# A Social Distancing-Based Facility Location Approach for Combating COVID-19

Suman Banerjee[1(✉)], Bithika Pal[2], and Maheswar Singhamahapatra[3]

[1] Indian Institute of Technology, Jammu, India
suman.banerjee@iitjammu.ac.in
[2] Indian Institute of Technology, Kharagpur, India
bithikapal@iitkgp.ac.in
[3] School of Business, FLAME University, Pune, India

**Abstract.** In this paper, we introduce and study the problem of facility location along with the notion of *'social distancing'*. The input to the problem is the road network of a city where the nodes are the residential zones, edges are the road segments connecting the zones along with their respective distance. We also have the information about the population at each zone, different types of facilities to be opened and in which number, and their respective demands in each zone. The goal of the problem is to locate the facilities such that the people can be served and at the same time the total social distancing is maximized. We formally call this problem as the Social Distancing-Based Facility Location Problem. We mathematically quantify social distancing for a given allocation of facilities and proposed an optimization model. As the problem is NP-Hard, we propose a simulation-based and heuristic approach for solving this problem. A detailed analysis of both methods has been done. We perform an extensive set of experiments with synthetic datasets. From the results, we observe that the proposed heuristic approach leads to a better allocation compared to the simulation-based approach.

**Keywords:** Facility location · Social distancing · Integer programming · COVID-19

## 1 Introduction

In recent times, the entire world is witnessing the pandemic Corona Virus Infectious Diseases (abbreviated as COVID-19) and based on the worldometers.com data, more than two hundred countries and territories are affected[1]. In this situation, every country tried with their best effort to combat this pandemic.

[1] https://www.worldometers.info/coronavirus/countries-where-coronavirus-has-spread/. Currently (around May 2021) the second wave has attacked countries like. As of now (dated 9th May, 2021), the number of deaths due to COVID-19 are 3304113, and the number of confirmed cases are 158863033.

© Springer Nature Switzerland AG 2021
O. Gervasi et al. (Eds.): ICCSA 2021, LNCS 12951, pp. 141–156, 2021.
https://doi.org/10.1007/978-3-030-86970-0_11

'Isolation', 'Quarantine', 'wearing of face mask' and 'Social Distancing' are the four keywords recommended by the *World Health Organization* (henceforth mentioned as WHO) to follow for combating this spreading disease. As per the topic of our study, here we only explain the concept of 'Social Distancing'. As per the WHO guidelines, the Corona virus spreads with the droplets that come out during the sneezing of an already infected person [27]. Hence, it is advisable to always maintain a distance of 1 meter (approximately 3 feet) from every other person.

Most of the countries (including India) were under temporary lock down and all the essential facilities including shopping malls, different modes of public transport, etc. were completely closed. However, for common people to survive some essential commodities (e.g., groceries, medicines, green vegetables, milk, etc.) are essential. It is better if the state administration can facilitate these essential services among the countrymen. However, as per the WHO guidelines, it is always advisable to maintain a distance of 1 meter from others whenever a person is out of his residence. Hence, it is an important question where to locate these facilities such that the demands of the customers can be served, and at the same time, social distancing among the people is maximized. Essentially, this is a facility location problem along with an additional measure 'social distancing'.

The study of the *Facility Location* problem goes back to the 17-th century and at that time this problem was referred to as the *Fermat-Weber Problem*[2]. This problem deals with selecting the location for placement of the facilities to best meet the demanded constraints of the customers [23]. Since its inception due to practical applications, this problem has been well studied from different points of view including algorithmic [9], computational geometry [6], the uncertainty of user demands [17], etc. However, to the best of our knowledge, the concept of social distancing is yet to be formalized mathematically, and also the facility location problem has not been studied yet with social distancing as a measure. It has been predicted that the third wave of COVID-19 is most likely to hit India sometime at the end of November or early December this year[3]. Hence, it is important to study the facility location problem considering social distancing as an important criterion.

In this paper, we study the problem of facility location under the social distancing criteria. The input to the problem is a road network (also known as graph) of a city where the nodes are small residential zone, edges are the streets connecting the zones. Additionally, we have how many different kinds of facilities are there and in which number they will be opened, population and demand of different kinds of facilities of every zone. The goal of this problem is locate the facilities in such a way that the basic amenities can be distributed to the people and the social distancing (mathematical abstraction is deferred until Sect. 3) is maximized. In particular we make the contributions in this paper.

[2] https://en.wikipedia.org/wiki/Geometric_median.

[3] https://www.indiatoday.in/coronavirus-outbreak/story/india-third-wave-of-covid-19-vaccine-prevention-1799504-2021-05-06.

- We mathematically formalize the notion of 'social distancing' and integrate it into the facility location problem. We formally call our problem as the Social Distancing-Based Facility Location Problem.
- We propose an optimization model for this problem which is a quadratic binary integer program.
- We propose two different approaches for solving the optimization problem. The first one is a simulation-based approach, and the other one is a heuristic solution.
- The proposed solution methodologies have been analyzed for time and space requirements.
- Finally, the proposed methodologies have been implemented with synthetic datasets and perform a comparative study between the proposed methodologies.

The remaining part of this paper is organized as follows: Sect. 2 presents some relevant studies from the literature. Section 3 describes our system's model, quantification of the 'social distancing', and introduce the Social Distancing-Based Facility Location Problem. Section 4 describes the mathematical model of our problem. In Sect. 5, we propose one simulation-based approach and a heuristic solution approach for solving this problem. Section 6 contains the experimental evaluation, and finally, Sect. 7 concludes our study.

## 2 Related Work

The problem of facility location was initially originated from the operations research community [10] and has been investigated by researchers from other communities as well including computational geometry [6], graph theory [28], Management Sciences [25], Algorithm Design [9], Geographical Information Systems [18] and many more. Also, this problem has been studied in different settings such as Uncapasitated [14], Capasited [29], Multistage [7], Hierarchical [12], Stochastic Demand [17], Unreliable Links [21], Under Disruptions [1] and their different combinations [22]. The problem has got applications in different domains of society including health care [2], defense [15], etc. However, to the best of our knowledge, there does not exist any literature that studies the facility location problem considering the social distancing as a measure.

The variant of facility location introduced and studied in this paper comes under the facility location under disruption. There exists an extensive set of literature in this context. Recently, Afify et al. [1] studied the reliable facility location problem under disruption, which is the improvisation of the well-studied Reliable $p$-Median Problem [20] and Reliable Uncapacitated Facility Location Problem. They developed an evolutionary learning technique to near-optimally solve these problems. Yahyaei and Bozorgi-Amiri [32] studied the problem of relief logistics network design under interval uncertainty and the risk of facility disruption. They developed a mixed-integer linear programming model for this problem and a robust optimization methodology to solve this. Akbari et al. [3] presented a new tri-level facility location (also known as $r$-interdiction median

model) model and proposed four hybrid meta-heuristics solution methodologies for solving the model. Rohaninejad et al. [24] studied a multi-echelon reliable capacitated facility location problem and proposed an accelerated *Benders Decomposition* Algorithm for solving this problem. Li and Zhang [16] studied the supply chain network design problem under facility disruption. To solve this problem they proposed a sample average approximation algorithm. Azizi [5] studied the problem of managing facility disruption in a hub-and-spoke network. They formulated mixed-integer quadratic program for this problem which can be linearized without significantly increasing the number of variables. For larger problem instances, they developed three efficient particle swarm optimization-based metaheuristics which incorporate efficient solution representation, short-term memory, and special crossover operator. Recently, there are several other studies in this direction [8,11]. As mentioned previously, there is an extensive set of literature on facility location problems on different disruption, and hence, it is not possible to review all of them. Look into [13,26] for a recent survey.

To the best of our knowledge, the problem of facility location has not been studied under the theme of social distancing, though this is the need of the hour. Hence, in this paper, we study the facility location problem with the social distancing criteria.

## 3 Preliminaries and the System's Model

In this section, we present the background and formalize the concept of social distancing. The input to our problem is the road network of a city represented by a simple, finite, and undirected graph $G(V, E, W)$. Here, the vertex set $V(G) = \{v_1, v_2, \ldots, v_n\}$ represents the small residential zones, the edge set $E(G) = \{e_1, e_2, \ldots, e_m\}$ denotes the links among the residential zones. $W$ denotes the edge weight function that maps the distance of the corresponding street, i.e., $W : E(G) \longrightarrow \mathbb{R}^+$. We use standard graph-theoretic notations and terminologies from [31]. For any zone $v_j \in V(G)$, $N_{<d}(v_j)$ denotes the set of neighbors of $v_j$ within the distance $d$. We denote the number of nodes and edges of $G$ by $n$ and $m$, respectively. Additionally, we have the following information. For each $v_j \in V(G)$, $P(v_j)$ denotes the people residing at zone $v_j$. In the rest of the paper, we use the term 'zone' and 'node', interchangeably. Let, $\mathcal{H} = \{h_1, h_2, \ldots, h_k\}$ denotes the $k$ different kinds of facilities (say, $h_1 \equiv$ 'groceries' , $h_2 \equiv$ 'medicine' and so on) need to be opened. For every zone $v_j \in V(G)$ and every kind of facility $h_i \in \mathcal{H}$, we know the corresponding demand $\mathcal{D}_{v_j}(h_i)$. Hence, the total demand for the facility $h_i$ is $\mathcal{D}(h_i) = \sum_{v_j \in V(G)} \mathcal{D}_{v_j}(h_i)$.

***Social Distancing.*** Now, we formally explain the meaning of the term 'social distancing' and express it as a mathematical expression. Suppose, a facility $h_i \in \mathcal{H}$ is opened for a particular duration $[t_1, t_2]$ in a day at the zone $v_j \in V(G)$. Let, $k^t_{i,j}$ denotes the number of people in the queue before this facility at time $t \in [t_1, t_2]$. From the geographical location, it is possible to determine the maximum number of people that can be present in the queue by maintaining WHO

recommended distance to each other. For the zone $v_j \in V(G)$ and facility $h_i \in \mathcal{H}$, this number is denoted as $\gamma_{i,j}$. If the number of people in the queue is within the threshold then the social distancing should be maximum and when it crosses the threshold the social distancing factor decreases linearly or exponentially depending upon the number of people. We denote the social distancing function by $S()$. The following equation mentions this.

$$S(k^t_{i,j}) = \begin{cases} \text{positive constant,} & \text{if } k^t_{i,j} < \gamma_{i,j} \\ \text{decreases linearly or exponentially,} & \gamma_{i,j} \leq k^t_{i,j} < 2 \cdot \gamma_{i,j} \\ \text{decreases linearly or exponentially,} & k^t_{i,j} \geq 2 \cdot \gamma_{i,j} \end{cases}$$

Once $k^t_{ij} > \gamma_{i,j}$, the value of social distancing function decreases in two ways: linearly and exponentially, which is described below.

1. **Linear Social Distance Factor**: Here, $S(.)$ decreases linearly with the increasing value current number of people in the queue at a particular facility of certain zone. We model this using Eq. 1.

$$S(k^t_{i,j}) = A - b * max\{k^t_{i,j} - \gamma_{i,j}, 0\}, \quad \text{where } A \text{ and } b \text{ are positive constant.} \tag{1}$$

2. **Exponential Social Distance Factor**: Here, the $S(.)$ decreases exponentially with the linear increment of the current queue strength at a particular facility of certain zone. We model this using Eq. 2.

$$S(k^t_{i,j}) = A - b \cdot [\gamma_{i,j} + exp(k^t_{i,j} - 2 \cdot \gamma_{i,j})^{\frac{1}{4}}], \quad \text{where } A > 1 \tag{2}$$

The goal of the social distancing-based facility location problem is to assign facilities in the zones such that the total value of the social distancing function is maximized. Formally, we present this problem as follows.

**Social Distancing-Based Facility Location Problem**

**Input:** The road network $G(V, E, W)$, Population at each zone $v_j \in V(G)$; i.e.; $P(v_i)$, Types of facilities $\mathcal{H} = \{h_1, h_2, \ldots, h_k\}$, Number of facilities of each type to be open; i.e.; $\{\ell_1, \ell_2, \ldots, \ell_k\}$, and their corresponding demand in each zone $v_j$; i.e.; $\{\mathcal{D}_{v_j}(h_1), \mathcal{D}_{v_j}(h_2), \ldots, \mathcal{D}_{v_j}(h_k)\}$ for all $v_j \in V(G)$.

**Problem:** Place maximum number of facilities to be opened of each type such that total social distancing is maximized.

Here, we note that each $\ell_i << n$. Symbols and notations used in this study are given in Table 1.

**Table 1.** Symbols with their interpretation

| Symbol | Interpretation |
|---|---|
| $G(V, E, W)$ | The road network under consideration |
| $V(G), E(G)$ | Residential zones and road segments joining them |
| $W$ | Edge weight function |
| $N_{<d}(v_j)$ | Neighbors of the zone $v_j$ within distance $d$ |
| $d_{max}$ | Maximum degree of a node in $G$ |
| $P(v_j)$ | Population at the zone $v_j$ |
| $P_{max}$ | Maximum population of a zone in $V(G)$ |
| $P$ | Total population of all zones of $V(G)$ |
| $\mathcal{H}$ | Set of different kinds of facilities |
| $k$ | Number of different kinds of facilities; i.e.; $\|\mathcal{H}\| = k$ |
| $\ell_i$ | Number of facilities opened of type $h_i$ |
| $\ell_{max}$ | Maximum number of facilities among all types |
| $[t_1, t_2]$ | Operation time window of the facilities |
| $\Delta$ | The difference between $t_2$ and $t_1$ |
| $k^t_{i,j}$ | Number of people in the queue at $v_j \in V(G)$, $h_i \in \mathcal{H}$, and $t \in [t_1, t_2]$ |
| $\gamma_{i,j}$ | Queue strength of the facility $h_i \in \mathcal{H}$ at $v_j \in V(G)$ |
| $S(.)$ | The social distancing function |
| $\mathcal{D}_{v_j}(h_i)$ | Demand of facility type $h_i$ at zone $v_j$ |

## 4 Mathematical Model Formulation

Now, we formulate a mathematical model of our problem. First, we define the decision variables involved in our formulation.

$$x_{i,j} = \begin{cases} 1, & \text{if the facility of type } h_i \text{ is opened at zone } v_j \\ 0, & \text{otherwise} \end{cases}$$

$$y^t_{i,j,p} = \begin{cases} 1, & \text{if the person } p \text{ visits the facility of type } h_i \text{ at zone } v_j \text{ at time } t \\ 0, & \text{otherwise} \end{cases}$$

Given the decision variables, total social distancing can be given by the following equation.

$$F = \sum_{h_i \in \mathcal{H}} \sum_{v_j \in V(G)} x_{i,j} \Big( \sum_{t \in [t_1, t_2]} \sum_{p \in P(v_j)} S(k^t_{i,j})\ y^t_{i,j,p} \Big) \tag{3}$$

In Eq. 3, four summations are involved. The first one is to sum up for all the facilities. Similarly, the second one is to sum up for all the nodes. Next, one is

used to sum up for all distinct time stamps within the operated time interval. Finally, the last one is for all the people of a node. The goal is to maximize the function mentioned in Eq. 3 subject to certain constraints. Now, we present the mathematical model, and next, we describe the meaning of each constraint.

**Maximizing Social Distancing**

$$max \sum_{h_i \in \mathcal{H}} \sum_{v_j \in V(G)} x_{i,j} \Big( \sum_{t \in [t_1,t_2]} \sum_{p \in P(v_j)} S(k^t_{i,j})\ y^t_{i,j,p} \Big)$$

$$\text{subject to,} \quad \sum_{v_j \in V(G)} x_{i,j} \leq \ell_i, \quad \forall h_i \in \mathcal{H} \tag{4}$$

$$\sum {\boldsymbol{y}^i_{i,j,p}}^T\ \boldsymbol{Y}^i_{i,:,p} = 1, \ \ \forall p \in \bigcup_{v_k \in V(G)} P(v_k), \ \ \forall v_j \in V(G), \ \ \forall h_i \in \mathcal{H} \tag{5}$$

$$\sum_{t \in [t_1,t_2]} \sum_{v_k \in \{v_j\} \cup N_{<d}(v_j)} y^t_{i,k,p} \geq 1, \forall h_i \in \mathcal{H}, \forall v_j \in V(G), \forall p \in P(v_j) \tag{6}$$

$$y^t_{i,j,p} \Big( \sum_{t' \in [t_1,t_2]} y^{t'}_{i,k,p} \Big) \{ t - \sum_{t' \in [t_1,t_2]} y^{t'}_{i,k,p} t' \} \leq 0, \ \forall h_i \in \mathcal{H}, \ \forall (j,k) \in E(G), \ \forall p \in P(v_j), \ \forall t \in [t_1,t_2] \tag{7}$$

$$x_{i,j} \in \{0,1\}, \ \ \forall h_i \in \mathcal{H}, \ \ \forall v_j \in V(G) \tag{8}$$

$$y^t_{i,j,p} \in \{0,1\}, \ \forall h_i \in \mathcal{H}, \ \forall v_j \in V(G), \ \forall t \in [t_1,t_2], \ \forall p \in \bigcup_{v_k \in V(G)} P(v_k) \tag{9}$$

The first constraint in Inequation 4 enforces that the total number of open facilities of a particular type $h_i$ should not be more than its allowance (which is $\ell_i$). All the remaining constraints are on $y^t_{i,j,p}$. The assumption is a person can visit a facility location at once. Then, he or she can wait there for a certain period for being served or visit the nearest neighboring facility location to get the service. Now, the quantification of the continuous time domain is done in such a way that $y^t_{i,j,p}$ can be one, only once for the entire duration of $[t_1, t_2]$. Another physical condition implies that one person can not be in two different facility locations at the same time. Both the scenarios are capture in the second constraint described in Inequation 5. Here, $\boldsymbol{Y}^i_{i,:,p}$ is the matrix of shape total duration by the number of facility locations denoting the presence of a person $p$ at any location in the entire time duration to avail $h_i$-type facility, and $\boldsymbol{y}^i_{i,j,p}$ is the one hot vector of the matrix $\boldsymbol{Y}^i_{i,:,p}$, which describes the presence at the particular location $v_j$. Now, the next constraint in Inequation 6 implies that the people should be served either at the location where they stay (e.g. $v_j$), or any nearby locations (e.g. $N_{<d}(v_j)$) within certain vicinity. The next constraint in Eq. 7 tells that if an arbitrary person $p$ visits the locations at $v_j$ and $v_k \in N_{<d}(v_j)$

to avail the facility type $h_i$ then he first visits the place where he stays, then the neighboring ones. Finally, Inequation 8 and 9 tells that the decision variables $x_{ij}$ and $y^t_{ijp}$, for all $h_i \in \mathcal{H}$, $v_j \in V(G)$, $t \in [t_1, t_2]$ take binary values.

**Note**: Observe that in this study we do not consider the cost of the facilities which is quite natural in case of traditional facility location problems [23]. This is due to the following reason. During the lock down period, a significant population are jobless, and hence government is providing the basic amenities free of cost. This is our assumption as mentioned in Sect. 1. We also assume that there is no dearth of supply.

## 5 Proposed Solution Approaches

As the facility location problem is NP-Hard, finding the optimal solution for our problem is also computationally intractable. Hence, we propose two different approaches to solve our problem. The first one is the simulation-based approach and the second one is a heuristic algorithm. It is important to note that none of these methods produces an optimal solution.

### 5.1 Simulation-Based Approach

A simulation-based approach has been adopted to obtain a solution for the developed model. Demand is generated for the people from all the nodes. As mentioned previously, a facility of type $h_i$ can have at most $\ell_i$ many ($\ell_i << n$). People of a node get the service of a facility from the same node if it is available there, otherwise move to the neighbor node where a service facility is opened. People's arrival to the service facility follows *Poisson Distribution* and service time follows *Exponential Distribution* [4]. People get service based on the first come first serve basis. We have simulated these features of how people are arriving and getting the service and during this process how social distance is being maintained based on the number of people in the queue at a given time. People from several nodes may come and form a queue based on the arrival time and get the service based on first come first serve (FCFS) basis. Due to the space limitations we do not discuss anything related to the queuing theory and can be found at [4].

Now, the arrival time of people of all the nodes is generated based on the predefined arrival rate. It is implemented based on the inter-arrival time using poisson distribution which is reciprocal of arrival rate. Next, for each node, where the facility is not opened, the nearby open facility is found based on the least distance. Arrival time for all the people in a given and its nearby nodes are sorted. Service times are generated for the nodes where a given type of service facility is opened using the exponential distribution. Social distance is calculated at each open node using Eq. 3. This process is repeated for a defined number of simulation runs. Finally, the locations corresponding to the maximum value of the social distancing function are returned as a solution. This process is shown in the form of pseudo code in Algorithm 1.

**Algorithm 1:** Pseudo Code for the Simulation-Based Approach

**Data:** Underlying Road Network $G(V, E, W)$, Population and Respective Demand of Every Zone, Number of Different Types of Facilities, Number of facilities of each type

**Result:** Location of Different Facilities

1 $Simulation_Run \longleftarrow 0$;
2 **while** $Simulation_Run$ *not reached* **do**
3 **Step 1**: For every type of facility (say $h_i$) the allowed number of them ( say $\ell_i$)are opened randomly;
4 **Step 2**: Arrival time of peoples of all the nodes are generated based on the predefined arrival rate.;
5 **Step 3**: For each node, where facility is not opened, the nearest open facility is found based on the least distance.;
6 **Step 4**: For a node $v_j$, the people of of this zone and those who are assigned to this facility, their arrival times are sorted and queue is formed.
7 **Step 5**: Service times are generated for the nodes where a given type of service facility is opened using the exponential distribution of service time.;
8 **Step 6**: Based on the arrival time and service time, number of people in queue is derived. ;
9 **Step 7**: Social distance is calculated at each queue and summed up.;
10 **Step 8**: $Simulation_Run$ is incremented by 1; i.e.; $Simulation_Run = Simulation_Run + 1$ ;
11 **end**
12 **Step 9**: Among all the simulation runs corresponding locations of the maximum is returned for locating the facilities.

Now, we analyze Algorithm 1 to understand its time and space requirement. Let, $\mathcal{R}$ denotes the number of simulation runs. Let, the number of different facilities are $k$ and $\ell_{max}$ denotes the maximum number of facilities among all types. Now, it is easy to observe that the running time of Step 1 is of $\mathcal{O}(k \cdot \ell_{max})$. Let, $P(v_i)$ denotes the number of people of at node $v_i$ and $P$ denotes the number of people in all the nodes in $G$; i.e.; $P = \sum_{v_i \in V(G)} P(v_i)$. As, for a people generating the arrival time requires $\mathcal{O}(1)$ time, hence, time requirement of Step 2 requires $\mathcal{O}(P)$ time. The number of nodes where the facility is not opened will be greater than equals to $(n - \ell_{max})$ and in the worst case it will be of $\mathcal{O}(n)$. Let $d_{max}$ denotes the maximum possible degree of a node in $G$; $d_{max} = \underset{v \in V(G)}{argmax}\ deg(v)$. For each node without facility to choose the nearest facility requires $\mathcal{O}(d_{max})$. So, for all the nodes and all types of facilities required computational time is of $\mathcal{O}(n \cdot k \cdot d_{max})$. This implies that the execution of Step 3 requires $\mathcal{O}(n \cdot k \cdot d_{max})$ time. Let, $P_{max}$ denotes the maximum number of people residing at any zone, hence, $P_{max} = \underset{v \in V(G)}{argmax}\ P(v)$. As, in any facility the people of its neighbors may come, and hence, the total number of people in any facility is of $\mathcal{O}((d_{max}+1)P_{max})$. Sorting this people based on the arrival time will require

$\mathcal{O}((d_{max} \cdot P_{max} + P_{max}) \log(d_{max} \cdot P_{max} + P_{max}))$ time. This implies that the execution of Step 4 requires $\mathcal{O}((d_{max} \cdot P_{max} + P_{max}) \log(d_{max} \cdot P_{max} + P_{max}))$ time. Assuming that the generating service time requires $\mathcal{O}(1)$, executing Step 5 requires $\mathcal{O}(k \cdot \ell_{max})$ time. For a single queue, calculating the number of people in it requires $\mathcal{O}(1)$ time. Hence, executing Step 6 requires $\mathcal{O}(k \cdot (d_{max} + 1) \cdot P_{max})$.

Assume that $t_2 - t_1 = \Delta$. Hence, from the objective function it is easy to follow that computing the social distancing requires $\mathcal{O}(n \cdot k \cdot \Delta \cdot P_{max})$ time. If in each iteration of the `while` loop, we update the maximum value of the social distancing and the corresponding location of the facilities of different types then after the last iteration we obtain the solution of the proposed simulation approach. This requires $\mathcal{O}(k \cdot \ell_{max})$ time. So, the total time required by the simulation procedure is of $\mathcal{O}(\mathcal{R}(k \cdot \ell_{max} + P + n \cdot k \cdot d_{max} + (d_{max} \cdot P_{max} + P_{max}) \log(d_{max} \cdot P_{max} + P_{max}) + k \cdot \ell_{max} + k \cdot (d_{max} + 1) \cdot P_{max} + n \cdot k \cdot \Delta \cdot P_{max} + k \cdot \ell_{max}))$. This reduces to $\mathcal{O}(\mathcal{R}(k \cdot \ell_{max} + P + n \cdot k \cdot d_{max} + (d_{max} \cdot P_{max} + P_{max}) \log(d_{max} \cdot P_{max} + P_{max}) + k \cdot d_{max} \cdot P_{max} + n \cdot k \cdot \Delta \cdot P_{max}))$. Additional space required by the proposed simulation approach is to store the arrival time of the people which is of $\mathcal{O}(n \cdot P_{max})$, service time of the queues which is of $\mathcal{O}(k \cdot \ell_{max})$, location the facilities which is of $\mathcal{O}(k \cdot \ell_{max})$, social distancing function value at each queue which is of $\mathcal{O}(k \cdot \ell_{max})$. So, the total space requirement is of $\mathcal{O}(n \cdot P_{max} + k \cdot \ell_{max})$. Hence, Theorem 1 holds.

**Theorem 1.** *The running time and space requirement of the proposed simulation-based approach is of* $\mathcal{O}(\mathcal{R}(k \cdot \ell_{max} + P + n \cdot k \cdot d_{max} + (d_{max} \cdot P_{max} + P_{max}) \log(d_{max} \cdot P_{max} + P_{max}) + k \cdot d_{max} \cdot P_{max} + n \cdot k \cdot \Delta \cdot P_{max}))$ *and* $\mathcal{O}(n \cdot P_{max} + k \cdot \ell_{max})$*, respectively.*

### 5.2 Heuristic Solution

In this section, we propose a heuristic solution to solve our problem. First, we describe the working principle, and subsequently, we present it as a step-by-step procedure. As a first step, we sort all the nodes based on demand in descending order. Based on this heuristic for any type of facility $h_i \in \mathcal{H}$, $\ell_i$ many facilities will be placed at the top demand zones. Now, it allocates the people to the facility in the following way: For the high demand locations the people of that zones are allocated to the same zone, and subsequently, we allocate the people of other zone in the reverse order such that the entire population of $V(G)$ is distributed as uniformly as possible. Without loss of generality, assume that after sorting the nodes based on the demand, the order of the nodes are $\rho =< v_1, v_2, \ldots, v_n >$. For any zone $v_j \in \{v_1, v_2, \ldots, v_i\}$, the people of zone $v_i$ are allocated to the facility at the same zone. Now, $v_{i+1}$ is allocated to $v_i$, $v_{i+2}$ is allocated to $v_{i-1}$, and $v_{2i}$ to $v_1$. Similarly, $v_{2i+1}$ is allocated to $v_2$, $v_{2i+2}$ is allocated to $v_3$ and so on. Next from Step 2 to Step 7 of Algorithm 1 is executed. Pseudo code is given in Algorithm 2.

Now, we analyze Algorithm 2 to understand its time and space requirement. Sorting of the nodes based on the demand requires $\mathcal{O}(n \log n)$ time. It is easy to observe that the allocation of the facilities requires $\mathcal{O}(n)$ time. As described in

**Algorithm 2:** Pseudo code for the Heuristic Algorithm

**Data:** Underlying Road Network $\mathcal{G}(\mathcal{V}, \mathcal{E})$, Population and Respective Demand of Every Zone, Number of Different Types of Facilities, Number of facilities of each type

**Result:** Location of Different Facilities

1 **Step 1**: All the nodes of the network are sorted based on the demand of the nodes;

2 **Step 2**: People of nodes are allocated in the facilities as described.;

3 **Step 3**: Execute Step 2 to Step 7 of Algorithm 1.;

Sect. 5.1, the total execution time from Steps 2 to 7 is of $\mathcal{O}(P+n\cdot k\cdot d_{max}+(d_{max}\cdot P_{max}+P_{max})\log(d_{max}\cdot P_{max}+P_{max})+k\cdot\ell_{max}+k\cdot d_{max}\cdot P_{max}+n\cdot k\cdot\Delta\cdot P_{max})$ time. Hence, total time requirement of Algorithm 2 is of $\mathcal{O}(n\log n+P+n\cdot k\cdot d_{max}+(d_{max}\cdot P_{max}+P_{max})\log(d_{max}\cdot P_{max}+P_{max})+k\cdot\ell_{max}+k\cdot d_{max}\cdot P_{max}+n\cdot k\cdot\Delta\cdot P_{max})$. It is easy to observe that the space requirement of Algorithm 2 is same as Algorithm 1 which is of $\mathcal{O}(n\cdot P_{max}+k\cdot\ell_{max})$. Hence, Theorem 2 holds.

**Theorem 2.** *The running time and space requirement of the proposed heuristic approach is of* $\mathcal{O}(n\log n+P+n\cdot k\cdot d_{max}+(d_{max}\cdot P_{max}+P_{max})\log(d_{max}\cdot P_{max}+P_{max})+k\cdot\ell_{max}+k\cdot d_{max}\cdot P_{max}+n\cdot k\cdot\Delta\cdot P_{max})$ *and* $\mathcal{O}(n\cdot P_{max}+k\cdot\ell_{max})$, *respectively.*

## 6 Experimental Evaluation

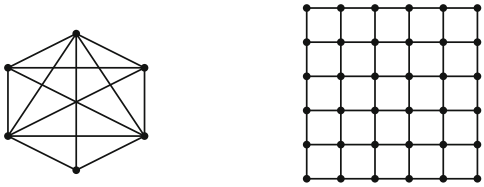
In this section, we describe the experimental evaluation of the proposed methodology. Initially, we start by describing the datasets.

### 6.1 Dataset Description

As it is difficult to find datasets that contains all the required information in this study, hence we create some synthetic datasets. We perform our experiments with two different kinds of networks topology:

- **Complete Network**: In this kind of network, every vertex is connected to every other vertex of the network. It is easy to observe that an undirected complete network with $n$ vertices will have $\frac{n(n-1)}{2}$ edges.
- **Rectangular Grid Network**: A rectangular grid network $G(V, E)$ of size $n \times n$ is defined with the vertex set $V(G) = [n] \times [n] = \{1, 2, \ldots, n\} \times \{1, 2, \ldots, n\}$ and two vertices $(i, j)$ and $(i^{'}, j^{'})$ will be connected by an edge if both $|i - i^{'}| \leq 1$ and $|j - j^{'}| \leq 1$. Many real-world road networks are of grid structure [19,30].

For the ease of understanding, we show a demo diagram of complete and grid network in Fig. 1.



Copmplete Network $(n = 6)$ Grid Network $(n = 6 \times 6)$

**Fig. 1.** Figure of a complete and grid network of size $n = 6$ and $n = 6 \times 6$, respectively.

### 6.2 Experimental Setup

In this section, we describe the experimental setup. Several parameters are there, whose value needs to be set. Here, we describe them one by one.

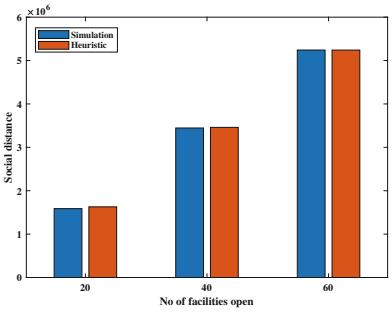
- **Network Size**: As mentioned, we use networks of two different kinds, namely complete and grid. For complete network, we use $n = 60$ and $n = 100$. On the other hand for the grid network we consider $n = 60 \times 60$ and $n = 100 \times 100$.
- **Population at Each Zone**: This has been chosen uniformly at random from the interval $[1000, 2000]$ for each zone.
- **Demand**: For simplicity, we consider the uniform and demand for all the people.
- **Parameters Related to Queuing Theory**: We consider the inter arrival time and service time are 1 min and 0.7 min, respectively.
- **Parameters in the Social Distancing Function**: As mentioned previously, there are three parameters in the social distancing function. They are $\gamma_{i,j}, A,$ and $b$. In this study, we consider $\gamma_{i,j} = 4$ for all $h_i \in \mathcal{H}$ and $v_j \in V(G)$, $A = 10$, and $b = 0.5$.

We implement our both solution approaches in MATLAB (Version MATLAB 9.0 R2016a) and all the simulation codes and the synthetic datasets can be found at https://github.com/BITHIKA1992/Facility_location_Covid19.

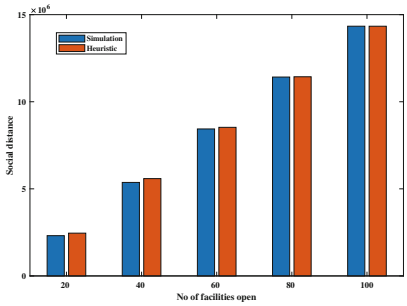
### 6.3 Experimental Results

Now, we discuss the experimental results. Our goal here is to study the impact of number of facilities on social distancing and average queue length.

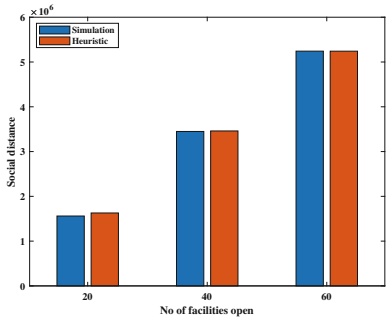
***Impact on the Social Distancing.*** Here, we discuss a comparative study of two different solution methodologies on social distancing. Figure 2 shows the plots for the number of opened facilities vs. social distancing for both kinds of network topologies with two different network sizes. From the figure, it has been observed that for a fixed network size when the number of opened facilities increases, social distancing also increases. As an example For a complete network with 100 nodes, when the number of opened facilities is 20, the value of social distancing due to the allocation of facilities by the heuristic method is $2.4 \times 10^6$.



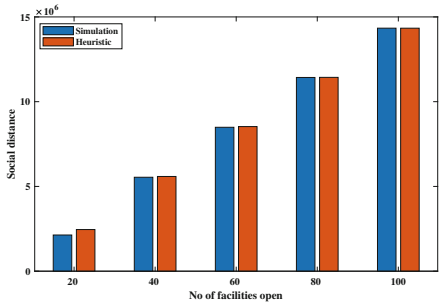

(a) Complete Network ($n = 60$ )




(b) Complete Network ($n = 100$)




(a) Grid Network ($n = 60 \times 60$)




(b) Grid Network ($n = 100 \times 100$)

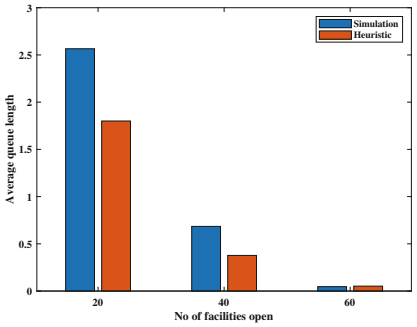
**Fig. 2.** Impact of the number of opened facilities on Social Distancing for Complete and Grid Network for two different sizes.

When the number of opened facilities are increased to 60, the value of social distancing also increased to $7.9 \times 10^6$.

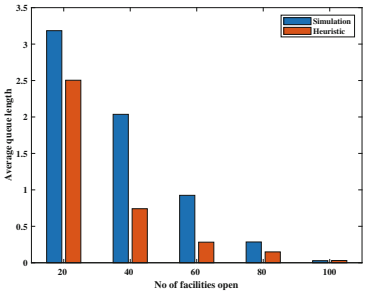
The impact of algorithms on social distancing can also be observed from Fig. 2. We can conclude that when the number of opened facilities is more, it does not matter which algorithm is used to locate facilities the value of social distancing will not change much. However, if the number of opened facilities is much less, the location suggested by the heuristic approach leads to more value of the social distancing. As an example, in case of the $100 \times 100$ rectangular grid network when the number of opened facilities is 20, the value of social distancing function due to facility placement by simulation-based approach and heuristic approach is $2.08 \times 10^6$ and $2.29 \times 10^6$, respectively. Hence, the difference is approximately 210000.

***Impact on Average Queue Length.*** Now, we discuss the impact of the number of opened facilities on the average queue length. Figure 3 shows the number of opened facilities vs. average queue length plots for both kinds of networks for two different sizes. From the figure, it has been observed that as the number of opened facilities increases naturally, the average queue length decreases. As an example, for a complete network with 60 nodes when the number of opened

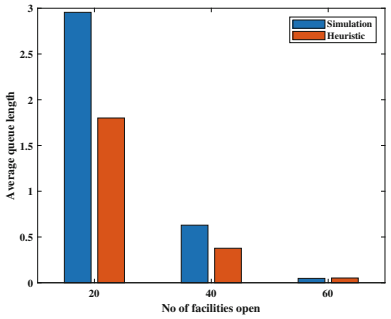
facilities is 20, the average queue length is 2.58 by the simulation-based approach. However, when it is increased to 40 the average queue length drops down to 0.68. It is important to observe that between the two proposed methodologies, the locations selected by the heuristic approach for placing the facilities lead to less average queue length. As an example, for a grid network of size $60 \times 60$ when the number of opened facilities is 20, the average queue length due to simulation-based and heuristic approach is 2.98 and 1.56, respectively.



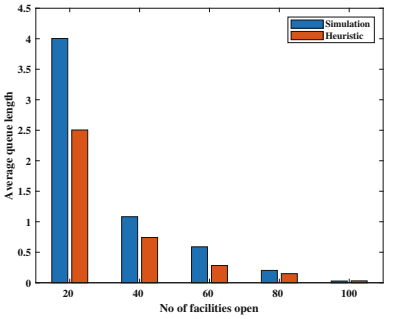

(a) Complete Network ($n = 60$ )




(b) Complete Network ($n = 100$)




(a) Grid Network ($n = 60 \times 60$)




(b) Grid Network ($n = 100$)

**Fig. 3.** Impact of the number of opened facilities on average queue length for complete and grid network for two different sizes.

From the experiments, we can observe that the allocation of facilities by the heuristic approach leads to lesser queue length and consequently more value of social distancing. This is due to the fact that the heuristic approach tries to assign the people to the facilities in such a way that their loads are balanced. Due to this reason, the social distancing achieved by the heuristic solution is more.

## 7 Concluding Remarks

In this paper, we have studied the problem of the Social Distancing-Based Facility Location Problem and proposed two solution methodologies. The first one is a simulation-based approach and the second one is a heuristic solution. From the experiments with synthetic datasets, we observe that the heuristic solution leads to the allocation of facilities causing more social distancing. Our study can be extended in several directions. One can consider more realistic scenarios like unknown demand, supply constraint, and many more. We can also extend our study for competitive situation where there will be more than one agency for providing the facilities of each type and their goal is to maximize their profit at the same time the social distancing needs to be maintained.

## References

1. Afify, B., Ray, S., Soeanu, A., Awasthi, A., Debbabi, M., Allouche, M.: Evolutionary learning algorithm for reliable facility location under disruption. Expert Syst. Appl. **115**, 223–244 (2019)
2. Ahmadi-Javid, A., Seyedi, P., Syam, S.S.: A survey of healthcare facility location. Comput. Oper. Res. **79**, 223–263 (2017)
3. Akbari-Jafarabadi, M., Tavakkoli-Moghaddam, R., Mahmoodjanloo, M., Rahimi, Y.: A tri-level r-interdiction median model for a facility location problem under imminent attack. Comput. Ind. Eng. **114**, 151–165 (2017)
4. Asmussen, S.: Applied Probability and Queues, vol. 51. Springer, New York (2008)
5. Azizi, N.: Managing facility disruption in hub-and-spoke networks: formulations and efficient solution methods. Ann. Oper. Res. **272**(1–2), 159–185 (2019)
6. Bespamyatnikh, S., Kedem, K., Segal, M., Tamir, A.: Optimal facility location under various distance functions. Int. J. Comput. Geom. Appl. **10**(05), 523–534 (2000)
7. Biajoli, F.L., Chaves, A.A., Lorena, L.A.N.: A biased random-key genetic algorithm for the two-stage capacitated facility location problem. Expert Syst. Appl. **115**, 418–426 (2019)
8. Fotuhi, F., Huynh, N.: Reliable intermodal freight network expansion with demand uncertainties and network disruptions. Netw. Spatial Econ. **17**(2), 405–433 (2017)
9. Hassin, R., Ravi, R., Salman, F.S., Segev, D.: The approximability of multiple facility location on directed networks with random arc failures. Algorithmica **82**, 2474–2501 (2020)
10. Hassin, R., Tamir, A.: Improved complexity bounds for location problems on the real line. Oper. Res. Lett. **10**(7), 395–402 (1991)
11. Hatefi, S.M., Jolai, F.: Robust and reliable forward-reverse logistics network design under demand uncertainty and facility disruptions. Appl. Math. Model. **38**(9–10), 2630–2647 (2014)
12. Helber, S., Böhme, D., Oucherif, F., Lagershausen, S., Kasper, S.: A hierarchical facility layout planning approach for large and complex hospitals. Flexible Serv. Manufact. J. **28**(1–2), 5–29 (2016)
13. Ivanov, D., Dolgui, A., Sokolov, B., Ivanova, M.: Literature review on disruption recovery in the supply chain. Int. J. Prod. Res. **55**(20), 6158–6174 (2017)
14. Krishnaswamy, R., Sviridenko, M.: Inapproximability of the multilevel uncapacitated facility location problem. ACM Trans. Alg. (TALG) **13**(1), 1–25 (2016)

15. Lessin, A.M., Lunday, B.J., Hill, R.R.: A bilevel exposure-oriented sensor location problem for border security. Comput. Oper. Res. **98**, 56–68 (2018)
16. Li, X., Zhang, K.: A sample average approximation approach for supply chain network design with facility disruptions. Comput. Ind. Eng. **126**, 243–251 (2018)
17. Marinakis, Y., Marinaki, M., Migdalas, A.: A hybrid clonal selection algorithm for the location routing problem with stochastic demands. Ann. Math. Artif. Intell. **76**(1–2), 121–142 (2016)
18. Miller, H.J.: GIS and geometric representation in facility location problems. Int. J. Geograph. Inf. Syst. **10**(7), 791–816 (1996)
19. Miyagawa, M.: Optimal hierarchical system of a grid road network. Ann. Oper. Res. **172**(1), 349 (2009)
20. Mladenović, N., Brimberg, J., Hansen, P., Moreno-Pérez, J.A.: The p-median problem: a survey of metaheuristic approaches. Eur. J. Oper. Res. **179**(3), 927–939 (2007)
21. Narayanaswamy, N.S., Nasre, M., Vijayaragunathan, R.: Facility location on planar graphs with unreliable links. In: Fomin, F.V., Podolskii, V.V. (eds.) CSR 2018. LNCS, vol. 10846, pp. 269–281. Springer, Cham (2018). https://doi.org/10.1007/978-3-319-90530-3_23
22. Ortiz-Astorquiza, C., Contreras, I., Laporte, G.: Multi-level facility location problems. Eur. J. Oper. Res. **267**(3), 791–805 (2018)
23. Owen, S.H., Daskin, M.S.: Strategic facility location: a review. Eur. J. Oper. Res. **111**(3), 423–447 (1998)
24. Rohaninejad, M., Sahraeian, R., Tavakkoli-Moghaddam, R.: An accelerated benders decomposition algorithm for reliable facility location problems in multi-echelon networks. Comput. Indus. Eng. **124**, 523–534 (2018)
25. Ross, G.T., Soland, R.M.: Modeling facility location problems as generalized assignment problems. Manag. Sci. **24**(3), 345–357 (1977)
26. Snyder, L.V., Atan, Z., Peng, P., Rong, Y., Schmitt, A.J., Sinsoysal, B.: OR/MS models for supply chain disruptions: a review. IIE Trans. **48**(2), 89–109 (2016)
27. Tahir, M., Shah, S.I.A., Zaman, G., Khan, T.: Stability behaviour of mathematical model MERS corona virus spread in population. Filomat **33**(12), 3947–3960 (2019)
28. Tamir, A.: The K-centrum multi-facility location problem. Discrete Appl. Math. **109**(3), 293–307 (2001)
29. Tran, T.H., Scaparra, M.P., O'Hanley, J.R.: A hypergraph multi-exchange heuristic for the single-source capacitated facility location problem. Eur. J. Oper. Res. **263**(1), 173–187 (2017)
30. Watanabe, D.: A study on analyzing the grid road network patterns using relative neighborhood graph. In: The Ninth International Symposium on Operations Research and Its Applications, pp. 112–119. Citeseer (2010)
31. West, D.B., et al.: Introduction to Graph Theory, vol. 2. Prentice Hall, Upper Saddle River (2001)
32. Yahyaei, M., Bozorgi-Amiri, A.: Robust reliable humanitarian relief network design: an integration of shelter and supply facility location. Ann. Oper. Res. **283**(1), 897–916 (2019)