Transportation Research Part B 106 (2017) 130–152

Contents lists available at ScienceDirect

Transportation Research Part B

journal homepage: www.elsevier.com/locate/trb

Benders-and-Price approach for electric vehicle charging station location problem under probabilistic travel range

Chungmok Lee[a], Jinil Han[b,*]

[a] Department of Industrial and Management Engineering, Hankuk University of Foreign Studies, 81 Oedae-ro, Mohyeon-myeon, Cheoin-gu, Yoingin-si, Gyeonggi-do 17035, Republic of Korea

[b] Department of Industrial and Information Systems Engineering, Soongsil University, 369 Sangdo-ro, Dongjak-gu, Seoul 06978, Republic of Korea

A R T I C L E I N F O

Article history:
Received 11 June 2017
Revised 18 September 2017
Accepted 23 October 2017
Available online 28 October 2017

Keywords:
Electric vehicles
Charging station location problem
Flow refueling location model
Probabilistic travel range
Benders decomposition
Column generation

A B S T R A C T

This paper investigates the optimal locations for refueling stations for electric vehicles. Electric vehicles have been successfully introduced into the market. However, their use seems to be limited to urban transport since recharging facilities are readily available only near home and work. Planning recharging infrastructure for electric vehicles is highly relevant because this will enable longer trips, including inter-state travel which requires multiple battery charges. Among various models to determine optimal locations of recharging stations, a flow refueling location model (FRLM) is considered in this study. It determines locations for recharging stations to maximize the flow that can travel between origin and destination pairs by refueling at built facilities. FRLM is extended by introducing a probabilistic consideration of the travel range which might vary depending on various factors including road conditions. We develop a mixed integer nonlinear programming formulation and propose a Benders-and-Price algorithm by combining the Benders decomposition and column generation to solve the proposed formulation. The proposed algorithm is validated using extensive computational experiments on two transport networks, including a real-life Texas highway network.

1. Introduction

Due to environmental and economic concerns, electric vehicles (EVs) are now gaining popularity. In particular, serious concerns regarding climate change at global scale discourage the use of conventional internal combustion engine vehicles that use fossil fuels while encouraging their replacement with more environmentally friendly transport such as electric vehicles. A recent study has estimated that the sale of the electric vehicles will maintain an annual growth rate above 25% by 2025 (International Energy Agency, 2016). In Norway, which is an environmentally conscious country, the market share of electric vehicles reached 23% in 2016. For a new buyer, electric vehicles are economically tempting because they offer lower running costs with various incentives provided by the government. For example, many countries have already established pro-electric-vehicle policies, including incentives at purchase, free parking, and waivers on access restrictions (Nie et al., 2016).

* Corresponding author.
E-mail addresses: chungmok@hufs.ac.kr (C. Lee), jinil.han@ssu.ac.kr (J. Han).

https://doi.org/10.1016/j.trb.2017.10.011

Even though battery costs have decreased by a factor of four since 2008, battery is one of the most expensive and inefficient components of electric vehicles (International Energy Agency, 2016). In particular, most electric vehicles have a very limited travel ranges without recharging the battery compared to conventional fossil fuel-powered vehicles. With current technology, it takes too much time to recharge a battery completely. Therefore, alternative methods have been proposed, including battery-swapping (Mak et al., 2013; Adler and Mirchandani, 2014; Hof et al., 2017) and dynamic wireless recharging (Chen et al., 2016; Liu and Wang, 2017). Nonetheless, compared to well-established infrastructure for fossil fuels (e.g., gas stations), electric charging stations are not readily available everywhere. This situation thus motivates a design problem to determine optimal locations for recharging stations.

There are two dominant approaches in the literature to model optimal locations for recharging stations: spatial approach and flow-based approach. Spatial approach is an adaptation of well-known models such as *p*-median extensively used for facility location (Toregas et al., 1971; Campbell, 1996; Drezner and Hamacher, 1995; An et al., 2014). The idea is to minimize the distance between the location of a facility and those of possible customers to be served by the facility. In the context of an electric vehicle, the facility corresponds to a recharging station while a customer is an electric vehicle that is to be served. The *p*-median model assumes that drivers of electric vehicles will purposely visit recharging stations before the battery is exhausted. The goal is to cover a given area by deploying recharging stations so that the total number of stations can be minimized while maximizing coverage. This is why it is called spatial approach. Unfortunately, limited travel range of electric vehicles makes complete coverage across the nation virtually impossible. Therefore, a *p*-median approach primarily targets urban transport planning problems (Frade et al., 2011).

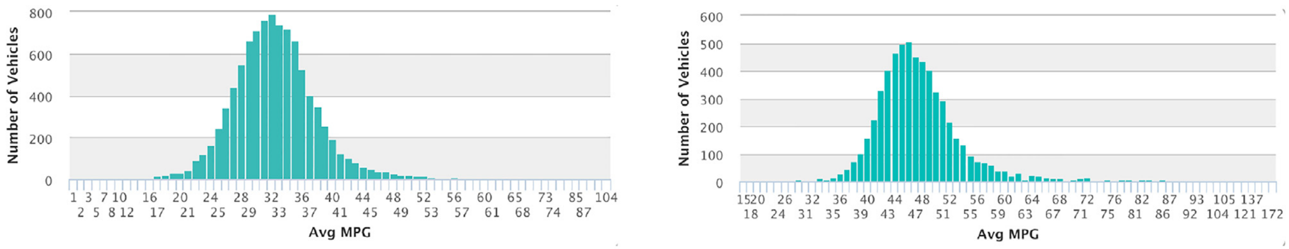
The flow-based approach *captures* electric vehicle traffic by inserting recharging stations on traffic routes. Each traffic flow is defined by an origin-destination (OD) pair. Electric vehicles are assumed to become a flow by traveling from their origins to their destinations. Kuby and Lim (2005) have proposed a flow refueling location model (FRLM). Its main idea is to divide a long route into several segments so that each of them can be traversed without recharging. It deploys recharging stations at each junction between two adjacent segments. Since there is no need to install recharging stations off the road, this approach tends to provide more efficient solutions than the spatial approach for a nation-wide highway road network (Upchurch et al., 2009). Many previous studies have utilized FRLM or a variant thereof to find optimal locations for recharging stations for a highway transport network (Kuby et al., 2009; Chung and Kwon, 2015; Kim and Kuby, 2012; Capar and Kuby, 2011; Kuby and Lim, 2005; Wang and Lin, 2013; 2009).

A recent study has found that travel range, cost, and charging infrastructure are the three main concerns for potential electric vehicle buyers (Egbue and Long, 2012). Even though the purchase cost is rapidly decreased, the other two concerns are still highly relevant. In particular, a long distance trip with inter-state travel as indicated by Franke et al. (2012) may induce stronger *range anxiety* because drivers need to make the right decisions regarding the distance that their cars can travel with their current battery status by determining a recharging plan with repeated recharging. Most previous studies have assumed that a recharging plan can be made with a simple rule that an electric vehicle can always travel a given distance R, i.e., a *deterministic* approach. Given the uncertain traffic condition in which electric vehicles operate, such deterministic assumption is unrealistic (Xie et al., 2017). In Nordic countries where electric vehicles are quite popular, a study has found that travel ranges of electric vehicles can vary by up to 60% depending on weather, temperature, use of in-car heater, average speed, and so on (Neubauer et al., 2013). In fact, the uncertainty of travel range (or mileage) is not limited to the electric vehicles. Conventional fuel engine vehicles also can travel considerably different distances with the same amount of fuel. Actual mileages recorded by real-world drivers can be found at http://www.fuelly.com/. Fig. 1 shows that travel ranges of the same types of vehicles can be dramatically different. An electric car maker Tesla published a detailed travel range report (Straubel, 2015), whose results showed that the travel range might be reduced by half when the speed is doubled (see Fig. 2). Tesla even provides an on-line range calculator (https://www.tesla.com/models), in which potential buyers can estimate the travel range by changing speed, outside temperature, operation of air conditioner, and/or types of tires. Especially, considering the uncertain travel range is critical for an infrastructure planner because it enables the planner to take into account the various driving conditions of drivers.

In this paper, we propose an optimization algorithm based on the extended FRLM with a probabilistic consideration of travel range. Hereafter, we refer to our approach as a flow-refueling location model with an uncertain travel range (FRLMwU). The resulting mathematical formulation for FRLMwU is a mixed-integer nonlinear programming (MINLP) problem. Due to difficulty in solving the mathematical formulation using off-the-shelf MINLP solvers, we develop a novel decomposition approach by combining Benders decomposition and column generation which we name it as *Benders-and-Price*.

This study contributes to the literature as follows: (1) by introducing the FRLMwU, we propose a method to incorporate probabilistic travel range into an optimal charging station location problem, (2) we present a mathematical formulation for the FRLMwU, (3) we develop the Benders-and-Price approach which is significantly faster than off-the-shelf solvers in solving real-life sized problems, (4) we also show that some classes of mixed integer nonlinear programming problem could be solved by using bi-objective optimization methodology, and (5) we conduct extensive computational experiments on networks including a real-life highway road network.

We first review the literature, with a special emphasis on the FRLM, in Section 2. The probabilistic expanded network and the mathematical formulation of the FRLMwU are introduced in Section 3. The Benders-and-Price algorithm to solve the FRLMwU is described in Section 4. Section 5 discusses computational experiments and analyzes obtained results. Finally, the conclusion is given in Section 6.



(a) Honda Civic: data obtained from 9,666 real-world drivers.

(b) Toyota Prius: data obtained from 6,067 real-world drivers.

Fig. 1. Distributions of MPG (miles per gallon) of Honda Civic and Toyota Prius (Graphs taken from http://www.fuelly.com/).

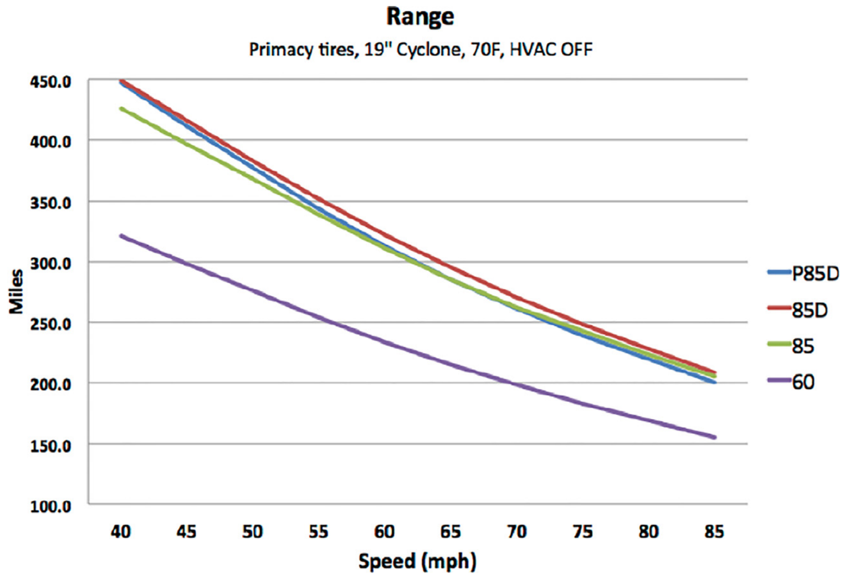


Fig. 2. Changes of travel range with different speeds of Tesla Model S (Straubel, 2015).

2. Literature review

As mentioned briefly in the previous section, the problem of locating recharging stations is closely related to problems in covering some components of a given graph. In graph theory, there are three main types of covering problems: node, arc, and path (Church and ReVelle, 1974). For example, in problems regarding facility location, a typical facility has some effective service ranges and nodes (or arcs) that should be within the range to be served, i.e., we need to cover nodes or arcs with the facilities. On the other hand, the path-covering problem involves paths that are considered as captured (or covered) if the facility is located at a node or arc used on the paths. Hodgson (1990) has proposed a model to capture paths (i.e., flows) on the network and named it flow capturing location model (FCLM). One notable difference with node or arc covering problems is that the facility that captures the flows does not have any effective ranges. It intercepts flows passing the facility instead. This seems to be an appropriate approach to model the behavior of drivers of electric vehicles.

One related line of research is the location and routing problem (LRP) where the location of facilities and routes of vehicles should be determined at the same time (see Nagy and Salhi, 2007 for an extensive survey). Worley et al. (2012) have presented an integer programming formulation for the problem of locating charging stations. They also designed EV routes. Lu et al. (2016) have investigated a resource-recharging station location and routing problem (RRS-LRP) to address various resources consumed as vehicles traveled. They proposed a resource-space-time network representation to generalize the well-known dynamic programming approach. A Lagrangian relaxation based heuristic algorithm utilizing the proposed resource-space-time network was developed. In this line of research, typically vehicles must leave from and return to the depot after visiting a number of customers, making the problem a variant of vehicle routing and/or traveling salesman problem. We notice that, contrast to the LRP models, the capturing or covering models generally do not have depots. In the capturing or covering models, every demand (or vehicle) has its own origin and destination. The goal of a demand (or vehicle) is not to visit customers during travel, but to complete the travel to the destination at minimum cost.

Capturing a flow means that an electric vehicle should be able to travel the entire path of the flow. Due to limited travel range of electric vehicles, it is sometimes impossible to capture the flow with a single facility. In this case, an electric vehicle should be recharged several times during a trip. This critical constraint must be taken into account. Kuby and Lim (2005) have addressed this by introducing the flow refueling location model (FRLM). They first divided the given path to multiple segments of sub-paths in which the distance of each segment was shorter than or equal to the travel range of electric vehicles. They then proposed a mathematical formulation to choose a specific segmentation for each flow path to ensure the installation of recharging stations at every joint node between two consequent segments. Two slightly different objective functions (maximizing the total sum of flows covered while minimizing the number of recharging stations) were considered in their study.

FRLM has been extended by many researchers to reflect different needs and assumptions. Originally, Kuby and Lim (2005) have assumed that the driver will only use the shortest route from the origin to the destination. However, Kim and Kuby (2012) have investigated the impact of "deviation-flow" when drivers do not follow the shortest path. They distributed the flow into multiple deviation paths depending on the difference (or deviation) in distance between the deviation path and the shortest distance. Recharging patterns were then considered for each deviation path with a distributed flow demand. Since the number of recharging patterns can grow exponentially for each path, considering more paths makes the problem even harder to solve. Such studies have assumed that recharging stations have an unlimited capacity. Upchurch et al. (2009) have considered the limited capacity of recharging facilities. They have ensured that the total sum of the flow using the facility does not exceed the facility's recharging capacity. This can be expressed using a knapsack type constraint. However, this model considers the capacity of a recharging station as a constant, not a decision variable.

These models presented above are based on full enumeration of recharging patterns. Before solving a mathematical formulation, one must generate all feasible patterns for each path, resulting in huge problem when the number of ODs is large. To overcome this difficulty, several mathematical formulations have been proposed to embed the recharging constraint. Wang (2007) has developed a mathematical formulation by embedding recharging decision variables and constraints to skip the generation of recharging patterns. Slightly different models are also used by the same authors for different applications (Wang, 2008; Wang and Lin, 2009). Recharging time is another issue that deserves serious study because fast-charging devices are more expensive than slow-charging ones. Therefore, the number of fast-charging devices should be maintained to a minimum to reduce the total investment. Wang and Lin (2013) have considered multiple types of charging stations with different charging speeds. Their computational experiments have showed that a mixed-station approach has advantages in terms of cost-effectiveness. Capar et al. (2013) have presented a new formulation of FRLM based on covering arcs that comprise each path. This formulation was later extended by Zhang et al. (2017) to account for the multi-period FRLM under the consideration of demand dynamics. Yıldız et al. (2016) have developed a branch-and-price approach for the path covering problem allowing deviations from the shortest path. The aim of that approach was to reduce the computational time by delaying the generation of recharging patterns. Path segments (i.e., a set of paths between two recharges) are generated on-the-fly by column generation.

FRLM involves two types of distinct decisions: location and route. Benders decomposition separates two different decision variables from the problem, resulting in two distinct optimization problems called Benders master and Benders subproblem, respectively. Zheng and Peeta (2017) have considered a variant of FRLM with a restriction on the maximum possible recharges in a single route using a Benders decomposition to solve the problem. They reported that the Benders decomposition approach could solve real-world instance in a reasonable amount of time. However, no performance comparison with other approaches was reported in that study. Another study by Arslan and Karaşan (2016) has utilized Benders decomposition to solve FRLM. They tried several alternative Benders cut generation techniques. Computational results indicated that Pareto-optimal cuts (originally developed by Magnanti and Wong, 1981) were the most effective in terms of computational times. Again, no comparison with other approaches was given in that study.

A novel approach utilizing a clever reformulation technique was proposed by MirHassani and Ebrazi (2012). They constructed a so-called "expanded network" by adding auxiliary arcs to the network when the distance between two nodes connected by the arc was not longer than the travel range. Each of these augmented arcs represents a segment of the recharging pattern of the FRLM approach. Hence, the recharging constraint can be expressed using well-known flow balance constraints where the existence of flow on the expanded network guarantees a feasible recharging pattern. Computational results showed that their approach was faster than the formulation proposed by Wang and Lin (2009) by multiple orders of magnitude. The major reason for the superiority of their approach is that they could remove the big-M constants from the formulation that often yields poor linear relaxation bounds at the expense of increased network size. Chung and Kwon (2015) have adopted an expanded network approach to solve the multi-period planning problem on Korean expressway network.

These models mentioned above can be seen as deterministic because data defining the problem are assumed to be certain. Unfortunately, there are few studies concerning uncertainty in the context of infrastructure planning for electric vehicles. One of a few exceptions consists of a robust optimization model proposed by Miralinaghi et al. (2017). In that study, the source of uncertainty was in route choice and demands. They proposed a mathematical formulation for the problem and employed a cutting-plane method to efficiently solve the problem. Lee et al. (2015) have considered a case where the initial battery state at the origin is uncertain. They presented a user equilibrium-state-based algorithm assuming that the initial battery status followed a certain probabilistic distribution. This model again assumed that the travel range was deterministic and that only the initial status was probabilistic, resulting in a split of the flow depending on the initial status. However, the flow's path became deterministic after splitting. A two-stage stochastic approach was developed by Faridimehr et al. (2017), where uncertainties in charging patterns, demand, and driver's behavior were considered. They developed a sample average approximation (SAA) based heuristic method and conducted a case study in Detroit midtown area in Michigan, USA. Their computational study showed that the proposed heuristic method could produce near-optimal solutions for large-scale problems. The uncertainty of travel range has also been addressed by de Vries and Duijzer (2017). The driving range was assumed as a random variable, not a fixed constant. Recognizing that joint probability of driving ranges could yield a highly non-linear formulation, they had to introduce restrictive assumptions with an approximation scheme. The resulting formulation was a relatively simple mixed integer programming (MIP) problem which was solved with an off-the-shelf MIP solver. Even though treatment of uncertain travel range was rather simplified, their computational experiments showed that the stochastic approach had clear merits over the deterministic approach. To the best of our knowledge, no studies have presented an exact algorithm addressing uncertainty in terms of travel range.

In reality, the travel range of electric vehicles varies for several reasons. The uncertainty regarding the travel range is probably the source of range anxiety because drivers feel more comfortable when they are certain of the range, as with daily trips between home and work (Franke et al., 2012). By definition, the length of each segment in a deterministic approach is below the travel range R, so the trip is always achievable. We relax this assumption by assigning a *probability of reachability* to each segment of a recharging pattern. Our probabilistic consideration is achieved by constructing a *probabilistic expanded network* to be explained in the following section.

3. FRLM with uncertain travel range

This section presents a formal definition of FRLMwU. We first introduce a probabilistic expanded network to address the recharging requirements for electric vehicles with uncertainty in their travel ranges. Then, a mathematical formulation is developed using the proposed probabilistic expanded networks.

3.1. Probabilistic expanded networks

First, we briefly review the expanded network proposed by MirHassani and Ebrazi (2012). We then develop a method to incorporate the uncertainty of the travel range into the expanded network. Expanded network models of electric vehicles' recharging requirements are based on the following assumptions:

(A1) A driver takes the shortest path from the origin to destination.
(A2) At the origin, half of the battery of the electric vehicle is full.
(A3) At the destination, the battery of the electric vehicle must not be less than half of its capacity.
(A4) The travel range of an electric vehicle is fixed and deterministic.

These assumptions were first proposed by Kuby and Lim (2005). They have been used extensively in the literature (Kuby et al., 2009; MirHassani and Ebrazi, 2012; Chung and Kwon, 2015; Kim and Kuby, 2012; Capar and Kuby, 2011; Wang and Lin, 2009). In particular, the third assumption (A3) implies that the travel can be round-trip, which is essentially a restatement of the second assumption. The last assumption (A4) does not mean that travel ranges of all types of electric vehicles are the same. If there are several types of electric vehicles with different travel ranges, one can define a separate traffic demand for each type of electric vehicle. This means that an electric vehicle can *always* travel a distance R, no more no less, where R is a given specification of the vehicle's travel range after a full charge.

We consider a given road network on which electric vehicles travel and charging stations are built, denoted as $G(N, A)$, where N and A are sets of nodes and arcs on the network, respectively. In this study, recharging stations are envisioned to be built along intercity highway network (or networks of similar size). This is because EV's travel range is extended significantly nowadays. Thus, drivers are rarely anxious about travel distance in urban transport. However, long distance trips along the intercity highway network still requires multiple recharging, thereby inducing range anxiety. It should be relieved by charging stations on the way. The travel distance between node i and j is denoted by d_{ij}. Without a loss of generality, we assume that any node in N can be a candidate for a recharging facility. Otherwise, we can remove such node and connect any pair of adjacent nodes using new arcs whose travel distance equals to the shortest distance between them. Let K be a set of origin-destination (OD) pairs (or demands) and D_k be the total number of electric vehicles corresponding to demand k. For any $k \in K$, the construction of the expanded network $G_k(N_k, A_k)$ involves the following steps:

Step 1 Obtain the shortest path p_k from the origin to the destination of demand k on the network $G(N, A)$. Construct a graph $G_k(N_k, A_k)$, where N_k is the set of nodes on path p_k and A_k is the set of arcs on path p_k.
Step 2 Create a source node s_k and a sink node t_k. Let $N_k \leftarrow N_k \cup \{s_k, t_k\}$ and $A_k \leftarrow A_k \cup \{(s_k, i), (j, t_k)\}$, where i and j are the first and last nodes on path p_k, respectively. Arcs (s_k, i) and (j, t_k) have zero distance.
Step 3 Let $A_k \leftarrow A_k \cup \{(s_k, i) \mid$ the travel distance from s_k to $i \in N_k \backslash \{s_k\}$ is no more than $\frac{R}{2}\} \cup \{(i, t_k) \mid$ the travel distance from $i \in N_k \backslash \{t_k\}$ to t_k is no more than $\frac{R}{2}\}$.
Step 4 Let $A_k \leftarrow A_k \cup \{(i, j) \mid$ the travel distance from $i \in N_k \backslash \{s_k\}$ to $j \in N_k \backslash \{t_k\}$ is no more than $R\}$.

Fig. 3 demonstrates the above procedure. Any path from s_k to t_k represents a feasible *recharging pattern* of an electric vehicle for demand k. Whenever an electric vehicle visits a node in $N_k \cap N$, the battery of the vehicle is fully recharged. For example, a path (s_k, b, c, t_k) in Fig. 3d means that the battery should be recharged two times at b and c. This also implies that the recharging stations should be installed at those nodes. Notice that **Steps 3** and **4** are based on (A4), in which added arcs mean that the electric vehicle can bypass some recharging stations only if it can reach the next recharging station within the travel range. Moreover, it can always travel distance R after complete recharging.

In reality, the travel range of electric vehicles can vary due to many factors beyond the drivers' control. In order to address the uncertainty of travel range, we replace assumption (A4) with the following:

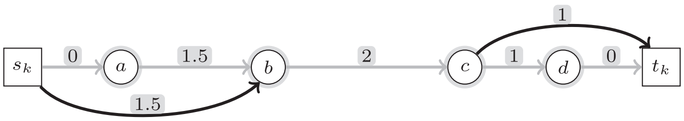
(A4′) The travel range of an electric vehicle can vary and follows some probability distribution.

Let $\tilde{z}$ denote a random variable representing the maximum travel distance of an electric vehicle with a fully charged battery. We assume that $\tilde{z}$ follows a probability distribution $p(z)$ such that $\lim_{z \to \infty} p(z) = 0$ and $\int_0^\infty p(z)dz = 1$. Then, we define a *reachability* function

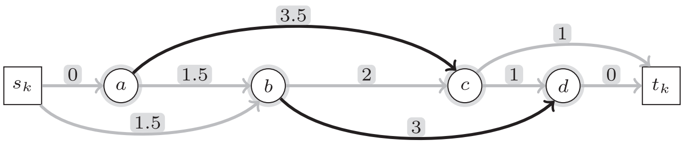
$$R(x) = P(\tilde{z} \geq x) = \int_x^\infty p(z)dz. \quad (1)$$

(a) **Step 1**. Shortest path for a demand $k \in K$ that originates at a and is bound to d. Let N_k be the nodes included in the shortest path, i.e., $N_k = \{a, b, c, d\}$, $A_k = \{(a, b), (b, c), (c, d)\}$

(b) **Step 2**. Create a source node s_k and a sink node t_k and connect them with arcs of distance zero to the first and the last nodes of the path, respectively. $N_k = \{s_k, a, b, c, d, t_k\}$, $A_k = \{(s_k, a), (a, b), (b, c), (c, d), (d, t_k)\}$



(c) **Step 3**. Connect s_k and $i \in N_k$ if the distance between them is no more than $\frac{R}{2}$. And also connect $i \in N_k$ and t_k if the distance between them is no more than $\frac{R}{2}$. $N_k = \{s_k, a, b, c, d, t_k\}$, $A_k = \{(s_k, a), (a, b), (b, c), (c, d), (d, t_k), (s_k, b), (c, t_k)\}$



(d) **Step 4**. Connect two nodes $i \in N_k$ and $j \in N_k$ if the distance between them is no more than the travel range $R = 4$. $N_k = \{s_k, a, b, c, d, t_k\}$, $A_k = \{(s_k, a), (a, b), (b, c), (c, d), (d, t_k), (s_k, b), (c, t_k), (a, c), (b, d)\}$

Fig. 3. An expanded network for demand from node a to node d ($R = 4$).

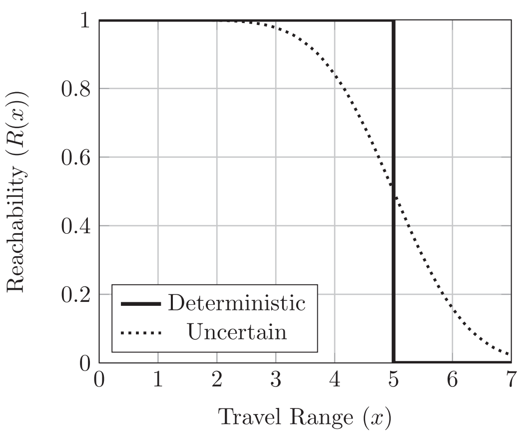


Fig. 4. Examples of reachability functions.

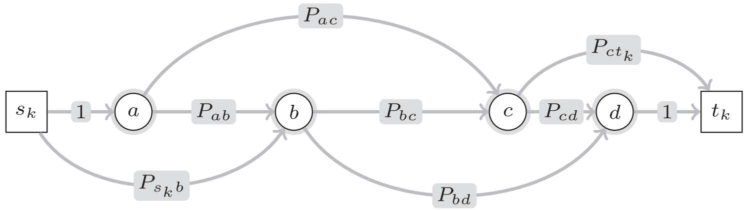


Fig. 5. A probabilistic expanded network for demand from node a to node d.

This represents the probability that an electric vehicle with a completely charged battery can reach at least distance x. Actually, assumption (A4) is a special case of (A4′) when $R(x)$ is 1 if $x \le R$, 0 otherwise. Fig. 4 shows two reachability functions: the deterministic case and the uncertain case of $p(x) = N(5, 1)$, where $N(\mu, \sigma)$ is a normal distribution with mean μ and standard deviation σ. Reachability functions in Fig. 4 are just two simplest cases shown for illustration purpose. In reality, the travel range of an EV is usually bounded above from some upper bound U which justifies the use of a truncated probability distribution. In this case, the reachability function $R(x)$ can be defined as following:

$$R(x) = P(\tilde{z} \ge x) = \int_x^U p(z)dz. \tag{2}$$

Estimating a precise real-world reachability function warrants state-of-the-art methodologies, including Big-data analysis, regression, data-mining, and so on, which is beyond the primary interests of this paper. However, any given reachability function can be used in the following exposition. We also note that the reachability function represents not only the possibility of reaching destination of a single driver but also the distribution of *effective travel ranges* of many electric vehicle drivers.

Now, we propose a probabilistic expanded network $\tilde{G}_k(N_k, A_k)$ for demand k by utilizing the reachability function. The following step assigns to each arc $(i, j) \in A_k$ a reachability P_{ij} to create the probabilistic expanded network.

Step 5 For each arc $(i, j) \in A_k$, assign $P_{ij} = R(d_{ij})$.

Fig. 5 graphically illustrates the probabilistic expanded network. In the figure, $P_{s_k a} = P_{dt_k} = 1$ because distances of the corresponding arcs are all 0. For a given path on the probabilistic expanded network, it is easy to calculate joint reachability of the path. For example, a path (s_k, b, c, t_k) has a reachability of $P_{s_k b} \cdot P_{bc} \cdot P_{ct_k}$. Under the deterministic assumption (A4), e.g., in Fig. 3d, two paths (s_k, b, c, t_k) and (s_k, a, c, t_k) are essentially equivalent in the sense that both require recharging twice and the total travel distance is the same. However, under assumption (A4′), they are different in terms of reachability: $P_{s_k b} \cdot P_{bc} \cdot P_{ct_k}$ vs. $P_{s_k a} \cdot P_{ac} \cdot P_{ct_k}$. Let q be a path on a probabilistic expanded network $\tilde{G}_k(N_k, A_k)$ and assume that a recharging pattern of electric vehicles for demand k follows path q. Each electric vehicle may or may not complete the travel by using path q. We denote random variable $\tilde{D}_k$ as the total number of electric vehicles that successfully complete the journey. Then, the expected number of electric vehicles that would reach the destination successfully is given as:

$$\mathbb{E}[\tilde{D}_k] = D_k \cdot \prod_{(i,j) \in A(q)} P_{ij}, \tag{3}$$

where $A(q)$ is a set of arcs used on path q on $\tilde{G}_k(N_k, A_k)$.

For an infrastructure planner, the expected demand covered can be interpreted as the number of electric vehicle drivers who may reach the destination with their driving behaviors (or habits). As mentioned in Section 1 with Figs. 1 and 2, each electric vehicle driver has his/her own travel range determined by his/her unique driving conditions. Some drivers can travel longer while other travel shorter. The reachability function represents how many drivers can travel a given distance. For example, having 30% of probability of reaching the destination successfully means that 30% drivers can complete the travel without much range anxiety. In other words, for 30% of drivers, the probability of reaching the destination is much higher.

Fig. 6 provides an example of the probabilistic expanded network where the random travel distance follows a normal distribution $N(3, \frac{3}{5})$. This implies that for an electric vehicle, the probabilities to travel more than 1, 2, and 3 are (approximately) 1, 0.95, and 0.5, respectively. If the distance of an arc is too large, it would be desirable to have recharging stations in the middle of arc. To represent such a case, some auxiliary nodes referred to as "midarc nodes" can be added to represent additional candidate locations for recharging facilities (Kuby and Lim, 2007; MirHassani and Ebrazi, 2012). In Fig. 6, three midarc nodes are added with distances between any two possible charging locations being 1. In this example, we can have $\mathbb{E}[\tilde{D}_k] = D_k$ by visiting 5 nodes ("the first midarc node between a and b" $\rightarrow$ "the second midarc node between a and b" $\rightarrow$ b $\rightarrow$ c $\rightarrow$ "the first midarc node between c and d") for recharging, which requires 5 charging stations to be installed. With a smaller number of recharging stations, e.g. 2, $\mathbb{E}[\tilde{D}_k] = D_k \cdot 0.5 \cdot 0.95$ can be seen to be the best recharging pattern maximizing the reachability, e.g., "the first midarc node between a and b" $\rightarrow$ c.

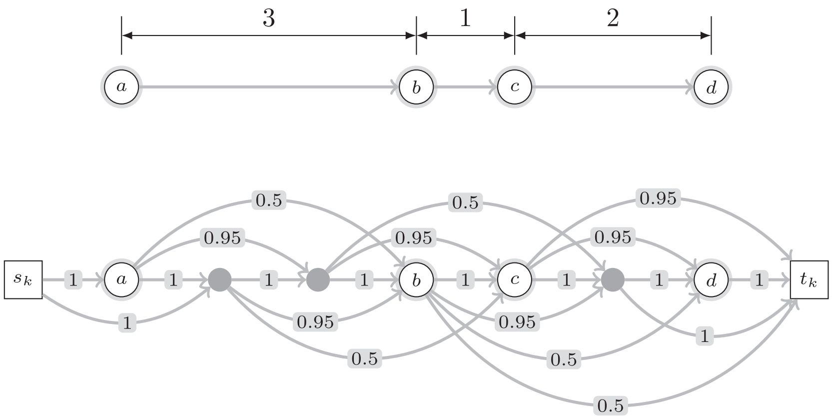


Fig. 6. An example of a probabilistic expanded network for demand from node a to node d with $p(x) = N(3, \frac{3}{5})$. Auxiliary gray nodes ("midarc" nodes) are added at the middle of the arcs by every single distance.

Table 1
Nomenclature.

Sets	
K	Set of OD pairs (or demands)
$G(N, A)$	Given road network, where N and A are sets of nodes and arcs
$G_k(N_k, A_k)$	Expanded network of demand k, where N_k and A_k are sets of nodes and arcs
$\tilde{G}_k(N_k, A_k)$	Probabilistic expanded network of demand k
L	Set of Benders cuts
Q_k	Set of all OD paths of demand k on $\tilde{G}_k(N_k, A_k)$
Parameters	
D_k	Total number of electric vehicles for demand k
C_i	Installation cost of recharging station at node i
B	Total budget
P_{ij}	Reachability of arc (i, j)
α_0^l, α^l	Coefficients of Benders cut l
P_q	Reachability of path q
δ_q^i	Indicator function on whether path q visit node i
Decision variables	
x_{ij}^k	Binary variable indicating if arc (i, j) is used for the path of demand k
y_i	Binary variable indicating if a charging station is built at node i
z_k	Binary variable indicating if the flow for demand k is covered
p_k	Reachability of the path for demand k
w_{ij}^k	Variable on the reachability of arc (i, j) for demand k (0 or P_{ij})
θ	Auxiliary variable used for Benders master problem
f_q^k	Binary variable indicating if path q is selected for demand k

3.2. Mathematical formulation

In this subsection, we present the mathematical formulation of FRLMwU. Table 1 provides the notation used in formulations introduced in this paper. Let C_i be the installation cost of a recharging station at node $i \in N$. The total budget B restricts the number of charging stations to be installed. For any $k \in K$, we construct the corresponding probabilistic expanded network $\tilde{G}_k(N_k, A_k)$. Then, the FRLMwU can be formulated as follows:

$$(\text{P}) \quad \max \quad \sum_{k \in K} D_k p_k z_k \tag{4}$$

$$\text{s.t.} \quad \sum_{(i,j) \in A_k} x_{ij}^k - \sum_{(j,i) \in A_k} x_{ji}^k = \begin{cases} z_k, & \text{if } i = s_k, \\ -z_k, & \text{if } i = t_k, \\ 0, & \text{otherwise} \end{cases}, \quad \forall i \in N_k, k \in K, \tag{5}$$

$$\sum_{(j,i) \in A_k} x_{ji}^k \le y_i, \quad \forall i \in N, k \in K, \tag{6}$$

$$p_k = \prod_{(i,j) \in A_k} w_{ij}^k, \quad \forall k \in K, \tag{7}$$

$$w_{ij}^k = P_{ij}x_{ij}^k + 1 - x_{ij}^k, \quad \forall (i,j) \in A_k, k \in K, \tag{8}$$

$$\sum_{i\in N} C_i y_i \le B, \tag{9}$$

$$x_{ij}^k \in \{0,1\}, \quad \forall (i,j) \in A_k, k \in K, \tag{10}$$

$$y_i \in \{0,1\}, \quad \forall i \in N, \tag{11}$$

$$z_k \in \{0,1\}, \quad \forall k \in K. \tag{12}$$

The decision variable z_k is 1 if there is a path from s_k to t_k for demand $k \in K$, 0 otherwise. If $z_k = 1$, the flow for demand k is covered. The path for demand $k \in K$ is represented by x_{ij}^k which has a value of 1 if the path uses arc $(i, j) \in A_k$, 0 otherwise. The decision variable y_i determines whether a charging station should be installed at $i \in N$. The reachability of the path for $k \in K$ is given by decision variable p_k. The decision variable w_{ij}^k equals to P_{ij} if the path for $k \in K$ uses arc $(i, j) \in A_k$, 1 otherwise. Note that, from constraints (7) and (8), for a selected path q of $k \in K$,

$$p_k = \prod_{(i,j)\in A_k} w_{ij}^k = \prod_{(i,j)\in A(q)} P_{ij} \tag{13}$$

holds.

The objective function (4) represents the total expected demand that is covered. Constraints (5) ensures a path for each $k \in K$ if it is covered. A path for $k \in K$ cannot visit node $i \in N$ if there is no charging station at i, which is guaranteed by constraints (6). Constraints (7) and (8) together calculate the reachability of $k \in K$. Constraints (9) restricts the total installation cost for charging stations to be no more than the budget. Note that the objective function and constraints (7) are nonlinear and not even convex, making it difficult to solve the formulation (P) with off-the-shelf solvers. Hence, in the next section, we propose an efficient decomposition approach to solve the formulation (P) to optimality.

4. Benders-and-Price approach

In this section, we develop a method to solve formulation (P) using Benders decomposition. The problem is decomposed into two distinct problems: (1) Benders master problem that determines where charging stations should be deployed, and (2) Benders subproblem that determines which OD pair (demand) to cover and recharging patterns for such OD pairs. Benders subproblem will be shown to be a mixed integer nonlinear programming problem that is computationally demanding. In order to overcome the computational burden, we will propose a column generation approach to be incorporated to solve the Benders subproblem. The column generation is based on the so-called *pricing* of the columns. Therefore, we call our approach “Benders-and-Price” approach.

4.1. Benders master problem

We now consider the following Benders master problem.

$$(\text{BM}) \quad \max \quad \theta \tag{14}$$

$$\text{s.t.} \quad \theta \le \sum_{k\in K} D_k, \tag{15}$$

$$\theta \le \alpha_0^l + \sum_{i\in N} \alpha_i^l y_i, \quad \forall l \in L, \tag{16}$$

$$\sum_{i\in N} C_i y_i \le B, \tag{17}$$

$$y_i \in \{0,1\}, \quad \forall i \in N, \tag{18}$$

where L is a set of Benders cuts added so far. Each Benders cut $l \in L$ is represented by α_0^l and α^l, where $\alpha_0^l \in \mathbb{R}$ and $\alpha^l \in \mathbb{R}^{|N|}$. At each Benders iteration, a new Benders cut $(\alpha_0^{l^*}, \alpha^{l^*})$ is generated in the Benders subproblem and is added to (BM), i.e.,

$L \leftarrow L \cup \{l^*\}$. Clearly, the objective value of (P) is bounded below from the total sum of demands $\Sigma_{k\in K} D_k$ and constraint (15) is added to prevent (BM) from becoming unbounded. Note that (BM) has decision variable y that determines the locations of the recharging stations while an auxiliary decision variable θ determines the objective value of (P). Formulation (BM) is a MIP problem that can be solved using off-the-shelf MIP solvers such as Cplex. We solve (BM) to optimality, and the obtained solution y^* is provided to the Benders subproblem. If the Benders subproblem finds a new Benders cut with a given y^*, the Benders cut is appended to (BM) and the procedure is iterated again. The iteration is terminated when Benders subproblem does not find any more Benders cut, and the solution of (BM) at the last iteration is an optimal solution to (P).

4.2. Benders subproblem

Let y^* and θ^* be the optimal solution of (BM) at a certain iteration. We need to solve Benders subproblem (BS) fixing $y = y^*$ to identify the Benders cuts. In general, there are two types of Benders cuts: *optimality* and *feasibility* cuts. We note that the following Bender subproblem is always feasible. Therefore, Benders feasibility cuts do not need to be generated.

$$\text{(BS)} \quad \xi(y^*) = \max \quad \sum_{k\in K} D_k p_k z_k \tag{19}$$

$$\text{s.t.} \quad \sum_{(i,j)\in A_k} x_{ij}^k - \sum_{(j,i)\in A_k} x_{ji}^k = \begin{cases} z_k, & \text{if } i = s_k, \\ -z_k, & \text{if } i = t_k, \\ 0, & \text{otherwise} \end{cases}, \quad \forall i \in N_k, k \in K, \tag{20}$$

$$\sum_{(j,i)\in A_k} x_{ji}^k \le y_i^*, \quad \forall i \in N, k \in K, \tag{21}$$

$$p_k = \prod_{(i,j)\in A_k} w_{ij}^k, \quad \forall k \in K, \tag{22}$$

$$w_{ij}^k = P_{ij} x_{ij}^k + 1 - x_{ij}^k, \quad \forall (i,j) \in A_k, k \in K, \tag{23}$$

$$x_{ij}^k \in \{0,1\}, \quad \forall (i,j) \in A_k, k \in K, \tag{24}$$

$$z_k \in \{0,1\}, \quad \forall k \in K. \tag{25}$$

Note that (BS) is still a mixed integer nonlinear programming problem. To obtain a Benders cut, the objective function of (BS) should be given in terms of solution of Benders master problem, y^*, which is easily achievable via the dual of (BS) if the strong duality holds. Since (BS) is neither a linear programming problem nor a convex programming problem, the strong duality generally does not hold. We now present a linearized Benders subproblem by introducing new decision variables to select the path for each demand $k \in K$ on the probabilistic expanded network $\tilde{G}_k(N_k, A_k)$. Let Q_k denote the set of all paths from s_k to t_k on $\tilde{G}_k(N_k, A_k)$. For any path $q \in Q_k$, we calculate the reachability of the path q and denote it by P_q, i.e., $P_q := \prod_{(i,j)\in A(q)} P_{ij}$. Then the so-called path based formulation for (BS) is stated as:

$$\text{(BSP)} \quad \max \quad \sum_{k\in K} \sum_{q\in Q_k} D_k P_q f_q^k \tag{26}$$

$$\text{s.t.} \quad \sum_{q\in Q_k} f_q^k \le 1, \quad \forall k \in K, \tag{27}$$

$$\sum_{q\in Q_k} \delta_q^i f_q^k \le y_i^*, \quad \forall i \in N, k \in K, \tag{28}$$

$$f_q^k \in \{0,1\}, \quad \forall q \in Q_k, k \in K, \tag{29}$$

where indicator function δ_q^i is 1 if a path $q \in Q_k$ visits node $i \in N$, 0 otherwise. Decision variable f_q^k is 1 if path $q \in Q_k$ is selected for $k \in K$, 0 otherwise. Note that (BSP) is an integer programming problem. It is easy to see that the linear programming (LP) relaxation of (BSP) has an integral polyhedron.

Proposition 1. *If y^* is an integer, the feasible set of the linear programming (LP) relaxation of (BSP) is an integral polyhedron.*

The above proposition implies that we can replace the integrality constraint (29) with $f_q^k \geq 0$ for all $q \in Q_k$, $k \in K$. Let $\pi \in \mathbb{R}_+^{|K|}$ and $\phi \in \mathbb{R}_+^{|N||K|}$ be the dual variables associated with constraints (27) and (28), respectively. Since LP relaxation of (BSP) has the same objective function as (BS), the following proposition is easily seen to hold.

Proposition 2. *For a given integer solution y^*, the optimal objective value of (BS) is given as:*

$$\xi(y^*) = \sum_{k\in K} \hat{\pi}_k + \sum_{k\in K}\sum_{i\in N} \hat{\phi}_i^k y_i^*,$$

where $\hat{\pi}$ and $\hat{\phi}$ are the dual optimal solution of LP relaxation of (BSP).

Note that there can be exponentially many paths for a given demand. This makes the problem (BSP) intractable. Therefore, in the next subsection, we develop a column generation approach for (BSP) to generate a subset of paths on an as-needed basis instead of having to solve (BSP) at once with all paths.

4.2.1. Column generation

First, we note that (BSP) is separable for $k \in K$. Then, for a given integer solution y^* and any demand $k \in K$, we consider the following problem with a restricted set of paths $\hat{Q}_k \subseteq Q_k$.

$$\text{(BSP-}k\text{)} \quad \xi_k(y^*) = \max \sum_{q\in\hat{Q}_k} D_k P_q f_q \tag{30}$$

$$\text{s.t.} \quad \sum_{q\in\hat{Q}_k} f_q \leq 1, \tag{31}$$

$$\sum_{q\in\hat{Q}_k} \delta_q^i f_q \leq y_i^*, \quad \forall i \in N, \tag{32}$$

$$f_q \geq 0, \quad \forall q \in \hat{Q}_k. \tag{33}$$

Let $\pi_k \in \mathbb{R}_+$ and $\phi^k \in \mathbb{R}_+^{|N|}$ denote dual variables associated with (31) and (32), respectively. To solve (BSP-k), we have to repeatedly generate columns (or paths). The reduced cost of path q for the dual solution $(\hat{\pi}_k, \hat{\phi}^k)$ is given as

$$D^k P_q - \hat{\pi}_k - \sum_{j\in N|(i,j)\in A(q)} \hat{\phi}_j^k. \tag{34}$$

and the dual solution is optimal if

$$D^k P_q - \hat{\pi}_k - \sum_{j\in N|(i,j)\in A(q)} \hat{\phi}_j^k \leq 0, \quad \forall q \in Q_k. \tag{35}$$

Then we can solve the following column generation subproblem to find paths with positive reduced cost for each demand $k \in K$.

$$\text{(CG1-}k\text{)} \quad Z_k = \max \quad D_k p - \hat{\pi}_k - \sum_{j\in N}\sum_{(i,j)\in A_k} \hat{\phi}_j^k x_{ij} \tag{36}$$

$$\text{s.t.} \quad \sum_{(i,j)\in A_k} x_{ij} - \sum_{(j,i)\in A_k} x_{ji} = \begin{cases} 1, & \text{if } i = s_k, \\ -1, & \text{if } i = t_k, \\ 0, & \text{otherwise} \end{cases}, \quad \forall i \in N_k, \tag{37}$$

$$p = \prod_{(i,j)\in A_k} w_{ij}, \tag{38}$$

$$w_{ij} = P_{ij} x_{ij} + 1 - x_{ij}, \quad \forall (i,j) \in A_k, \tag{39}$$

$$x_{ij} \in \{0,1\}, \quad \forall (i,j) \in A_k. \tag{40}$$

Note that, at optimality, $w_{ij} = P_{ij}$ if the optimal path uses (i, j), i.e., $x_{ij} = 1$, $w_{ij} = 1$ otherwise. Therefore, by constraints (38), p is the joint probability of all arcs in the path successfully traveled. Note that, at some iteration, (BSP-k) is optimal if $Z_k \leq 0$. If $Z_k > 0$, the path defined by the optimal solution x^* of (CG1-k) becomes a new column that may improve the objective function of (BSP-k) because it has a positive reduced cost. Let q^* be such path with a positive reduced cost. We set $\hat{Q}_k \leftarrow \hat{Q}_k \cup \{q^*\}$ and solve (BSP-k) with an updated set of paths $\hat{Q}_k$. The column generation is terminated if we have

$Z_k \le 0$ for all $k \in K$. Since (CG1-k) is a nonlinear integer programming problem, it is difficult to solve the problem directly using an MIP solver. However, without the constraints for reachability p of a path, the problem becomes a classical shortest path problem that can be very efficiently solved. The difficulty comes from reachability p given as a multiplication of other decision variables w_{ij}.

The objective function (36) of (CG1-k) consists of two different criteria: (1) $D_k p$ measures how much the path is reachable. We prefer a higher p because it means that the recharging pattern represented by the solution would be less risky, i.e., a higher probability of making the whole travel, and (2) $-\sum_{i\in N}\hat{\phi}_i^k \sum_{(i,j)\in A_k} x_{ij}$ represents how well the path utilizes the recharging stations during the whole travel. Note that $\hat{\phi}_i^k$ will be 0 if $y_i^* = 1$, which means there is no penalty in using this node for recharging. If it is impossible to complete the travel to the destination due to a lack of recharging stations, some new recharging facilities should be deployed somewhere. This results in a penalty from positive $\hat{\phi}_i^k$. This observation leads us to the possibility of addressing those two objectives separately.

We now consider the following formulation of a *bi-objective shortest path problem* (BSP) (Raith and Ehrgott, 2009; Martins, 1984):

$$\text{(CG2-}k\text{)} \quad \min \quad Z_k(x) = \begin{cases} Z_k^1(x) = -\sum_{(i,j)\in A_k} (\log P_{ij}) x_{ij}, \\ Z_k^2(x) = \hat{\pi}_k + \sum_{j\in N}\sum_{(i,j)\in A_k} \hat{\phi}_j^k x_{ij}, \end{cases} \tag{41}$$

$$\text{s.t.} \quad \sum_{(i,j)\in A_k} x_{ij} - \sum_{(j,i)\in A_k} x_{ji} = \begin{cases} 1, & \text{if } i = s_k, \\ -1, & \text{if } i = t_k, \\ 0, & \text{otherwise} \end{cases}, \quad \forall i \in N_k, \tag{42}$$

$$x_{ij} \in \{0, 1\}, \quad \forall (i, j) \in A_k. \tag{43}$$

There are two independent objective functions $Z_k^1(x)$ and $Z_k^2(x)$ that often represents two objectives as trade-offs. We want to find a feasible solution for problem (CG2-k) for which one objective function cannot be improved without making the other objective worse. Such solutions are formally defined as:

Definitions 1. A feasible solution x^* of (CG2-k) is *efficient* if there is no other feasible solution $\hat{x}$ such that $Z_k^1(\hat{x}) < Z_k^1(x^*)$ and $Z_k^2(\hat{x}) < Z_k^2(x^*)$.

Let $X_E \subseteq X$ be the set of all efficient solutions, where X is the set of all feasible solution described by (42) and (43).

Theorem 1. *Let (x^*, w^*, p^*) be the optimal solution of (CG1-k), then $x^* \in X_E$.*

Proof. Assume that, for a contradiction, there is an optimal solution (x^*, w^*, p^*) of (CG1-k) such that $x^* \notin X_E$. Surely x^* is a feasible solution of (CG2-k) because it satisfies constraint (42) and (43). By Definition 1, there must be $\hat{x}$ that is feasible for (CG2-k) and $Z_k^1(\hat{x}) < Z_k^1(x^*)$ and $Z_k^2(\hat{x}) < Z_k^2(x^*)$. It is clear that there should be $\hat{p}$ and $\hat{w}$ such that $(\hat{x}, \hat{w}, \hat{p})$ is feasible for (CG1-k). Note that

$$\log \hat{p} = \log \prod_{(i,j)\in A_k | \hat{x} \text{ uses } (i,j)} P_{ij} = \sum_{(i,j)\in A_k | \hat{x}_{ij}=1} \log P_{ij} = -Z_k^1(\hat{x})$$

holds, which implies that $D_k \hat{p} > D_k p^*$ because $Z_k^1(\hat{x}) < Z_k^1(x^*)$. Moreover, we have

$$\hat{\pi}_k + \sum_{j\in N}\sum_{(i,j)\in A_k} \hat{\phi}_j^k \hat{x}_{ij} < \hat{\pi}_k + \sum_{j\in N}\sum_{(i,j)\in A_k} \hat{\phi}_j^k x_{ij}^*,$$

which implies that $(\hat{x}, \hat{w}, \hat{p})$ is strictly better than (x^*, w^*, p^*) for (CG1-k). This gives a contradiction, thus completing the proof. □

Various algorithms can be used to obtain X_E of (CG2-k), such as label correction and label setting algorithms (Skriver and Andersen, 2000; Sedeño-Noda and Raith, 2015; Raith and Ehrgott, 2009). Once X_E is obtained, we can enumerate (see Algorithm 1) all efficient solutions in X_E to find feasible solutions for (CG1-k) with positive reduced costs. It is easy to see that the following theorem holds.

Theorem 2. *Algorithm 1 finds a column with a positive reduced cost if it exists.*

We note that the purpose of the column generation is obtaining the dual optimal solution of (BSP-k). Sometimes a "primal" optimal solution of (BSP-k) can be found easily without solving the bi-objective shortest path problem. For example, it is easy to see that a (primal) optimal solution of (BSP-k) is a path $q \in Q_k$ with the largest P_q, which can be found by solving a shortest path problem with setting arc costs to $-\log P_{ij}$ for all $(i, j) \in A_k$. Even with the known the primal optimal solution, the current dual solution of the restricted master problem of (BSP-k) is not dual feasible unless it satisfies the optimality condition (35). If it is not dual feasible, a column (a constraint for the dual problem) is added and (BSP-k) is solved again to

Algorithm 1 Identifying columns with positive reduced costs for (CG1-k).

procedure ColumnGeneration($k, \hat{\pi}, \hat{\phi}$)
  Obtain X_E by solving (CG2-k) on probabilistic expanded network $\tilde{G}_k(N_k, A_k)$
  for $x \in X_E$ **do**
    q is the path represented by x
    $Z \leftarrow D_k e^{-Z_k^1(x)} - Z_k^2(x)$
    if $Z > 0$ **then**
      Add column q, i.e., $\hat{Q}_k \leftarrow \hat{Q}_k \cup \{q\}$
    end if
  end for
end procedure

find a different dual solution. Therefore, the column generation must be iterated while the optimality condition (35) is not satisfied, even the primal optimal solution was found earlier.

For any $k \in K$, if Algorithm 1 finds a column with a positive reduced cost, the column is added to (BSP-k). The column generation is terminated if there is no added column for all $k \in K$. Then, for a given solution (y^*, θ^*) of (BM), by Proposition 2 we have

$$\xi(y^*) = \sum_{k\in K} \xi_k(y^*) = \sum_{k\in K} \hat{\pi}_k + \sum_{k\in K}\sum_{i\in N} \hat{\phi}_i^k y_i^*. \tag{44}$$

A Benders cut is identified if

$$\theta^* > \xi(y^*) = \sum_{k\in K} \hat{\pi}_k + \sum_{k\in K}\sum_{i\in N} \hat{\phi}_i^k y_i^*, \tag{45}$$

where a new Benders cut $\hat{l}$ is defined as

$$\alpha_0^{\hat{l}} = \sum_{k\in K} \hat{\pi}_k, \tag{46}$$

$$\alpha_i^{\hat{l}} = \sum_{k\in K} \hat{\phi}_i^k, \quad \forall i \in N. \tag{47}$$

Therefore, the following constraint is added to (BM).

$$\theta \le \sum_{k\in K} \hat{\pi}_k + \sum_{i\in N}\sum_{k\in K} \hat{\phi}_i^k y_i, \tag{48}$$

and let $L \leftarrow L \cup \{\hat{l}\}$.

Note that the two objectives of (CG2-k) all have nonnegative coefficients because $P_{ij} \le 1$ for all $(i, j) \in A_k$ and $\hat{\phi}_i^k \ge 0$ for all $i \in N$. This will make obtaining X_E much easier because any solution $x \in X_E$ does not contain a cycle. Moreover, each column generation problem (BSP-k) can be independently solved. This allows for parallelized solving of many (BSP-k) problems for different ks.

5. Computational experiments

This section presents the results of the computational experiments. The proposed algorithm was implemented using Python and Cplex 12.6. All experiments were conducted on a MacPro machine with a 3.5 GHz Intel Xeon E5 and 32 GB of RAM.

5.1. Test networks

We used two distinct networks for computational experiments. The first network (Fig. 7) is a so-called 25NODE network that was originally proposed by Simchi-Levi and Berman (1988) for traveling salesperson problem. This network has been used extensively in the literature in the context of electric vehicles (MirHassani and Ebrazi, 2012; Kim and Kuby, 2012; Yıldız et al., 2016; Kuby and Lim, 2005). There are 25 nodes and 43 edges (86 arcs) in the network. Every node is considered as an origin, destination, and candidate location for the recharging facility, yielding a total of 300 ODs (i.e., $\binom{25}{2} = 300$). Distances of arcs range from 2 to 9. Midarc nodes for additional locations of recharging facilities were created so that the distance between any two candidate locations was 1. The total number of candidate locations for recharging facilities was 180 (25 OD nodes + 155 midarc nodes). The traffic volume (demand) for each OD was estimated using the gravity model as proposed by Hodgson (1990): Let W_i be the weight (or population) of node i, then the traffic D_{ij} between two nodes i and j was estimated as $(\frac{W_i W_j}{sd_{ij}})^{1.5}$ where sd_{ij} was the shortest distance between i and j on the network.

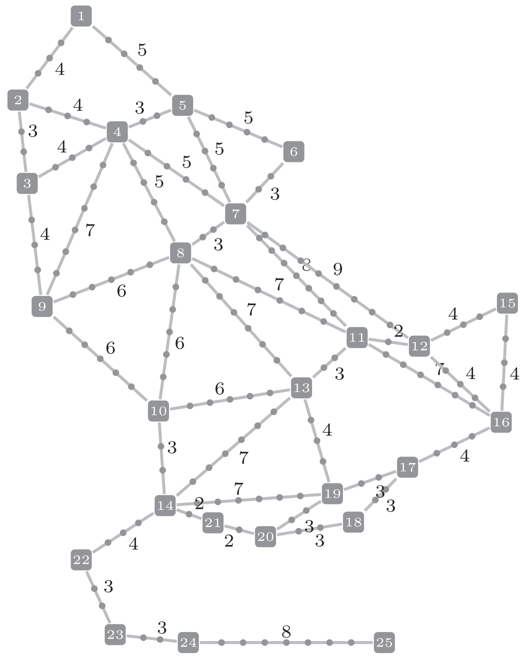


Fig. 7. The 25NODE network. Filled rectangular and small dot nodes are OD and midarc nodes, respectively.

The second network (Fig. 8) models the highway road network in Texas, US. We created the network by digitizing a map of Texas highways that is available online (https://www.tourtexas.com/texas-maps/map-of-texas-highways). There are 124 nodes and 238 edges (476 arcs) in this network. Every node is a candidate location for the recharging facility. Only 30 nodes are the origin and destination, yielding a total of 435 ODs (i.e., $\binom{30}{2} = 435$). A total of 516 midarc nodes were created using the same method as the 25NODE network. The number of candidates for recharging facilities was 640 (124 nodes + 516 midarc nodes). In Figs. 7 and 8, filled rectangular nodes, hollow rectangular nodes, and small gray dots represent OD nodes, intersection nodes, and midarc nodes, respectively. In the Texas Highway network, an arc with a distance of 10 corresponds to approximately 250 km in real-world. The demand for each OD was estimated using the same gravity model used for the 25NODE network, with population of the city (i.e., node) used as weight.

In this computational study, we assume that all nodes (OD nodes + midarc nodes) are candidates of the recharging facilities. In reality, however, locations of charging facilities can be more restrictive such as only interaction nodes. Determining candidate locations is beyond the scope of our paper because our aim is to present a general methodology. We wanted to make our model valid for any choice of candidate locations. Moreover, the assumption that recharging stations can be built in any nodes is the worst case scenario in terms of problem size because it makes the problem the most difficult. In reality, if candidate locations are more restricted, the problem would become easier to solve.

We set $C_i = 1$ for all $i \in N$, which means that the budget constant B will constrain the total number of recharging stations to be installed on the network.

5.2. Performance analysis

For each test network, we created probabilistic expanded networks for all OD demands. We assumed that the random travel distance followed a normal distribution $N(\hat{R}, \frac{\hat{R}}{5})$, where $\hat{R}$ was a given mean travel range. The proposed approach (Benders-and-Price, BnP) approach involves solving a mixed integer programming problem (BM) and a linear programming problem (BSP-k), for which we used Cplex. In order to solve the problem (BSP) for a given y^*, we needed to solve $|K|$ column generation problems (i.e., (BSP-k) for all $k \in K$). We found that the most time consuming step was to get a set of efficient solutions X_E for (CG2-k) since the number of ODs was relatively large. We implemented a label-correcting algorithm (Skriver and Andersen, 2000) to obtain X_E by using Python and Numba (Lam et al., 2015). Problems (BSP-k) for all $k \in K$ were distributed to 5 independent processes running concurrently using `ipyparallel` library (Bussonnier, 2015). Results of column generation problems were then gathered by using map-reduce type parallelism. We did not employ any load balancing scheme. Therefore, each of all 5 spawned processes should solve at most $\lceil \frac{|K|}{5} \rceil$ (BSP-k) problems. Of course, we can improve the degree of parallelism by increasing the number of processes to spawn. Our computational experiences, however, showed that too many processes were not necessarily beneficial due to communication overhead between different processes.

Fig. 8. The Texas highway network. Filled rectangular and small dot nodes are OD and midarc nodes, respectively. Hollow nodes are intersections. A recharging facility can be deployed at any of the nodes.

Table 2
Computational performance for the Benders-and-Price approach with parallel and sequential solving of subproblems (BSP-k) for the 25NODE network.

B	$\hat{R}$	obj. value	Parallel (5 CPUs)			Sequential (1 CPU)			$\frac{\text{time}_{\text{Sequential}}}{\text{time}_{\text{Parallel}}}$
			Iterations	Columns	Time(s)	Iterations	Columns	Time(s)	
5	4	4930.2	23	12,498	14.3	18	10,495	32.8	2.3
5	6	9289.6	24	43,125	40.8	21	41,157	94.2	2.3
5	8	12109.0	13	124,028	151.2	12	116,388	220.5	1.5
6	4	5795.0	18	13,805	12.9	17	12,164	30.9	2.4
6	6	9868.2	30	74,516	66.2	25	60,013	127.7	1.9
6	8	12802.0	19	209,211	226.0	18	259,927	542.8	2.4
7	4	6375.5	32	19,099	21.3	34	17,609	56.8	2.7
7	6	10418.6	33	80,555	68.2	35	92,430	171.8	2.5
7	8	13387.7	21	276,442	371.2	24	254,596	492.1	1.3

Table 3
Computational performance for the Benders-and-Price approach with parallel and sequential solving of subproblems (BSP-k) for the Texas highway network.

B	$\hat{R}$	obj. value	Parallel (5 CPUs)			Sequential (1 CPU)			$\frac{\text{time}_{\text{Sequential}}}{\text{time}_{\text{Parallel}}}$
			Iterations	Columns	Time(s)	Iterations	Columns	Time(s)	
6	5	11410.8	26	64,995	69.3	30	77,691	204.4	2.9
6	6	12004.5	33	213,212	286.0	30	172,483	408.3	1.4
6	7	12117.1	23	594,443	1291.3	25	558,114	1783.4	1.4
8	5	11667.9	36	104,772	106.8	37	120,434	303.9	2.8
8	6	12229.9	44	251,699	335.0	51	346,384	969.5	2.9
8	7	12362.0	31	565,328	1318.9	31	388,652	1141.2	0.9
10	5	11950.2	40	145,954	192.2	34	125,844	278.8	1.5
10	6	12469.4	43	245,107	266.8	37	358,015	918.3	3.4
10	7	12596.5	40	627,325	1449.0	41	573,472	1940.8	1.3

Tables 2 and 3 show computational performance of our Benders-and-Price approach for each network. The column "iterations" denotes the number of Benders iterations. Total number of generated columns is represented under the header "columns". As mentioned in Section 4.2.1, column generation problems of (CG2-k) can be parallelized. We populated 5 processes to solve the column generation problems (CG2-k) for all $k \in K$ in parallel. For comparison, we also included results obtained without parallelism ("Sequential" columns in Tables 2 and 3). These results clearly indicate that the parallelized column generation greatly helps us to reduce computational times. The number of Benders iterations and the total number of generated columns are different between parallelized and sequential algorithms, even though they give the same optimal objective value. This behavior might be due to the possibility of degeneracy in the dual optimal solutions of (BSP). This results in slightly different columns and different Benders cuts eventually.

As expected, it took longer to solve the problems when the budget constant B and/or the mean travel range $\hat{R}$ becomes larger. However, the number of Benders iterations does not strongly correlate with computational time. Rather, the number of generated columns has a greater impact on the computational time. We also note that the time spent solving the Benders master problem is rather small compared to the total computational time. Most of the computational time is spent solving the column generation subproblems (CG2-k). Since the parallelized algorithm clearly outperforms the sequential version, all experiments presented below were performed using the parallelized algorithm.

We compared our Benders-and-Price approach with an MINLP solver which solves the formulation (P) directly by using Couenne (Belotti et al., 2009). Couenne is a branch-and-bound solver that aims at finding global optima of nonconvex MINLPs. Although Couenne implements a sophisticated branch-and-bound method with many additional enhancements, it was found that Couenne could not solve any of the problems solved by the BnP approach (Tables 2 and 3). We had to reduce the number of OD pairs to make the problems solvable with Couenne. Test problems were created to have smaller numbers of ODs by taking some of the largest demands only. The number of demands ranged from 20 to 40, which was less than approximately 10% that of the original. We set the time limit for Couenne to be 1 h. Results are shown in Tables 4 and 5. As shown in these tables, BnP was significantly faster than solving the formulation (P) directly. Couenne could not solve most of the problems with $|K| > 20$ for the 25NODE network. When it could, BnP was faster by several orders of magnitude. Also, solution times for Couenne were inconsistent. It could solve problems that were supposedly harder to solve (e.g., $|K| = 30, B = 6, \hat{R} = 6$ for the Texas Highway network). However, it failed to solve supposedly easier problems (e.g., $|K| = 30, B = 6, \hat{R} = 4$ for the Texas Highway network). On the other hand, BnP showed relatively consistent results for problems with various difficulties.

5.3. Comparison with banch-and-Price approach

The proposed algorithm utilized column generation in the context of the Benders decomposition. Column generation is often incorporated into the branch-and-bound framework, resulting in the so-called branch-and-price approach. We note that it might be possible to use formulation (BSP) with constraint (9) and decision variable y for the column generation master problem and the branch-and-price algorithm. We implemented the branch-and-price algorithm for comparison with our Benders-and-Price approach, since it is not difficult to develop the branch-and-price algorithm by exploiting the probabilistic expanded network proposed in this paper. The sketch of the branch-and-price approach is as follows. The linear relaxation of formulation (BSP) with constraint (9) and decision variable y becomes column generation master problem. The dual optimal solution of the column generation master problem is then fed to the column generation subproblem presented in Section 4.2.1. If there is no more column to generate and the optimal solution of the column generation master problem is fractional, we can create two nodes in the branch-and-bound tree to be evaluated. Due to Proposition 2, it is easy to see that branching on some fractional valued y is enough to warrant an optimal integer solution. Note that, for the branch-and-price, we used the same column generation code developed for the Benders-and-Price because the goal was to compare the performance between Benders-and-Price and branch-and-price, and to show which approach might be more appropriate for solving the problem.

Table 4
Comparison of computational times for the proposed approach (BnP) and Couenne for the 25NODE network.

$\|K\|$	B	$\hat{R}$	obj. value	time(sec.)		$\frac{\text{Couenne}}{\text{BnP}}$
				BnP	Couenne	
20	5	4	4442.6	0.7	41.2	61.5
20	5	6	7468.4	0.5	69.7	128.7
20	5	8	8375.9	0.6	190.7	338.4
20	6	4	4973.9	0.7	59.6	91.0
20	6	6	7738.6	0.9	56.6	60.4
20	6	8	8596.3	0.6	97.8	170.0
20	7	4	5489.5	1.3	70.0	52.0
20	7	6	8066.0	1.1	188.2	164.9
20	7	8	8599.3	1.4	117.2	83.5
30	5	4	4661.1	2.0	217.6	108.4
30	5	6	8444.5	1.6	Time limit	> 2187.9
30	5	8	9768.6	2.1	Time limit	> 1683.9
30	6	4	5424.4	2.2	261.9	117.0
30	6	6	8956.1	2.1	Time limit	> 1685.6
30	6	8	9989.0	2.4	Time limit	> 1531.4
30	7	4	5930.1	4.1	378.1	92.7
30	7	6	9445.0	2.0	Time limit	> 1828.5
30	7	8	10203.6	3.2	Time limit	> 1113.7
40	5	4	4661.1	2.0	489.8	242.3
40	5	6	8444.5	1.5	2,646.1	1799.2
40	5	8	10379.2	1.7	Time limit	> 2100.4
40	6	4	5424.4	2.2	640.4	296.0
40	6	6	8956.1	1.9	Time limit	> 1849
40	6	8	10721.0	2.2	Time limit	> 1642.5
40	7	4	5930.1	4.2	807.3	190.9
40	7	6	9445.0	1.8	Time limit	> 2021.3
40	7	8	10992.7	2.8	Time limit	> 1294.4

Table 5
Comparison of computational times for the proposed approach (BnP) and Couenne for the Texas highway network.

$\|K\|$	B	$\hat{R}$	obj. value	time(sec.)		$\frac{\text{Couenne}}{\text{BnP}}$
				BnP	Couenne	
30	6	4	10830.6	2.5	Time limit	> 1424.7
30	6	6	11110.1	1.8	207.3	114.5
30	6	8	11273.0	1.4	68.5	48.6
30	7	4	10938.8	1.0	Time limit	> 3595.3
30	7	6	11187.8	1.6	205.9	132.2
30	7	8	11279.8	5.2	77.0	14.7
30	8	4	10969.9	1.7	Time limit	> 2066.4
30	8	6	11265.3	1.2	128.6	111.2
30	8	8	11357.9	2.2	83.5	37.9
40	6	4	11227.4	1.3	Time limit	> 2694.7
40	6	6	11519.0	1.4	645.1	464.4
40	6	8	11684.3	3.6	936.5	261.6
40	7	4	11264.9	2.1	Time limit	> 1722.2
40	7	6	11532.2	5.2	Time limit	> 686.9
40	7	8	11745.3	3.9	1720.8	444.7
40	8	4	11342.0	2.0	Time limit	> 1786.5
40	8	6	11629.7	2.9	2456.2	835.3
40	8	8	11802.6	9.1	2786.2	305.6

Results of performance comparison between the proposed algorithm (Benders-and-Price) and the branch-and-price algorithm are shown in Tables 6 and 7. For the both approaches, we used parallel implementation for column generation. The header “master(s,%)” represents the time (in seconds and percentage of total time) spent in solving Benders master problem and column generation master problem, respectively. The header “bnb” denotes the number of branch-and-bound nodes solved during the branch-and-price algorithm. These results clearly show that the proposed Benders-and-Price approach is significantly faster than the branch-and-price approach, up to 10 times in average. Given that the master problem for the Benders-and-Price approach is an integer programming problem while the column generation master problem for the branch-and-price approach is an LP, it is interesting to see so much longer time is needed to solve the master problem with the branch-and-price approach. This might be due to the huge-sized column generation master problem for the

Table 6
Comparison of computational times for the proposed approach (BnP) and branch-and-price for the 25NODE network.

B	$\hat{R}$	obj. value	Proposed (Benders-and-Price)				Branch-and-price				Time ratio
			iter.	Columns	Master(s, %)	Time(s)	bnb	Columns	Master(s, %)	Time(s)	
5	4	4930.2	23	12,498	2.1 (14.9%)	14.3	83	62,119	54.0 (19.9%)	270.7	18.9
5	6	9289.6	24	43,125	1.7 (4.1%)	40.8	9	95,620	529.7 (83.2%)	636.4	15.6
5	8	12109.0	13	124,028	0.9 (0.6%)	151.2	1	124,539	711.0 (87.4%)	813.4	5.4
6	4	5795.0	18	13,805	1.4 (11.0%)	12.9	1	36,419	5.8 (14.9%)	38.7	3.0
6	6	9868.2	30	74,516	3.9 (5.9%)	66.2	1	95,549	1446.4 (95.0%)	1521.7	23.0
6	8	12802.0	19	209,211	1.8 (0.8%)	226.0	1	121,269	1127.3 (91.5%)	1232.6	5.5
7	4	6375.5	32	19,099	4.0 (18.8%)	21.3	9	38,039	22.2 (25.4%)	87.5	4.1
7	6	10418.6	33	80,555	5.6 (8.1%)	68.2	7	93,094	597.3 (85.9%)	695.5	10.2
7	8	13387.7	21	276,442	2.0 (0.5%)	371.2	5	147,220	1461.8 (89.9%)	1625.3	4.4
											avg.: 10.0

Table 7
Comparison of computational times for the proposed approach (BnP) and branch-and-price for the Texas highway network.

B	$\hat{R}$	obj. value	Proposed (Benders-and-Price)				Branch-and-price				Time ratio
			iter.	Columns	Master(s, %)	Time(s)	bnb	Columns	Master(s, %)	Time(s)	
6	5	11410.8	26	64,995	5.0 (7.2%)	69.3	3	256,475	98.0 (14.7%)	664.6	9.6
6	6	12004.5	33	213,212	7.6 (2.7%)	286.0	1	419,480	334.6 (28.3%)	1180.5	4.1
6	7	12117.1	23	594,443	4.6 (0.4%)	1291.3	1	491,921	1183.4 (54.3%)	2180.8	1.7
8	5	11667.9	36	104,772	12.4 (11.6%)	106.8	71	530,993	2346.9 (53.9%)	4353.7	40.8
8	6	12229.9	44	251,699	19.0 (5.7%)	335.0	3	472,288	522.3 (33.6%)	1553.5	4.6
8	7	12362.0	31	565,328	7.8 (0.6%)	1318.9	7	642,601	3048.8 (65.7%)	4639.9	3.5
10	5	11950.2	40	145,954	23.4 (12.2%)	192.2	3	233,311	129.6 (19.4%)	667.7	3.5
10	6	12469.4	43	245,107	34.4 (12.9%)	266.8	1	501,108	1486.7 (58.6%)	2537.6	9.5
10	7	12596.5	40	627,325	28.4 (2.0%)	1449.0	1	518,498	1742.6 (63.0%)	2767.5	1.9
											avg.: 8.8

branch-and-price. Since y is a decision variable in the column generation master problem, it is impossible to decompose the problem with $k \in K$, thus yielding a huge optimization problem. The size of master problem is getting bigger as column generation and searching the branch-and-bound tree proceed. This makes it even more difficult to solve the problem. On the other hand, in the Benders-and-Price approach, the formulation (BSP) is separable with $k \in K$ because y^* is no longer a decision variable, but a fixed value. This implies that the column generation master problem for Benders cut generation problem is much smaller in terms of size. Therefore, it is significantly easier to solve.

Another aspect to consider is the numbers of columns generated. Our results indicate that the Benders-and-Price approach requires considerably less columns to obtain optimal solutions. Note that the column generation for the Benders-and-Price approach always uses a fixed and integral y^*, making dual optimal solution be more sparse. On the other hand, the column generation master problem of the branch-and-price approach tends to produce more dense dual optimal solutions because there can be many non-zero and fractional y values. In column generation literature (Lee and Park, 2011; Lübbecke and Desrosiers, 2005), it is well-known that dense dual optimal solutions may make solving column generation deteriorated, especially for labeling algorithms (Briant et al., 2008).

As a summary, the superior performance of the Benders-and-Price approach over the branch-and-price approach mainly comes from three aspects: (1) The Benders master problem is easy to solve; (2) Column generation master problems in Benders cut generation are separable and easy to solve; and (3) The labeling algorithm for the bi-objective shortest path problem performs better due to the sparse dual optimal solutions.

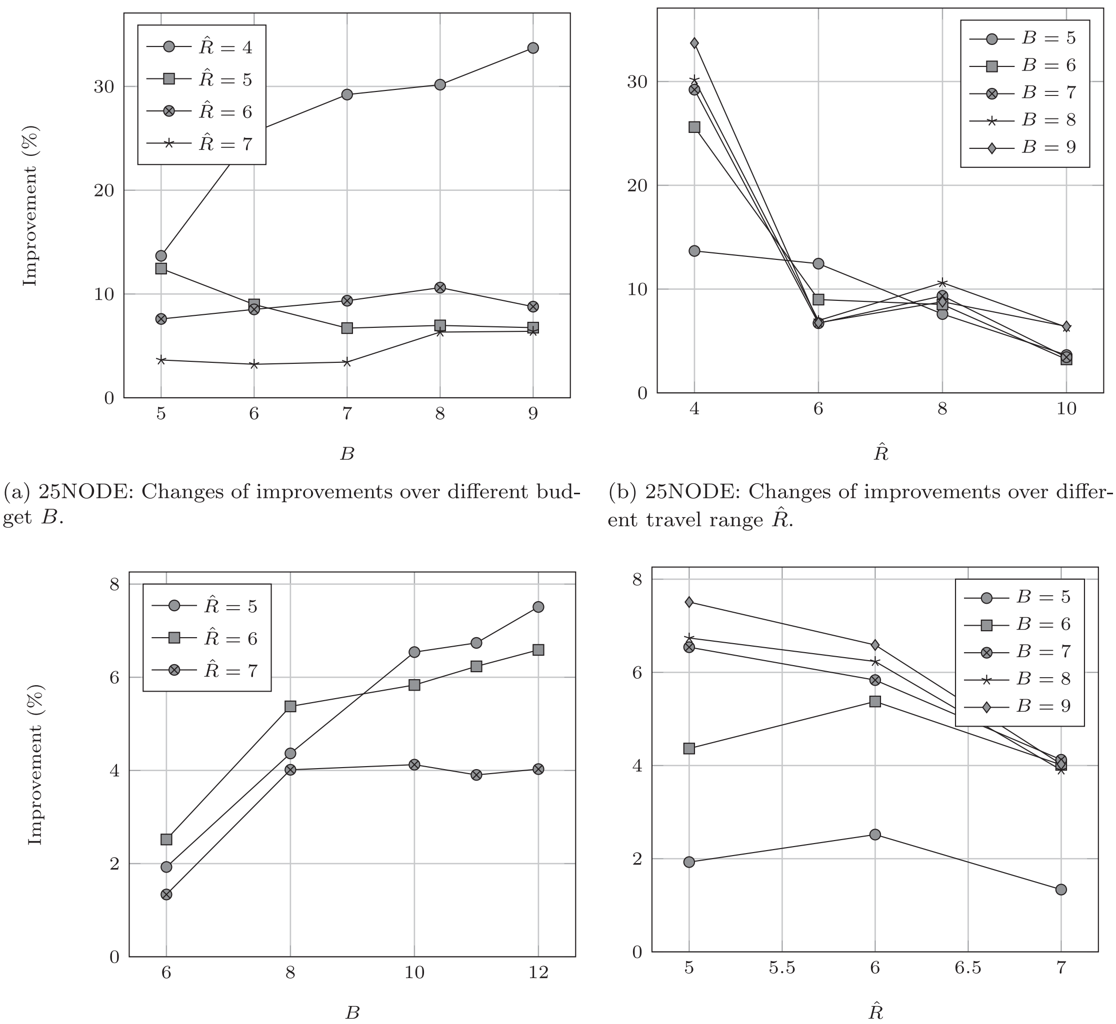
5.4. Benefits of a probabilistic approach

Without a probabilistic consideration, the problem FRLMwU is reduced to a deterministic FRLM problem that can be given as the following:

$$\text{(FRLM)} \quad \max \quad \sum_{k \in K} D_k z_k \tag{49}$$

$$\text{s.t.} \quad \sum_{(i,j) \in A_k} x_{ij}^k - \sum_{(j,i) \in A_k} x_{ji}^k = \begin{cases} z_k, & \text{if } i = s_k, \\ -z_k, & \text{if } i = t_k, \\ 0, & \text{otherwise} \end{cases}, \quad \forall i \in N_k, k \in K, \tag{50}$$

$$\sum_{(j,i) \in A_k} x_{ji}^k \le y_i, \quad \forall i \in N, k \in K, \tag{51}$$



(a) 25NODE: Changes of improvements over different budget B.

(b) 25NODE: Changes of improvements over different travel range $\hat{R}$.

(c) Texas: Changes of improvements over different budget B.

(d) Texas: Changes of improvements over different travel range $\hat{R}$.

Fig. 9. Improvements in the objective values from those of the deterministic approach to those of the proposed approach.

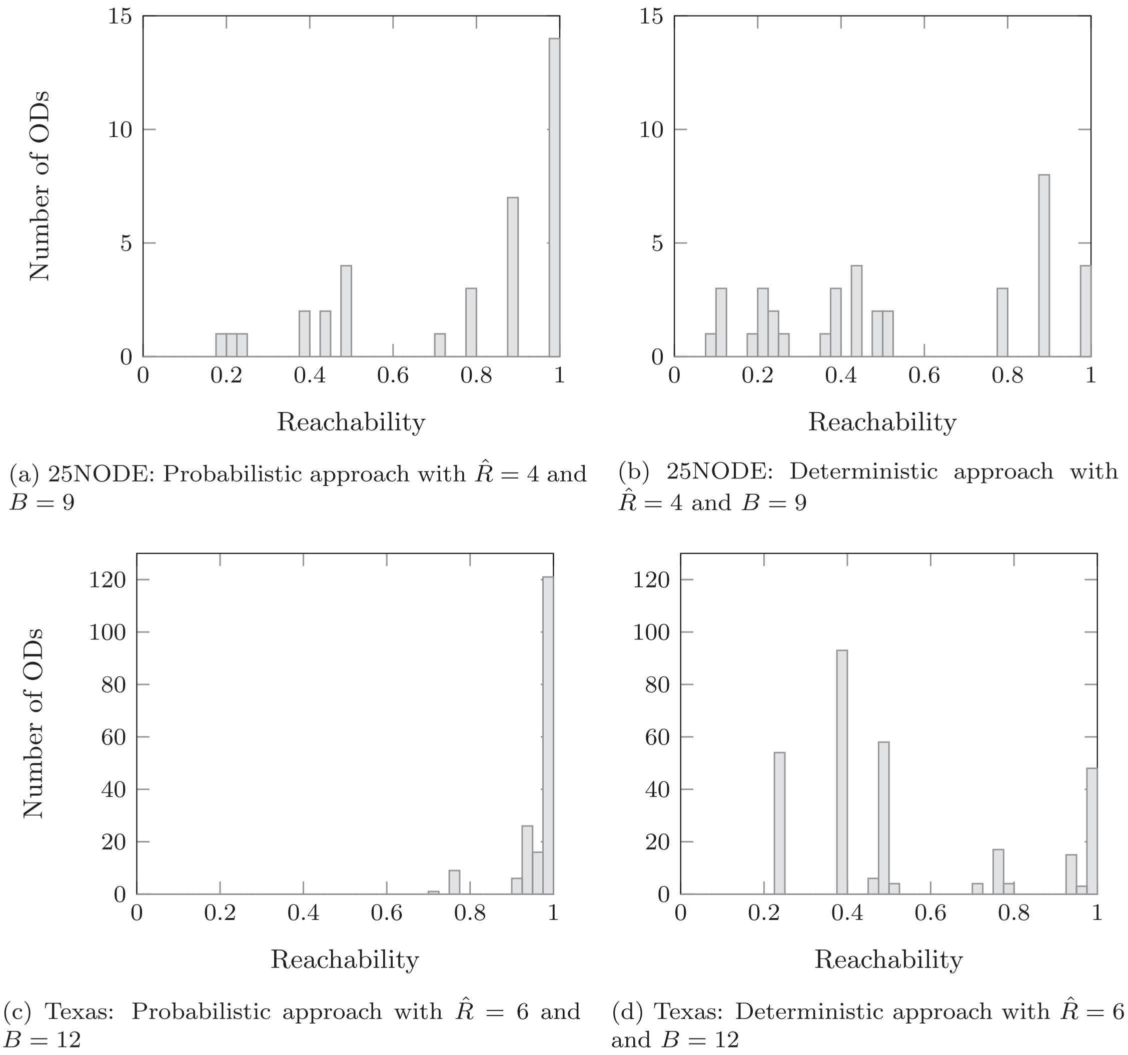
$$\sum_{i \in N} C_i y_i \le B, \tag{52}$$

$$x_{ij}^k \in \{0, 1\}, \quad \forall (i, j) \in A_k, k \in K, \tag{53}$$

$$y_i \in \{0, 1\}, \quad \forall i \in N, \tag{54}$$

$$z_k \in \{0, 1\}, \quad \forall k \in K. \tag{55}$$

The formulation (FRLM) is an integer programming problem that can be solved very efficiently using Cplex. We resolved all problems shown in Tables 2 and 3 by using the formulation (FRLM). We then evaluated reachability of the recharging pattern for each OD demand by using the same reachability function as in FRLMwU. We also calculated the value of the objective function (4) by using the optimal solution obtained from (FRLM). The purpose of this posterior evaluation is to assess the benefits of the probabilistic approach. Fig. 9 plots the advantage of the probabilistic approach over the deterministic method. The y-axis represents the extent to which the optimal objective values of the BnP are better than those of (FRLM)



(a) 25NODE: Probabilistic approach with $\hat{R} = 4$ and $B = 9$

(b) 25NODE: Deterministic approach with $\hat{R} = 4$ and $B = 9$

(c) Texas: Probabilistic approach with $\hat{R} = 6$ and $B = 12$

(d) Texas: Deterministic approach with $\hat{R} = 6$ and $B = 12$

Fig. 10. Histogram of reachability of covered ODs.

based on posterior evaluation, i.e., improvement $:= \frac{Z_{\text{BnP}} - Z_{\text{FRLM}}}{Z_{\text{BnP}}} \times 100$, where Z_{BnP} and Z_{FRLM} are the optimal objective value of BnP and the value of the posterior evaluation of (FRLM), respectively.

This figure shows that the proposed probabilistic approach has a clear advantage over the deterministic approach based on probabilistic assumption of the maximum travel range. Its benefits then become greater as the budget constant B increases. Longer travel distance makes results of the two approaches become closer. We observed much smaller improvement for the Texas highway network with the proposed approach. This might be due to the fact that many big cities in the Texas highway network are located close to each other, e.g., nodes 1–8 in Fig. 8. Demand between these clustered cities does not require recharging at all. This makes the two approaches produce the same expected value of the demand. On the contrary, OD nodes in the 25NODE network are not spatially clustered. Therefore, the traffic demand would require frequent recharging. The probabilistic method can lead to evenly distributed recharging locations in the paths of the demand, yielding more reachable routes.

Fig. 10 contrasts the reachability of each OD pair between the probabilistic approach and the deterministic approach. It is desired for ODs to have higher reachability values close to 1. Histograms show that solutions obtained using our probabilistic approach (Fig. 10a and c) are significantly better in terms of reachability than those obtained from the deterministic method (Fig. 10b and d) since there are more ODs with higher reachability values.

6. Conclusion

In this paper, we proposed a novel optimization approach to determine optimal locations for electric vehicle charging stations. Previous studies considered that the travel range of electric vehicles was deterministic. In reality, the maximum travel distance can vary according to environmental factors including traffic conditions, weather, temperature, and so on. Our approach is based on a probabilistic expanded network. It is developed by incorporating a probabilistic reachability function into the existing expanded network. By using the probabilistic expanded network, we first developed a mixed-integer nonlinear programming problem that took too much time to solve it using state-of-the-art solvers. To overcome this difficulty, we presented a Benders-and-Price approach by combining Benders decomposition and column generation methods. The resulting column generation problem turned out to be a nonlinear integer programming problem. We showed that the column generation problem could be solved using an efficient combinatorial algorithm inspired by the bi-objective shortest path problem.

Computational experiments were conducted on the benchmark network (25NODE) extensively used in the literature and a real-world road network (Texas highway). Our results clearly showed that the proposed Benders-and-Price approach outperformed the state-of-the-art MINLP solver and the branch-and-price approach. Beside its advantages regarding computational time, the proposed approach produced better solutions under the assumption of an uncertain travel range.

The problem considered here did not take the capacity of recharging stations into account. One major drawback of electric vehicles is that recharging takes too much time while fast-charging facilities are too expensive to deploy. To reduce the waiting time of drivers, the capacity of the recharging stations should be carefully designed to meet traffic demand for a particular location. The proposed Benders-and-Price approach might be applicable to this problem by penalizing waiting time in the Benders subproblem.

Another line of research is to consider the route choice behavior of EV drivers. Given the uncertain travel range, locations of charging stations can have significant influence on EV driver's decision on which route to take (He et al., 2015). For example, Yang et al. (2016) showed that EV drivers tend to choose the routes with charging stations being closer to origin and consistent with travel direction. To capture such an interdependency between the locations of charging stations and route choice behavior of drivers when planning charging infrastructure, the charging station location problem should be modeled as a two player game between the infrastructure planner and the EV drivers. Usually, this game is formulated as a following bi-level programming problem:

$$\min_x \quad F(x, y) \tag{56}$$

$$\text{s.t.} \quad G(x, y) \leq 0 \tag{57}$$

$$y \in \operatorname{argmin}_z \{f(z) : g(x, z) \leq 0\}. \tag{58}$$

The upper level problem determines the locations of charging stations, which are represented by decision variable x, and the lower level problem describes the route choice behavior of EV drivers, which is represented by decision variable y. In principle, the bi-level programming problem is known to be very difficult because of its nonconvexity nature. Also, the difficulty of the problem may increase due to the fact that the lower level problem should be characterized by the user equilibrium condition (Gao et al., 2005). Note also that, to address the route choice behavior in the lower level problem, it is necesssary to consider alternate routes for a given OD pair. The current study assumed the use of the shortest path. Although this assumption can be relaxed by introducing alternative routes, the complexity of the problem may increase significantly. Therefore, developing a concrete mathematical formulation of the bi-level problem and an efficient solution algorithm for solving it merits further investigation.

Acknowledgments

This research was supported by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Science, ICT & Future Planning (NRF-2015R1C1A1A01054606). Chungmok Lee was supported by Hankuk University of Foreign Studies Research Fund of 2017.

References

Adler, J.D., Mirchandani, P.B., 2014. Online routing and battery reservations for electric vehicles with swappable batteries. Transp. Res. Part B 70, 285–302.
An, Y., Zeng, B., Zhang, Y., Zhao, L., 2014. Reliable p-median facility location problem: two-stage robust models and algorithms. Transp. Res. Part B 64, 54–72.
Arslan, O., Karaşan, O.E., 2016. A benders decomposition approach for the charging station location problem with plug-in hybrid electric vehicles. Transp. Res. Part B 93, 670–695.
Belotti, P., Lee, J., Liberti, L., Margot, F., Waechter, A., 2009. Branching and bounds tightening techniques for non-convex MINLP. Optim. Methods Softw. 24 (4–5), 597–634.
Briant, O., Lemaréchal, C., Meurdesoif, P., Michel, S., Perrot, N., Vanderbeck, F., 2008. Comparison of bundle and classical column generation. Math. Program. 113 (2), 299–344.
Bussonnier, M., 2015. Interactive parallel computing in python. https://github.com/ipython/ipyparallel.
Campbell, J.F., 1996. Hub location and the p-hub median problem. Oper. Res. 44 (6), 923–935.
Capar, I., Kuby, M., 2011. An efficient formulation of the flow refueling location model for alternative-fuel stations. IIE Trans. 44 (8), 622–636.
Capar, I., Kuby, M., Leon, V.J., Tsai, Y.-J., 2013. An arc cover–path-cover formulation and strategic analysis of alternative-fuel station locations. Eur. J. Oper. Res. 227 (1), 142–151.
Chen, Z., He, F., Yin, Y., 2016. Optimal deployment of charging lanes for electric vehicles in transportation networks. Transp. Res. Part B 91, 344–365.

Chung, S.H., Kwon, C., 2015. Multi-period planning for electric car charging station locations: a case of korean expressways. Eur. J. Oper. Res. 242 (2), 677–687.
Church, R., ReVelle, C., 1974. The maximal covering location problem. In: Papers of the Regional Science Association, vol. 32. Springer, pp. 101–118.
Drezner, Z., Hamacher, H.W., 1995. Facility Location. Springer-Verlag, New York, NY.
Egbue, O., Long, S., 2012. Barriers to widespread adoption of electric vehicles: an analysis of consumer attitudes and perceptions. Energy Policy 48, 717–729.
Faridimehr, S., Venkatachalam, S., Chinnam, R.B., 2017. A stochastic programming approach for electric vehicle charging network design. arXiv preprint arXiv:1701.06723.
Frade, I., Ribeiro, A., Gonçalves, G., Antunes, A., 2011. Optimal location of charging stations for electric vehicles in a neighborhood in lisbon, portugal. Transp. Res. Rec. 2252, 91–98.
Franke, T., Neumann, I., Bühler, F., Cocron, P., Krems, J.F., 2012. Experiencing range in an electric vehicle: understanding psychological barriers. Appl. Psychol. 61 (3), 368–391.
Gao, Z., Wu, J., Sun, H., 2005. Solution algorithm for the bi-level discrete network design problem. Transp. Res. Part B 39 (6), 479–495.
He, F., Yin, Y., Zhou, J., 2015. Deploying public charging stations for electric vehicles on urban road networks. Transp. Res. Part C 60, 227–240.
Hodgson, M.J., 1990. A flow-capturing location-allocation model. Geograph. Anal. 22 (3), 270–279.
Hof, J., Schneider, M., Goeke, D., 2017. Solving the battery swap station location-routing problem with capacitated electric vehicles using an AVNS algorithm for vehicle-routing problems with intermediate stops. Transp. Res. Part B 97, 102–112.
International Energy Agency, 2016. Global EV Outlook 2016. Technical Report.
Kim, J.-G., Kuby, M., 2012. The deviation-flow refueling location model for optimizing a network of refueling stations. Int. J. Hydrogen Energy 37 (6), 5406–5420.
Kuby, M., Lim, S., 2005. The flow-refueling location problem for alternative-fuel vehicles. Socioecon. Plann. Sci. 39 (2), 125–145.
Kuby, M., Lim, S., 2007. Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs. Netw. Spat. Econ. 7 (2), 129–152.
Kuby, M., Lines, L., Schultz, R., Xie, Z., Kim, J.-G., Lim, S., 2009. Optimization of hydrogen stations in florida using the flow-refueling location model. Int. J. Hydrogen Energy 34 (15), 6045–6064.
Lam, S.K., Pitrou, A., Seibert, S., 2015. Numba: a LLVM-based python JIT compiler. In: Proceedings of the Second Workshop on the LLVM Compiler Infrastructure in HPC.
Lee, C., Park, S., 2011. Chebyshev center based column generation. Discret. Appl. Math. 159 (18), 2251–2265.
Lee, Y.-G., Kim, H.-S., Kho, S.-Y., Lee, C., 2015. User equilibrium-based location model of rapid charging stations for electric vehicles with batteries that have different states of charge. Transp. Res. Rec. 2454, 97–106.
Liu, H., Wang, D.Z., 2017. Locating multiple types of charging facilities for battery electric vehicles. Transp. Res. Part B 103, 30–55.
Lu, G., Zhou, X., Peng, Q., He, B., Mahmoudi, M., Zhao, J., 2016. Solving resource recharging station location-routing problem through a resource-space-time network representation. arXiv preprint arXiv:1602.06889.
Lübbecke, M.E., Desrosiers, J., 2005. Selected topics in column generation. Oper. Res. 53 (6), 1007–1023.
Magnanti, T.L., Wong, R.T., 1981. Accelerating benders decomposition: algorithmic enhancement and model selection criteria. Oper. Res. 29 (3), 464–484.
Mak, H.-Y., Rong, Y., Shen, Z.-J.M., 2013. Infrastructure planning for electric vehicles with battery swapping. Manage. Sci. 59 (7), 1557–1575.
Martins, E.Q.V., 1984. On a multicriteria shortest path problem. Eur. J. Oper. Res. 16 (2), 236–245.
Miralinaghi, M., Lou, Y., Keskin, B.B., Zarrinmehr, A., Shabanpour, R., 2017. Refueling station location problem with traffic deviation considering route choice and demand uncertainty. Int. J. Hydrogen Energy 42 (5), 3335–3351.
MirHassani, S., Ebrazi, R., 2012. A flexible reformulation of the refueling station location problem. Transp. Sci. 47 (4), 617–628.
Nagy, G., Salhi, S., 2007. Location-routing: issues, models and methods. Eur. J. Oper. Res. 177 (2), 649–672.
Neubauer, J., Wood, E., Pesaran, A., 2013. Project Milestone. Analysis of Range Extension Techniques for Battery Electric Vehicles. Technical Report. National Renewable Energy Laboratory (NREL), Golden, CO (United States).
Nie, Y.M., Ghamami, M., Zockaie, A., Xiao, F., 2016. Optimization of incentive polices for plug-in electric vehicles. Transp. Res. Part B 84, 103–123.
Raith, A., Ehrgott, M., 2009. A comparison of solution strategies for biobjective shortest path problems. Comput. Oper. Res. 36 (4), 1299–1331.
Sedeño-Noda, A., Raith, A., 2015. A dijkstra-like method computing all extreme supported non-dominated solutions of the biobjective shortest path problem. Comput. Oper. Res. 57 (C), 83–94.
Simchi-Levi, D., Berman, O., 1988. A heuristic algorithm for the traveling salesman location problem on networks. Oper. Res. 36 (3), 478–484.
Skriver, A.J., Andersen, K.A., 2000. A label correcting approach for solving bicriterion shortest-path problems. Comput. Oper. Res. 27 (6), 507–524.
Straubel, J., 2015. Driving range for the model S family.
Toregas, C., Swain, R., ReVelle, C., Bergman, L., 1971. The location of emergency service facilities. Oper. Res. 19 (6), 1363–1373.
Upchurch, C., Kuby, M., Lim, S., 2009. A model for location of capacitated alternative-fuel stations. Geograph. Anal. 41 (1), 85–106.
de Vries, H., Duijzer, E., 2017. Incorporating driving range variability in network design for refueling facilities. Omega 69, 102–114.
Wang, Y.-W., 2007. An optimal location choice model for recreation-oriented scooter recharge stations. Transp. Res. Part D 12 (3), 231–237.
Wang, Y.-W., 2008. Locating battery exchange stations to serve tourism transport: a note. Transp. Res. Part D 13 (3), 193–197.
Wang, Y.-W., Lin, C.-C., 2009. Locating road-vehicle refueling stations. Transp. Res. Part E 45 (5), 821–829.
Wang, Y.-W., Lin, C.-C., 2013. Locating multiple types of recharging stations for battery-powered electric vehicle transport. Transp. Res. Part E 58, 76–87.
Worley, O., Klabjan, D., Sweda, T.M., 2012. Simultaneous vehicle routing and charging station siting for commercial electric vehicles. In: Electric Vehicle Conference (IEVC), 2012 IEEE International. IEEE, pp. 1–3.
Xie, C., Wang, T.-G., Pu, X., Karoonsoontawong, A., 2017. Path-constrained traffic assignment: modeling and computing network impacts of stochastic range anxiety. Transp. Res. Part B 103, 136–157.
Yang, Y., Yao, E., Yang, Z., Zhang, R., 2016. Modeling the charging and route choice behavior of BEV drivers. Transp. Res. Part B 65, 190–204.
Yıldız, B., Arslan, O., Karaşan, O.E., 2016. A branch and price approach for routing and refueling station location model. Eur. J. Oper. Res. 248 (3), 815–826.
Zhang, A., Kang, J.E., Kwon, C., 2017. Incorporating demand dynamics in multi-period capacitated fast-charging location planning for electric vehicles. Transp. Res. Part B 103, 5–29. doi:10.1016/j.trb.2017.04.016.
Zheng, H., Peeta, S., 2017. Routing and charging locations for electric vehicles for intercity trips. Transport. Plan. Techn. 40 (4), 393–419.