Connected Placement of Disaster Shelters in Modern Cities

Huanyang Zheng and Jie Wu
Department of Computer and Information Sciences
Temple University, USA
{huanyang.zheng, jiewu}@temple.edu

ABSTRACT

This paper is motivated by the fact that modern cities are surprisingly vulnerable to large-scale emergencies, such as the recent terrorist attacks on Paris that resulted in the death of 130 people. Disaster shelters are one of the most effective methods to handle large-scale emergencies. Hence, this paper establishes disaster shelters with bounded costs. The objective is to minimize the total establishment costs of disaster shelters under three constraints. The first constraint is a distance constraint, which requires that people must be assigned to disaster shelters within a certain range. The second constraint is a capacity constraint, which requires that disaster shelters must have the capacity to hold incoming people. The third constraint is a connection constraint, which requires that disaster shelters should be connected to avoid being isolated. Two bounded algorithms are proposed to efficiently establish disaster shelters. Real data-driven experiments are conducted to demonstrate the efficiency and effectiveness of the proposed algorithms.

CCS Concepts

•**Networks** → Cyber-physical networks;

Keywords

Disaster shelters; smart city; connected placement; greedy algorithm; approximation ratio;

1. INTRODUCTION

Recently, France has witnessed unprecedented attacks on Paris that resulted in the death of 130 people [1]. While effective policies are necessary to resolve this crisis, we are motivated by the fact that modern cities are vulnerable to large-scale emergencies, such as terrorist attacks and refugee crises. Meanwhile, disaster shelters are one of the most effective methods to handle large-scale emergencies. They are defensive buildings with food storages and medical facilities.

CHANTS'16, October 03-07, 2016, New York City, NY, USA

DOI: http://dx.doi.org/10.1145/2979683.2979690

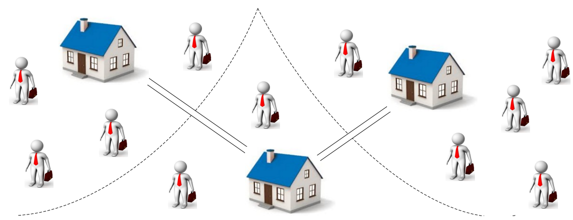


Figure 1: Scenario of the disaster shelter placement.

Hawaii has established underground disaster shelters to protect people from hurricanes [2]. While disaster shelters can effectively save lives in emergencies, few research efforts were made. Establishment costs of disaster shelters are currently unbounded in existing literatures [3]. In contrast, this paper explores the bounded placement of disaster shelters to minimize their establishment costs. Our scenario is shown in Fig. 1, which is a three-dimensional Euclidean surface representing the ground of the city. People are distributed in the Euclidean surface (known a priori). Disaster shelters of different capacities are established with different establishment costs. A large-capacity disaster shelter has a higher cost than a small-capacity one.

The objective of this paper is to minimize the total establishment costs of disaster shelters under three constraints. The number, locations, and capacities of disaster shelters serve as variables that need to be determined. The first constraint (distance constraint) is that people must be assigned to disaster shelters within a certain range. Otherwise, people may not have enough time to reach disaster shelters in large-scale emergencies. The second constraint (capacity constraint) is that disaster shelters must have the capacity to hold incoming people. A crowded disaster shelter can lead to epidemic spreadings, insufficient amount of foods, and evacuation space problems. The third constraint is that disaster shelters should be "connected" to avoid isolations (connection constraint, formally defined later in the problem formulation). This constraint aims to maintain the communication between different disaster shelters, such that searches and rescues can be facilitated. For example, family members separated in different disaster shelters can be found through those communications. Rescue teams can also be supported by nearby disaster shelters.

The remainder of this paper is organized as follows. Section II surveys related works. Section III describes the model and then formulates the problem. Section IV analyzes the problem. Section V proposes bounded solutions. Section VI includes the experiments. Section VII concludes the paper.

2. RELATED WORK

Few research efforts have been made with respect to the placement of the disaster shelters. Park et al. [3] proposed a combinational optimization method to determine locations of tsunami vertical evacuation shelters. Genetic algorithms were used to solve the placement problem without a bounded performance. Gama et al. [4] studied the shelter placement for mitigating urban flood disasters. A coverage model was proposed to maximize a gradual demand coverage function, which represents a trade-off between a full coverage objective and a distance objective. Heßler and Hamacher [5] explored the sink location problem in evacuation planning, such that disaster shelters are reachable for people. Their results were evaluated by numerical tests including random data as well as real world data from the city of Kaiserslautern, Germany. The above works are conducted by the transportation research community. Few bounded results were presented.

Due to the recent development of Delay Tolerant Networks (DTNs), the search and rescue after disasters were studied by the wireless network community. Mase and Okada [6] designed Unmanned Aerial Vehicles (UAVs) for message deliveries in large-scale disaster environments. Messages are sent and received through wireless links between devices and the UAV. DTNs can be built based on the UAVs. Yang et al. [7] focused on the capacity constrained Voronoi diagram, which partitions the network to minimize imbalanced traffic loads among disaster shelters. We additionally consider the communication between disaster shelters. Our disaster shelter placement problem extends the classic set cover problem in terms of the connection constraint. Given some elements and a collection of sets of elements, the classic set cover problem aims to select minimum sets to cover all given elements [8]. Elements in a set are covered if this set is selected. In our problem, disaster shelters are mapped to sets and people are mapped to elements. Geometric set cover techniques [9] are used, while the connection constraint is posed.

3. PROBLEM FORMULATION

Our scenario is based on a three-dimensional Euclidean surface that represents the ground of a city. Let p denote a person that is planned for disaster shelters. P is the set of people, i.e., $P = \{p\}$. Locations of people are known a priori. Let s denote a disaster shelter. S is the set of the disaster shelters, i.e., $S = \{s\}$. Then, the location, capacity and establishment cost of s are denoted by l_s, c_s, and e_s, respectively. We consider that the capacity of a disaster shelter is bounded, i.e., $c_{\min} \leq c_s \leq c_{\max}$. The establishment cost of a disaster shelter depends on its capacity. Clearly, a large-capacity disaster shelter has a higher establishment cost than a small-capacity one. We use $e_s = F(c_s)$ to denote the pre-known establishment cost function. Let $x_{p,s}$ denote a boolean decision variable. $x_{p,s} = 1$ means that the person, p, is assigned to the disaster shelter, s. Finally, $D(\cdot)$ is a pre-known distance function.

Our objective is to minimize the total establishment costs, $\sum_s e_s$, of the disaster shelters. The number, locations, and capacities of the disaster shelters (i.e., $|S|$, l_s, and c_s) are variables. We determine the assignment plan of $x_{p,s}$. Three constraints are posed. The first constraint (distance constraint) is that people must be assigned to disaster shelters within a certain range of r, i.e., $\sum_s x_{p,s} = 1$ and $D(p,s) \cdot x_{p,s} \leq r$. This constraint guarantees that people have enough time to reach disaster shelters in large-scale emergencies. The second constraint (capacity constraint) is that disaster shelters must have the capacity to hold incoming people. We have $\sum_p x_{p,s} \leq c_s$ to ensure sufficient resources for people in disaster shelters. The third constraint is that disaster shelters should be "connected" to avoid isolations. We say two disaster shelters are neighboring, if the distance between them is no larger than a threshold that is denoted by R. The third constraint (connection constraint) means that the induced neighboring graph of the disaster shelters includes all disaster shelters and is connected. This aims to maintain temporary communication between different disaster shelters, since disasters may destroy wired and wireless communication infrastructures. Vehicles equipped with wireless devices can move around disaster shelters to provide temporary communication based on DTNs.

As a trade-off, we need to balance the number and capacities of the disaster shelters. For the same cost, we can either establish many small-capacity disaster shelters or a few large-capacity disaster shelters, depending on the distribution of the people. This trade-off becomes even more complex, when the location problem is involved. Since we require that disaster shelters should be connected, the establishment of a disaster shelter can be used to hold people or serves as an intermediate relay to connect other disaster shelters. Hence, our problem faces unique challenges.

4. PROBLEM ANALYSIS

We start with the problem hardness:

Theorem 1. *Our placement problem is NP-hard.*

Proof: The proof is done through two special assumptions, under which our problem becomes equivalent to the geometry set cover problem [10]. Given some elements and a collection of sets of elements, the set cover problem aims to select a minimum number of sets to cover all given elements. The geometry set cover problem is a special case, in which elements are distributed in the Euclidean space and sets are geometric balls. An element is covered by a set, if it is geometrically included by the ball of that set.

Let us consider the disaster shelter placement problem in the Euclidean space. The first assumption is that the establishment cost of a disaster shelter barely depends on its capacity. As a result, the capacity of each shelter can be large enough to hold all incoming people, i.e., the second constraint (capacity constraint) is relaxed. The second assumption is that R is large enough, such that we can relax the third constraint (connection constraint). When R goes to infinity, all disaster shelters are neighboring and connected. Then, let us reduce the geometry set cover problem to our disaster shelter placement problem. This reduction is done by mapping elements and sets to people and disaster shelters, respectively. The geometric ball of each set has a radius of r, i.e., the disaster shelter will hold all people within the range of r. Our problem can also reduce to the geometry set cover problem in the same manner (they become equivalent to each other). Since the geometry set cover problem is NP-hard, our problem is also NP-hard. ■

We apply a discrete framework to solve the disaster shelter placement problem. The three-dimensional Euclidean surface, which models the ground of the city, is horizontally discretized by a grid, i.e., the vertical projection of the discretized surface would be a grid at the horizontal plane. An

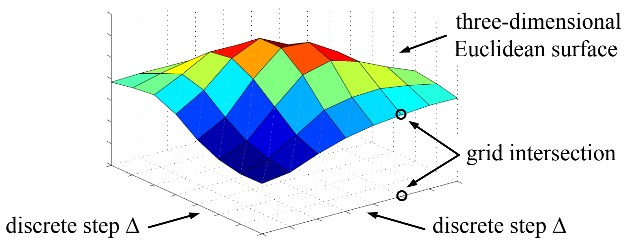


Figure 2: An illustration of the discrete framework.

example of the discretization is shown in Fig. 2. Then, each square in the grid has a size of $\Delta \times \Delta$, in which Δ is the discrete step measuring the distance between neighboring grid intersections. Let L denote the set of locations at grid intersections on the discretized three-dimensional Euclidean surface (the ground of the city). We assume that disaster shelters are only established at locations that belong to L. Since the slope of the ground is limited, such a discretization brings a limited information loss. Δ brings a trade-off between the accuracy and the time complexity. A smaller Δ brings a higher accuracy at the cost of a higher time complexity, since the set L becomes larger. In addition, the establishment cost function, $e_s = F(c_s)$, is also discretized. This is because the shelter capacity is discretized, i.e., the number of people that it can hold is an integer. We have:

Definition 1. *The coverage radius, γ_s, of a disaster shelter, s, is the maximum possible distance that satisfies: (i) $\gamma_s \leq r$, and (ii) s has a sufficient capacity to hold all people within γ. People within the coverage radius of s are covered by s, and the set of covered people is denoted by P_s.*

In Definition 1, r comes from the distance constraint, in which people must be assigned to disaster shelters within a range of r. A disaster shelter will not cover people without the range of r. In addition, the coverage radius is restricted by the capacity constraint, in which disaster shelters must have the capacity to hold incoming people. Note that the area within the coverage radius of a shelter is not necessarily a disk, since the ground of the city may not be flat. For a disaster shelter at a given location, its coverage radius can be determined once its capacity is known.

5. BOUNDED SOLUTIONS

Two bounded algorithms are proposed for our problem. The first and second algorithms independently and cooperatively consider the connection constraint, respectively.

5.1 Two-Stage Algorithm

We propose a Two-Stage Algorithm (TSA) to place disaster shelters. The first stage resolves the distance and capacity constraints, and then, the second stage places additional disaster shelters to satisfy the connection constraint. TSA is presented as Algorithm 1. Lines 1 to 4 correspond to the first stage. In lines 1 to 3, disaster shelters at all possible locations with each possible capacity are mapped to a collection of sets. People are mapped elements, and people that are covered by a disaster shelter are mapped to elements in the corresponding set. In line 4, an existing approximation algorithm is used to solve the geometric set cover problem.

Algorithm 1 Two-Stage Algorithm (TSA)

Input: A set of people, P, and their locations,
An establishment cost function, $F(\cdot)$,
A distance function on Euclidean surface, $D(\cdot)$,
A set of discrete locations, L,
A range, r, and a connection threshold, R.

Output: S, their locations, and capacities.

1: **for** each location, $l \in L$ **do**
2: **for** each discrete capacity, $c_{\min} \leq c \leq c_{\max}$ **do**
3: Calculate the set of covered people, P_s, for the disaster shelter, s, with $l_s = l$ and $c_s = c$. Map s to a set and P_s to elements in s. Use the establishment cost, $e_s = F(c_s)$, as the set weight.
4: For the mapped sets and elements, use an approximation algorithm to solve the geometric set cover problem. The resultant collection of selected sets are mapped back to the disaster shelters, including locations and capacities.
5: Map locations in L as nodes in a graph (nodes are connected if their distance is no larger than R). Use an approximation algorithm to solve a Steiner tree problem, which selects nodes to connect the terminals (disaster shelters in line 4). Selected nodes are recorded as new disaster shelters with minimum capacities.
6: **return** the resultant disaster shelters in lines 4 and 5.

The sets, which are selected by the approximation algorithm, are mapped back to disaster shelters at given locations with given capacities. The resultant disaster shelters satisfy the distance and capacity constraints, but may not satisfy the connection constraint. To satisfy the connection constraint, we introduce the second stage, which includes line 5. The resultant disaster shelters of the first stage are regarded as terminals in a Steiner tree problem. Through another existing approximation algorithm, they can be connected by placing additional disaster shelters with minimum capacities. TSA is bounded:

Theorem 2. *Let a_1 and a_2 denote the approximation ratios of existing algorithms for the geometric set cover problem and the Steiner tree problem, respectively. Then, the approximation ratio of TSA is $a_1 + a_2 \times (1 + \lfloor 2r/R \rfloor \times a_1)$.*

Proof: Let TSA_1 and TSA_2 denote the cost of the disaster shelters placed by the first and second stages of TSA, respectively. Let OPT_1 and OPT_2 denote the optimal cost of the disaster shelters for the first and second stages of TSA, respectively. Let OPT denote the optimal cost of the disaster shelters in our problem. Based on the definition of the approximation ratio, we have the following inequalities:

$$\text{OPT}_1 \leq \text{TSA}_1 \leq a_1 \times \text{OPT}_1 \tag{1}$$

$$\text{OPT}_2 \leq \text{TSA}_2 \leq a_2 \times \text{OPT}_2 \tag{2}$$

Since the optimal placement in our problem is also a solution to the geometric set cover problem in the first stage, we have:

$$\text{TSA}_1 \leq a_1 \times \text{OPT}_1 \leq a_1 \times \text{OPT} \tag{3}$$

Let us consider a special solution to the Steiner tree problem in the second stage, based on the optimal placement in our problem. Since people must be assigned to disaster shelters within a certain range of r, the distance between a disaster shelter in TSA_1 and its closest disaster shelter in OPT must be no larger than $2r$. Consequently, we can use $\lfloor 2r/R \rfloor$ disaster shelters with minimum capacities (and thus minimum establishment costs) to connect an arbitrary disaster shelter in TSA_1 to disaster shelters in OPT. Hence, we can use a total cost of $\text{OPT} + \lfloor 2r/R \rfloor \times \text{OPT}_1$ as a special solution to the Steiner tree problem. The cost of this special solution is no smaller than OPT_2 by the definition. We have:

$$\text{TSA}_2 \leq a_2 \times \text{OPT}_2 \leq a_2 \times (\text{OPT} + \lfloor \frac{2r}{R} \rfloor \times \text{TSA}_1) \quad (4)$$

If we combine Eqs 3 and 4, we have:

$$\begin{aligned} \text{TSA}_1 + \text{TSA}_2 &= \text{TSA}_1 + a_2 \times (\text{OPT} + \lfloor \frac{2r}{R} \rfloor \times \text{TSA}_1) \\ &\leq \big[a_1 + a_2 \times (1 + \lfloor \frac{2r}{R} \rfloor \times a_1)\big] \times \text{OPT} \quad (5) \end{aligned}$$

The proof completes. ■

Let $|P|$ denote the number of people (set cardinality of P). Since there exist approximation algorithms [10, 11] with $a_1 = \log |P|$ and $a_2 = \frac{3}{2}$, we have the following corollary:

Corollary 1. *The approximation ratio of TSA can be* $\frac{3}{2} + (1 + \frac{3}{2}\lfloor \frac{2r}{R} \rfloor) \cdot \log |P|$. *When* $2r < R$, *it is* $\frac{3}{2} + \log |P|$.

The bound of TSA depends on $\lfloor 2r/R \rfloor$, which represents the tightness of the connection constraint. When R is large enough, the connection constraint can be relaxed, and then, we can focus on the distance and capacity constraints. The time complexity of TSA depends on the time complexities of approximation algorithms for the geometric set cover problem and the Steiner tree problem. TSA solves the distance and capacity constraints by the geometric set cover problem, and then, solves the connection constraint by the Steiner tree problem. TSA solves constraints independently.

5.2 Tree Growth Algorithm

This subsection presents a Tree Growth Algorithm (TGA) to place disaster shelters. The distance, capacity, and connection constraints are cooperatively considered. The idea of TGA is to iteratively place a disaster shelter that can connect to an existing disaster shelter, until all people are covered. We start with the following definition:

Definition 2. *The adjacent location is defined based on a set of the disaster shelters,* S. l *is adjacent to* S, *if it is (i) in the range of* R *of at least one disaster shelter in* S, *and (ii) not in the coverage radius of a disaster shelter in* S.

The coverage radius is in Definition 1, and R is the threshold for the connection constraint. If we place a disaster shelter at an adjacent location of S, then it is connected to S, and therefore satisfying the connection constraint. Meanwhile, since people in adjacent locations of S are not covered by S, the placement of a disaster shelter at an adjacent location can also provide efficient coverage to people (the distance and capacity constraints are stisfied). To quantify the connection constraint, we slightly abuse the notation, and use $D(l, S)$ to denote the distance between l and its closest disaster shelter in S. If l is adjacent to S, then we have $\min_s \gamma_s \leq D(l, S) \leq R$.

TGA is presented in Algorithm 2. In line 1, it initializes $S = \emptyset$, i.e., no disaster shelter has been placed. In line 2, among all the locations, it places the first disaster shelter that minimizes $\frac{F(c_s)}{|P_s|}$. Here, $e_s = F(c_s)$ is the establishment cost

Algorithm 2 Tree Growth Algorithm (TGA)

Input: A set of people, P, and their locations,
An establishment cost function, $F(\cdot)$,
A distance function on Euclidean surface, $D(\cdot)$,
A set of discrete locations, L,
A range, r, and a connection threshold, R.
Output: S, their locations, and capacities.

1: Initialize $S = \emptyset$.
2: Among locations $l \in L$, place a disaster shelter, s, with a capacity, c_s that minimizes $\frac{F(c_s)}{|P_s|}$.
3: Add s to S, and then, remove P_s from P.
4: **while** $P \neq \emptyset$ **do**
5: Find the set of locations, L', that are adjacent to S.
6: Among locations $l \in L'$, place a disaster shelter, s, with a capacity, c_s, that minimizes $\frac{F(c_s) \cdot D(l,S)}{|P_s|}$.
7: Add s to S, and then, remove P_s from P.
8: **return** S, their locations, and capacities.

of a disaster shelter with a capacity, c_s, and $|P_s|$ is the number of people covered by the disaster shelter, s. Therefore, $\frac{F(c_s)}{|P_s|}$ represents the "cost-to-benefit" ratio, which should be minimized. Note that $\frac{F(c_s)}{|P_s|}$ involves the distance and capacity constraints, but not the connection constraint. This is because line 2 only places the first disaster shelter. Line 3 updates the disaster shelter placement, and removes covered people. Lines 4 to 7 iteratively place a disaster shelter, until all people are covered. Line 5 calculates the set of adjacent locations of the existing disaster shelters, S. Among all the adjacent locations of S, line 6 places a disaster shelter that minimizes $\frac{F(c_s) \cdot D(l,S)}{|P_s|}$, which represents an improved "cost-to-benefit" ratio. Here, $D(l, S)$ is the distance between l and its closest disaster shelter in S. It additionally incorporates the connection constraint into $\frac{F(c_s)}{|P_s|}$. Line 7 updates the disaster shelter placement, and removes covered people. The iteration terminates when all people are covered.

In TGA, the distance and capacity constraints are satisfied by the definition of the coverage radius. Meanwhile, the connection constraint is satisfied by the definition of the adjacent location. Let $\min_s \gamma_s$ denote the minimum coverage radius in TGA. Then, TGA is bounded as follows:

Theorem 3. *In terms of minimizing the total establishment cost, TGA has an approximation ratio of* $\frac{R}{\min_s \gamma_s} \log |P|$ *to the optimal disaster shelter placement algorithm.*

Proof: A new concept of the person's weight is introduced to prove this theorem. TGA is an iterative algorithm, and in each iteration, a disaster shelter is established to minimize $\frac{F(c_s) \cdot D(l,S)}{|P_s|}$. Note that people, who have already been covered in previous iterations, are not included in P_s. The weight of each person in P_s is defined as $\frac{F(c_s) \cdot D(l,S)}{|P_s|}$, and is denoted by w_p. To unify line 2, the weight of each person covered by the first disaster shelter is specially defined as $\frac{F(c_s) \cdot R}{|P_s|}$. Then, the total weights of people in TGA is:

$$\begin{aligned} \sum_{p \in P} w_p = \sum_s \sum_{p \in P_s} w_p &= \sum_s \sum_{p \in P_s} \frac{F(c_s) \cdot D(l, S)}{|P_s|} \\ &\geq \min_s \gamma_s \times \sum_s F(c_s) \quad (6) \end{aligned}$$

This is because $\sum_{p\in P_s}\frac{1}{|P_s|}=1$. We have $D(l,S)\geq \min_s \gamma_s$ based on the definition of adjacent locations. Let s^* denote a disaster shelter in the optimal solution, S^*. We have:

$$\sum_{p\in P} w_p = \sum_{s^*}\sum_{p\in P_{s^*}} w_p \tag{7}$$

We claim that, for each disaster shelter in the optimal solution, the following inequality is satisfied:

$$\sum_{p\in P_{s^*}} w_p \leq \log|P|\cdot F(c_{s^*})\cdot R \tag{8}$$

To prove our claim, let us consider how TGA covers people in P_{s^*}. Let n_k be the number of uncovered people in P_{s^*}, after the k-th iteration. For a people, p, who is covered by s in TGA and by s^* in the optimal solution, we have:

$$w_p = \frac{F(c_s)\cdot D(l,S)}{|P_s|} \leq \frac{F(c_{s^*})\cdot D(l,S^*)}{|P_{s^*}|} \leq \frac{F(c_{s^*})\cdot R}{n_k} \tag{9}$$

The first inequality results from the greediness, in which TGA minimizes $\frac{F(c_s)\cdot D(l,S)}{|P_s|}$ among all possible placements. The second inequality results from $|P_{s^*}|\geq n_k$ and $D(l,S^*)\leq R$. As a result, we have [11]:

$$\begin{aligned}\sum_{p\in P_{s^*}} w_p &\leq \sum_k (n_k - n_{k+1})\frac{F(c_{s^*})\cdot R}{n_k}\\ &\leq \log|P|\cdot F(c_{s^*})\cdot R\end{aligned} \tag{10}$$

The proof of Eq. 8 completes. Combining Eqs. 6, 7, and 8, the following bound can be obtained:

$$\begin{aligned}\min_s \gamma_s \times \sum_s F(c_s) \leq \sum_{p\in P} w_p &= \sum_{s^*}\sum_{p\in P_{s^*}} w_p\\ &\leq \sum_{s^*}\log|P|\cdot F(c_{s^*})\cdot R\end{aligned} \tag{11}$$

Eq. 11 can be rewritten as:

$$\sum_s F(c_s) \leq \frac{R}{\min_s \gamma_s}\log|P| \times \sum_{s^*} F(c_{s^*}) \tag{12}$$

Note that $\sum_s F(c_s)$ and $\sum_{s^*} F(c_{s^*})$ are the total establishment costs of TGA and the optimal solution, respectively. Therefore, the proof completes. ■

TGA's bound depends on its minimum coverage radius, which in turn depends on the distribution of the people and the establishment cost function, $F(\cdot)$. When $F(\cdot)$ is a constant function, the establishment cost of a disaster shelter is a constant, and is not related to the capacity. We have $\min_s \gamma_s = r$, since disaster shelters can have sufficient and free capacities to hold all people within r. Hence, we have:

Corollary 2. *The approximation ratio of TGA is $\frac{R}{r}\log|P|$, when $F(\cdot)$ is a constant establishment cost function.*

The time complexity of TGA is $O(|L|^2 + |L||P||c_{\max} - c_{\min}|)$. $|L|$ is the number of locations, $|P|$ is the number of people, and $|c_{\max}-c_{\min}|$ is the number of different capacities. This time complexity can be obtained by pre-computations. $O(|L|^2)$ comes from pre-computations of the distances between all pairs of locations in L, in order to determine the adjacent location. $O(|L||P||c_{\max} - c_{\min}|)$ comes from pre-computations of the coverage radius for all possible disaster shelters with all possible capacities.

6. EXPERIMENTS

6.1 Settings

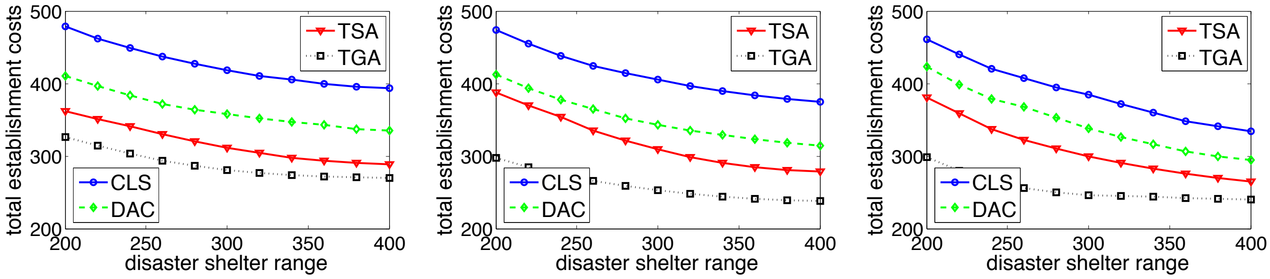
Our dataset is based on the city of Cannon Beach, Oregon, United States [12]. The dataset information has been illustrated in Park's work [12]. It includes a long and narrow area (6.1km by 1.5km). The locations of 1,382 houses are collected. For simplicity, we assume each house corresponds to a single person in our experiments. Algorithms 1 and 2 are denoted as TSA and TGA, respectively. Two baseline algorithms are used as comparisons. (i) CLS stands for a bounded clustering method in [13]. To guarantee that each person is within in a range, r, of the disaster shelter, the radius of each cluster is r. Disaster shelters are established at the center of each cluster to satisfy the distance and capacity constraints. To satisfy the connection constraint, a spanning tree is additionally built. (ii) DAC is a divide and conquer method. DAC iteratively divides people into two clusters through a k-means algorithm [14]. The centers of these two clusters are connected through placing disaster shelters with minimum capacities. The iteration terminates when the distance and capacity constraints are satisfied.

We set $\Delta = 100$m to discretize the locations. Note that the dataset includes a long and narrow area of 6.1km by 1.5km. Therefore, such a discretization could maintain the accuracy. To reveal the capacity constraint, we set $c_{\min} = 10$ and $c_{\max} = 100$. Two establishment cost functions are used: $F(c_s) = \sqrt{c_s}$ and $F(c_s) = 2.4 + 0.076\times c_s$. Coefficients of 2.4 and 0.076 are used, since $F(c_{\min})$ and $F(c_{\max})$ can be the same for the first and second establishment cost functions. Therefore, fair comparisons can be obtained.

We do not use hyper-linear establishment cost functions, since the establishments of multiple small-capacity disaster shelters can be always better than those of a few large-capacity disaster shelters. Moreover, hyper-linear establishment cost functions are also empirically impractical. Parameters, r and R, are tuned to represent the impacts of the distance and connection constraints, respectively. While r ranges from 200m to 400m, we set R to be 400m, 600m, and 800m. In this setting, we have $R \geq r$ to guarantee that the connection constraint is not ignored or relaxed.

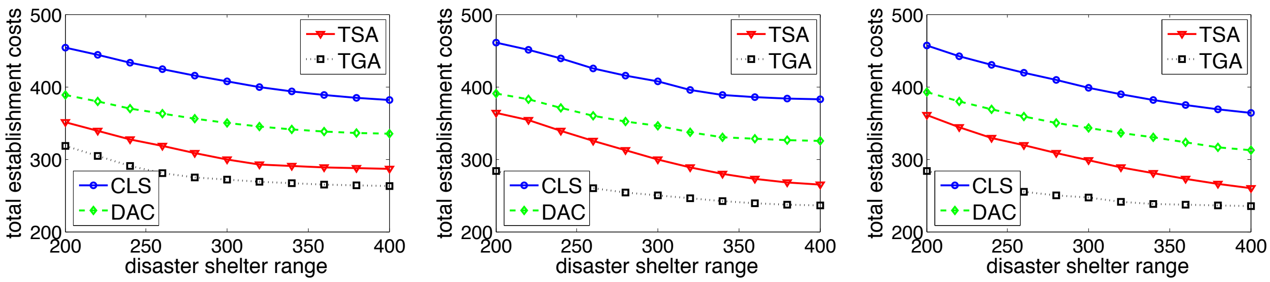
6.2 Evaluation Results

The evaluation results under $F(c_s) = \sqrt{c_s}$ are shown in Fig. 3. Three subfigures of Fig. 3(a), Fig. 3(b), and Fig. 3(c) correspond to $R = 400$m, $R = 600$m, and $R = 800$m for the connection constraint, respectively. TSA and TGA outperform CLS and DAC under all the settings. TGA outperforms TSA, since it cooperatively considers the connection constraint. For all algorithms, the total establishment cost, $\sum_s e_s$, decreases with respect to the disaster shelter range, r. The distance constraint requires that people must be assigned to disaster shelters within a certain range of r. A larger r means that the distance constraint is more relaxed, and thus, the total establishment cost is smaller. When r goes to infinity, the distance constraint is completely relaxed. Another notable point is that, when the connection threshold, R, becomes larger, the total establishment cost also slightly decreases. Similarly, a larger R means that the connection constraint can be more relaxed, and thus, the total establishment cost is smaller. When R goes to infinity, the connection constraint can be ignored. As shown in Fig. 3, the performance gap between TSA and TGA is small when



(a) $R = 400$m. (b) $R = 600$m. (c) $R = 800$m.

Figure 3: Evaluation results under $F(c_s) = \sqrt{c_s}$.



(a) $R = 400$m. (b) $R = 600$m. (c) $R = 800$m.

Figure 4: Evaluation results under $F(c_s) = 2.4 + 0.076 \times c_s$.

$R = 400$m, but is large when $R = 600$m and $R = 800$m. This is because TSA and TGA independently and cooperatively consider the connection constraint, respectively.

The evaluation results under $F(c_s) = 2.4 + 0.076 \times c_s$ are shown in Fig. 4. Three subfigures of Fig. 4(a), Fig. 4(b), and Fig. 4(c) correspond to $R = 400$m, $R = 600$m, and $R = 800$m for the connection constraint, respectively. The results in Fig. 4 are similar to results in Fig. 3. The only difference is that the total establishment costs of TSA and TGA decrease more slowly in Fig. 4. Under $F(c_s) = \sqrt{c_s}$, we significantly favor the establishments of large-capacity disaster shelters, since they are cost-efficient, especially when r is large (a disaster shelter can cover more people due to the distance constraint). On the other hand, when $F(c_s) = 2.4 + 0.076 \times c_s$, we only slightly favor the establishments of large-capacity disaster shelters, since the establishment cost function scales linearly with respect to the disaster shelter capacity.

7. CONCLUSION

This paper minimizes the total establishment costs of the disaster shelters under three constraints: the distance constraint, the capacity constraint, and the connection constraint. The distance constraint requires that people must be assigned to disaster shelters within a certain range. The capacity constraint requires that disaster shelters must have the capacity to hold incoming people. The connection constraint requires that disaster shelters should be connected to avoid isolations. Two bounded algorithms are proposed to establish disaster shelters. Experiments demonstrate the efficiency and effectiveness of the proposed algorithms.

ACKNOWLEDGEMENT

This research was supported in part by NSF grants CNS 1449860, CNS 1461932, CNS 1460971, CNS 1439672, CNS 1301774, and ECCS 1231461.

8. REFERENCES

[1] A. Sazonov, "Paris attacks and europeańs response to the refugee crisis," *Huffpost Politics*, 2015.

[2] K. Kim, P. Pant, and E. Yamashita, "Evacuation planning for plausible worst case inundation scenarios in Honolulu, Hawaii." *Journal of Emergency Management*, 2014.

[3] S. Park, J. W. Van de Lindt, R. Gupta, and D. Cox, "Method to determine the locations of tsunami vertical evacuation shelters," *Natural Hazards*, 2012.

[4] M. Gama, M. P. Scaparra, and B. Santos, "Optimal location of shelters for mitigating urban floods," in *European Working Group on Transportation*, 2014.

[5] P. Heßler and H. W. Hamacher, "Sink location to find optimal shelters in evacuation planning," *EURO Journal on Computational Optimization*, 2014.

[6] K. Mase and H. Okada, "Message communication system using unmanned aerial vehicles under large-scale disaster environments," in *IEEE PIMRC*, 2015.

[7] K. Yang, A. H. Shekhar, D. Oliver, and S. Shekhar, "Capacity-constrained network-voronoi diagram," *IEEE Transactions on Knowledge and Data Engineering*, 2015.

[8] A. Deshpande, L. Hellerstein, and D. Kletenik, "Approximation algorithms for stochastic boolean function evaluation and stochastic submodular set cover," in *ACM-SIAM SODA*, 2014.

[9] S. Stergiou and K. Tsioutsiouliklis, "Set cover at web scale," in *ACM SIGKDD*, 2015.

[10] K. L. Clarkson and K. Varadarajan, "Improved approximation algorithms for geometric set cover," *Discrete & Computational Geometry*, 2007.

[11] Z. Zhang, X. Gao, and W. Wu, "Algorithms for connected set cover problem and fault-tolerant connected set cover problem," *Theoretical Computer Science*, 2009.

[12] S. Park, "Understanding and mitigating tsunami risk for coastal structures and communities," *PhD Thesis, Colorado State University*, 2007.

[13] T. F. Gonzalez, "Clustering to minimize the maximum intercluster distance," *Theoretical Computer Science*, 1985.

[14] J. Wang, J. Wang, Q. Ke, G. Zeng, and S. Li, "Fast approximate k-means via cluster closures," in *Multimedia Data Mining and Analytics*, 2015.