher, Jason Miklian, and Håvard Mokleiv Nygård. Conflict, Peacekeeping, and Humanitarian Security: Understanding Violent Attacks Against Aid Workers. *International Peacekeeping*, 24(4):538–565, aug 2017.

[Huang and Song, 2018] Xiaoxia Huang and Liying Song. An emergency logistics distribution routing model for unexpected events. *Annals of Operations Research*, 269(1-2):223–239, 2018.

[Israeli and Wood, 2002] Eitan Israeli and R. Kevin Wood. Shortest-Path Network Interdiction. *Networks*, 40(2):97–111, sep 2002.

[Li *et al.*, 2020] Xiaocheng Li, Chunlin Sun, and Yinyu Ye. Simple and fast algorithm for binary integer and online linear programming. *arXiv preprint arXiv:2003.02513*, 2020.

[Ozdamar *et al.*, 2018] Linet Ozdamar, Dilek Tuzun Aksu, Elifcan Yasa, and Biket Ergunes. Disaster relief routing in limited capacity road networks with heterogeneous flows. *Journal of Industrial and Management Optimization*, 14(4):1367–1380, 2018.

[Sabouhi *et al.*, 2019] Fatemeh Sabouhi, Ali Bozorgi-Amiri, Mohammad Moshref-Javadi, and Mehdi Heydari. An integrated routing and scheduling model for evacuation and commodity distribution in large-scale disaster relief operations: a case study. *Annals of Operations Research*, 283(1-2):643–677, dec 2019.

[Schwartz, 2020] David A. Schwartz. Being Pregnant during the Kivu Ebola Virus Outbreak in DR Congo: The rVSV-ZEBOV Vaccine and Its Accessibility by Mothers and Infants during Humanitarian Crises and in Conflict Areas. *Vaccines*, 8(1):38, jan 2020.

[Sinha *et al.*, 2018] Arunesh Sinha, Fei Fang, Bo An, Christopher Kiekintveld, and Milind Tambe. Stackelberg security games: Looking beyond a decade of success. IJCAI, 2018.

[Smith and Song, 2020] J. Cole Smith and Yongjia Song. A survey of network interdiction models and algorithms. *European Journal of Operational Research*, 283(3):797–811, 2020.

[Usanov *et al.*, 2013] Artur Usanov, Marjolein Ridder, Willem L. Auping, Stephanie Lingemann, Luis A Tercero Espinoza, Magnus Ericcson, Masuma Farooki, Sievers Henrike, and Liedtke Maren. *Coltan, Congo & Conflict*. 2013.

[WHO, 2020] WHO. Ebola virus disease – democratic republic of the congo, 2020.

[Xiao *et al.*, 2018] Kaiming Xiao, Cheng Zhu, Junjie Xie, Yun Zhou, Xianqiang Zhu, and Weiming Zhang. Dynamic defense strategy against stealth malware propagation in cyber-physical systems. In *IEEE INFOCOM 2018-IEEE Conference on Computer Communications*, pages 1790–1798. IEEE, 2018.

[Xu *et al.*, 2017] Kai Xu, Kaiming Xiao, Quanjun Yin, Yabing Zha, and Cheng Zhu. Bridging the gap between observation and decision making: Goal recognition and flexible resource allocation in dynamic network interdiction. In *Twenty-Sixth International Joint Conference on Artificial Intelligence*, pages 4477–4483, 2017.

[Yuan and Wang, 2009] Yuan Yuan and Dingwei Wang. Path selection model and algorithm for emergency logistics management. *Computers and Industrial Engineering*, 56(3):1081–1094, apr 2009.

[Zhang *et al.*, 2019] Youzhi Zhang, Qingyu Guo, Bo An, Long Tran-Thanh, and Nicholas R Jennings. Optimal interdiction of urban criminals with the aid of real-time information. In *Proceedings of the AAAI Conference on Artificial Intelligence*, volume 33, pages 1262–1269, 2019.

[Zhu *et al.*, 2019] Li Zhu, Yeming Gong, Yishui Xu, and Jun Gu. Emergency relief routing models for injured victims considering equity and priority. *Annals of Operations Research*, 283(1-2):1573–1606, 2019.