UCI-ITS-WP-13-2

Stochastic dynamic itinerary interception refueling location problem with queue delay for electric taxi charging stations

UCI-ITS-WP-13-2

Jaeyoung Jung, Ph.D.
Joseph Y.J. Chow, Ph.D.
R. Jayakrishnan, Ph.D.
Ji Young Park, Ph.D.

Institute of Transportation Studies
University of California, Irvine; Irvine, CA 92697-3600, U.S.A.

jaeyounj@uci.edu, joseph.chow@ryerson.ca, rjayakri@uci.edu, parkjy@koti.re.kr

July, 2013

Institute of Transportation Studies
University of California, Irvine
Irvine, CA 92697-3600, U.S.A.
http://www.its.uci.edu

Stochastic dynamic itinerary interception refueling location with queue delay for electric taxi charging stations

Jaeyoung Jung[1*], Joseph Y.J. Chow[2], R. Jayakrishnan[1], and Ji Young Park[3]

[1]*Institute of Transportation Studies, Department of Civil and Environmental Engineering, University of California, Irvine, California, United States*
[2]*Department of Civil Engineering, Ryerson University, Toronto, Ontario, Canada*
[3]*Department of National Transport Strategy Planning, The Korea Transport Institute, Gyeonggi-do, Korea*

ABSTRACT

A new facility location model and a solution algorithm are proposed that feature 1) itinerary-interception instead of flow-interception; 2) stochastic demand as dynamic service requests; and 3) queueing delay. These features are essential to analyze battery-powered electric shared-ride taxis operating in a connected, centralized dispatch manner. The model and solution method are based on a bi-level, simulation-optimization framework that combines an upper level multiple-server allocation model with queueing delay and a lower level dispatch simulation based on earlier work by Jung and Jayakrishnan. The solution algorithm is tested on a fleet of 600 shared-taxis in Seoul, Korea, spanning 603 km^2, a budget of 100 charging stations, and up to 22 candidate charging locations, against a benchmark "naïve" genetic algorithm that does not consider cyclic interactions between the taxi charging demand and the charger allocations with queue delay. Results show not only that the proposed model is capable of locating charging stations with stochastic dynamic itinerary-interception and queue delay, butt that the bi-level solution method improves upon the benchmark algorithm in terms of realized queue delay, total time of operation of taxi service, and service request rejections. Furthermore, we show how much additional benefit in level of service is possible in the upper-bound scenario when the number of charging stations approaches infinity.

KEYWORDS: Electric Vehicle, Shared-Taxi, EV charging, Bi-Level Optimization, Facility Location, Stochastic Demand, Simulation, Refueling

1. INTRODUCTION

Carbon-based emissions and greenhouse gases (GHG) are critical issues that policy-makers have sought to address in a generally global effort since the Kyoto Protocol in 1998. The transportation sector is a major culprit, contributing approximately 30 percent of the GHG emissions in the United States (U.S. EPA, 2006). One technology-oriented solution is to switch vehicle fuels to alternative, sustainable sources and wean consumers off their oil dependency. To that end, automobile manufacturers produced a range of different types of alternative-fuel vehicles (AFVs) including hybrid vehicles (HVs), plug-in hybrid electric vehicles (PHEVs), electric vehicles (EVs), and fuel cell vehicles (FCVs). It has been shown that EVs can cost as little as 1.2 to 1.9 cents per km (2 to 3 cents per mile) compared to 8 cents per km (13

[*] Corresponding Author. Tel.:+1 949-824-5989
E-mail address: jaeyounj@uci.edu

cents per mile) for vehicles powered by Internal Combustion Engines (ICE), and EVs cause 50 percent less CO2 emissions per mile traveled compared to ICE (Mak *et al.*, 2012).

Nevertheless, there are several obstacles to EV adoption; the most important being the chicken-and-egg problem related to the electric charging infrastructure. The lack of penetration in the EV market (and other AFVs as well) limits the value of fuel infrastructure investments, which increases the "range anxiety" and reduces consumer demand. As a result, much of the recent research in AFVs has placed optimal location and allocation of refueling infrastructure at the forefront (e.g. Kuby and Lim, 2005; Lin *et al.*, 2008; Wang and Lin, 2009; Nourbakhsh and Ouyang, 2010; Kameda and Mukai, 2011; Kang and Recker, 2012; MirHassani and Ebrazi, 2012; Chung and Kwon, 2012, Jiang *et al.*, 2012; He *et al.*, 2013; Xi *et al.*, 2013). This problem is particularly relevant to EVs because the vehicles tend to have limited ranges, and recharging can take on the order of an hour or more.

One implementation strategy is to invest in EV flexible transit services (FTS) (Koffman, 2004) and taxi fleets. FTS encapsulate a wide range of transit and taxi services that include demand responsive transit (DRT), flexi-route transit, car-sharing systems, and ride-sharing. There are several advantages to implementing EV FTS for initial EV adoption: 1) it takes the individual consumer behavior out of the picture since the decision would be based on fleet managers and/or public agencies; 2) "range anxiety" is less of an issue because transit fleets operate in the same urban region; 3) the reduction in energy consumption would be most realized by transit vehicles operating in a congestion-heavy region (Tong *et al.*, 2000) as opposed to long distance rural trips; and 4) specific to FTS, the vehicles can be passenger cars so that the same infrastructure can also serve the consumer market. Even if EVs are not considered ready to replace private ICE vehicles for the general public, they may be viable options for these fleet applications (Barth and Todd, 2001; Better Place, 2012; FedEx, 2012). The Seoul Metropolitan Government in Korea unveiled its "2014 Master Plan for Electric Vehicle" in 2011, which will commercialize 30,000 units of EVs on the road by 2014 with buses and taxicabs (eGlobal Travel Media, 2011). One of the key goals in the plan is to deploy 1,000 EV taxicabs by 2014. The EV taxis will be able to operate over a distance of 200 km to 400 km a day and will require battery charging stations or battery replacing facilities.

The proliferation of computing and mobile technologies for connected vehicles has made it easier to operate FTS (Lau *et al.*, 2011), resulting in a growing literature in that area: Cortés and Jayakrishnan (2002), Zhao and Dessouky (2008), Hadas and Ceder (2008), Quadrifoglio and Li (2009), Mulley and Nelson (2009), Crainic *et al.* (2010), Alshalalfah and Shalaby (2012), Nourbakhsh and Ouyang (2012), Kim and Schonfeld (2012), Jung (2012), Jung and Jayakrishnan (2012), and Jung *et al.* (2013). Surprisingly, there has not been much research in the facility location problem specifically for EV taxi and FTS fleets.

Facility location problems for EV transit and taxi fleets have a set of unique characteristics that cannot be addressed by the solutions developed thus far in the literature. First, EV taxis, as service vehicles, need to place much more weight on the cost of the recharging time. Whereas consumers can recharge when they return home for the night, a taxi may need to be out of commission for a significant period of time during the day if it runs out of charge. With limited infrastructure, queue delay is an important problem that has been largely neglected in the AFV facility location literature to date. Second, many of the AFV location problems focused on fuel stations covering paths formed by origin-destination (OD) pairs in a network. However, connected taxis operate with a central dispatch as a dynamic vehicle routing system. In other words, each taxi follows an itinerary that is determined dynamically. Trip-based demand, even modeled with methods such as path flow-interception as described in the next section, cannot capture this dynamic itinerary. Lastly, passenger demand is stochastic, which results in a *stochastic dynamic itinerary* for each taxi. The AFV location literature has generally ignored such stochastic elements with a few exceptions.

A new model is proposed for locating infrastructure for electric taxis: the stochastic dynamic itinerary-interception refueling location problem with queue delay (SDIRQ). The model can also be applied to many other problems in urban logistics. The problem is presented as a bi-level simulation-optimization model, and tested on a network in Seoul, Korea spanning 603 km^2 (233 mi^2) with 600 taxicabs, a budget of 100 charging stations, and 22 candidate refueling locations. Section 2 provides a more in-depth literature review to show the motivation. Section 3 presents the bi-level SDIRQ model and two solution methods. A numerical study is provided in Section 4, followed by a conclusion in Section 5.

2. LITERATURE REVIEW

2.1. Refueling Location Problems

Facility location problems are optimization models used to determine a set of locations to serve neighboring demand in a network at a minimum or constrained cost. The cost is usually defined by the distances between the demand nodes and the server nodes. These problems have a long history and extensive literature spanning many different fields. Comprehensive reviews and introductions to the subject are available from Owen and Daskin (1998), Drezner and Hamacker (2002), and Snyder (2006).

In the studies on conventional-fuel vehicles, it is well recognized that conventional models with node-based demand do not handle the refueling station siting problem very well. This is because people generally will not make a trip from their home to the station solely for refueling. Both Upchurch and Kuby (2010) and Kang and Recker (2012) have convincingly demonstrated this point by comparing against the conventional models. Instead, three alternative approaches have been considered.

The first, and the more popular approach, has been the flow-capture or flow-interception location problem (Hodgson, 1990; Berman *et al.*, 1992). The problem is designed as a path-based version of Church and ReVelle's (1974) conventional maximal covering location problem. The shortest path between each OD pair is considered covered if it passes through at least one node that contains a server. The *ij* notation is switched to *p* path notation over the set of all shortest paths between each OD pair. Kuby and Lim (2005) extended the model to a flow-refueling location problem (FRLP), which serves demand from origin-destination flows along their shortest paths rather than demand at their end points. FRLP selects optimal refueling locations to maximize coverage of the these flows by considering vehicles traveling the full round trip path with a range constraint not present in the prior formulation. Kuby and Lim (2007) further extended the problem to allow solutions on arcs because the optimality condition of being on nodes was only true for the flow-interception problem, not the range-constrained flow-refueling location problem. Upchurch *et al.* (2009) added capacity constraints to the problem. Kim and Kuby (2012) relaxed the FRLP to allow penalized or maximum deviations from the shortest paths to get to fuel stations. Jiang *et al.* (2012) proposed a flow-based model that combined capacity as well as deviation flow.

Because of the need to enumerate facility combinations, the FRLP was not very tractable. Wang and Lin (2009) proposed a variant path-based set covering model that avoided the combination enumeration by using variables to track remaining fuel and range. Wang and Wang (2010) further combined path and node-based passenger demand as a multiobjective problem. Capar and Kuby (2012) reformulated their model in an effort to make it more efficient, also resorting to constraints based on range variables. MirHassani and Ebrazi (2012) proposed a new model based on both Kuby's and Wang's models that was more computationally efficient. Their model was extended by Chung and Kwon (2012) to handle multi-period staging of investment strategies.

A second approach is to ignore ODs altogether and just relate the value of a location based on vehicle-miles traveled (VMT) indicators (Lin *et al.*, 2008). While this simplifies the problem somewhat, it ignores the range from a given location to a fuel station and simplifies assumptions about the probability of a particular vehicle's location.

The third and most recent approach is to locate based on an itinerary, or trip chain of multiple ODs, instead of just a trip OD. Nourbakhsh and Ouyang (2010) integrated the path-based refueling location problem with the fuel scheduling problem for freight locomotives, which was an initial step in this direction in the refueling literature. Kang and Recker (2012) introduced a model to locate hydrogen stations where households respond to location decisions by rescheduling their daily itineraries using an activity-based variant of the pickup and delivery problem (Recker, 1995). This approach is more high resolution than the trip-based location solutions because it includes departure and arrival times throughout the day and maintains trip chains for each household member. It is the first study of an *itinerary-interception* refueling location model. The model is based on the location-routing problem (Perl and Daskin, 1985; Laporte *et al.*, 1988), and has been generalized to an activity-based network design problem (Kang *et al.*, 2013). The differences between these classes of models and their hypothetical solutions are shown in Figure 1.

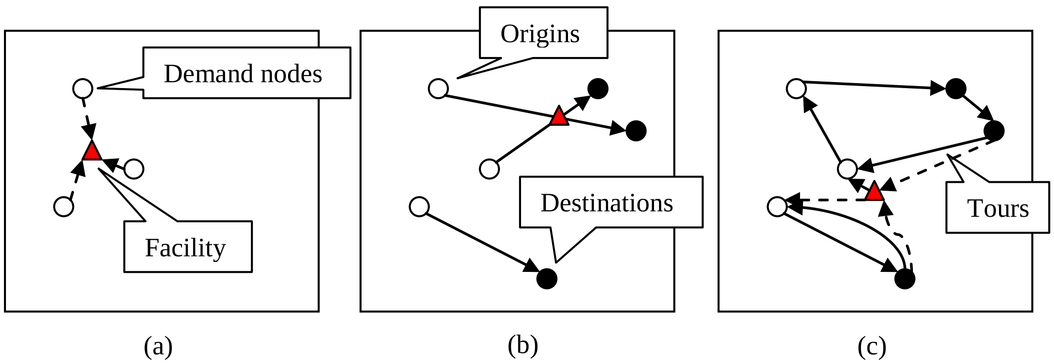


Fig 1. (a) Node-based, (b) flow-intercept, and (c) itinerary-intercept.

Figure 1(a) shows conventional node-based models, such as the maximal cover location problem. Figure 1(b) shows the flow interception models that try to cover as many OD flows as possible, acknowledging that travelers tend to refuel as part of a trip. In this second scenario, it would make more sense to have the single facility located further to the right where two out of three paths can be "intercepted". Figure 1(c) shows the itinerary-interception models where refueling is scheduled into the trip chains. The proposed SDIRQ is a new itinerary-interception location problem for refueling FTS and taxis in particular considering two added complexities not currently in this literature: 1) stochastic demand from customers that occur dynamically over time; and 2) queueing due to daytime charging for the service fleet.

2.2. Uncertainty

There have not been many attempts at stochastic variations of the FRLP, perhaps due to its complexity and also the belief that it is sufficient to assume routine consumer demands. Berman *et al.* (1995) extended their flow-interception model to include flows that are not known with certainty, using Dial's algorithm to help solve the Markov Decision Process at each node. Mak *et al.* (2012) was one of the few to consider demand uncertainty in a flow-interception location model for EV battery swapping stations. They used a robust optimization approach to address uncertainty in demand.

2.3. Queueing

Stochastic queueing has not been considered at all in the AFV refueling station location literature. Most refueling location problems have not considered stochastic demand, which explains why there have not been any efforts to evaluate stochastic models of standard queues such as the well-known Kendall-notation queues (e.g. M/M/c). Facility location problems that feature queueing are defined such that optimal locations do not just minimize or constrain distances from a demand node to a server node; they also include the delay caused by a queue at the server node. This is an important issue to include for limited infrastructure and long recharging times, especially in a much more stochastic environment such as in the fleet context with real-time dispatch that we focus on.

One of the earliest efforts to incorporate queues in a location model is Larson's (1974) hypercube queueing model. More recently, Boyaci and Geroliminis (2011) extended the model to handle large scale spatial queueing. Spatial queueing treats the whole coverage of demand as a service, and are more appropriate for mobile servers that relocate from one node to another.

For stationary location siting that covers a large neighborhood of nodes, the challenge is to find some way to quantify the steady state delay at a server node to include in the objective function. This is trivial for single server nodes that assume Markov arrivals and service times, but becomes cumbersome for multiple co-located servers. In AFV refueling location problems, there is usually a set number of refueling or charging stations at a given server node. It becomes a highly nonlinear function of the demand arriving at the server node, which leads to an intractable facility location problem if it is included in the objective function. In a special context of emergency vehicles' dispatch, Marianov and ReVelle (1996) linearized this expression by moving it to the constraint of a maximal availability location problem (MALP) and showing how an equivalent linear form relative to a reliability measure can be obtained. Marianov and Serra (1998) further extended the linearized queueing constraint to other location models. To date, there has not been such an extension to the path-based FRLP, much less to itinerary-based demand.

A similar area of research is on an allocation problem that considers queueing. In this problem, the locations are already set and the problem is to allocate the demand to the server nodes (or to allocate servers to given demand) in an optimal manner. For example, Xi *et al.* (2013) proposed an allocation model with exogenous deterministic demand using simulation to evaluate the deterministic multiple-server delay and level of service. The allocation problem with stochastic queueing is even more challenging and has its roots in queueing networks, where researchers sought to determine the optimal allocations of servers in a network to minimize total network cost. Approximations for the multiple server delay evaluation were considered very early on. Iglehart (1965) presented a proof of an approximation for an M/M/c model. Rolfe (1971) extended the approximation to M/D/c queues with deterministic service times. Kimura (1983) generalized the approximation to M/G/c queues. Lovell *et al.* (2012) introduced a continuum approximation approach to approximate M/M/1 queues.

Berman and Drezner (2007) proposed a multiple server allocation model that accounted for queues at the server nodes, assuming demand is allocated to the closest server node. The multiple server allocation problem can be solved exactly using a greedy algorithm to incrementally assign servers, as has been demonstrated in the queueing network literature (e.g. Rolfe, 1971).

2.4. Shared-Taxi Operations

In short, the literature on AFV refueling stations has been growing in recent years but they have dealt primarily with *consumer passenger vehicles*, using deterministic approaches that do not account for queueing. Furthermore, one investment strategy that may help overcome the chicken-and-egg problem—investing in publicly operated or regulated AFV fleets and infrastructure to primarily support them—has been largely disregarded. Investing in infrastructure for public fleets can reduce the risk of lack of demand. At the same time, the infrastructure would provide a positive feedback force that could drive the demand for the consumer AFV market.

Jung and Jayakrishnan (2012) showed that the shared-ride taxi (or shared-taxi) is a less constrained, and hence more generalized, operation than the standard taxi. A shared-taxi allows different groups of passengers to share a taxi ride. This trend is gaining traction in many countries in Asia. In China, the Beijing government recently allowed taxi-sharing due to the shortage of taxicabs during rush hours (China.org, 2012). In Singapore, Taiwan, and Japan, dynamic shared-taxi services are provided to link passengers who travel to the same area (Asiaone, 2012; Tao, 2007; Tsukuda and Takada, 2005).

Jung and Jayakrishnan (2012) proposed a real-time shared-taxi system operated by an online dispatch center with the help of communication technologies and geo-location services. The advanced shared-taxi service is capable of taking random service requests and updating vehicle schedules. Further details of the shared-taxi concept and simulation procedure, as well as a more in-depth review of the shared-taxi literature, are given by Jung and Jayakrishnan (2012).

2.5. BEV Taxi Charging

Unlike HEVs or PHEVs, BEVs are powered exclusively by the electricity from its battery pack. Most prevailing 5-seater vehicles produced by major automobile manufacturers are equipped with up to 24 kWh battery types that allow up to a maximum of 100 miles per charge depending on driving conditions. The EV charging station is commonly called EVSE (Electric Vehicle Supply Equipment), which has three levels of SAE (Society of Automotive Engineers) classification: (1) Level 1 with 120 volt outlet; (2) Level 2 with 240 volts; and (3) Level 3 with higher voltage DC power for fast charging, which takes less than one hour. These considerations are made in the computational study in Section 4.

3. PROPOSED MODEL

The SDIRQ model simultaneously seeks a set of locations and number of co-located servers at each of those locations to minimize the average delay—defined as the sum of travel to a facility, wait time at the facility to be served, and service time—for a set of random demand itineraries. This type of problem has many applications where demand can be characterized as an itinerary of trip chains that occur with a degree of uncertainty, including taxi fleets, activity based land use planning, military logistics, supply chain logistics, humanitarian logistics, among others. The specific application addressed in this study is the location of charging stations to minimize the delay to taxi fleets' itineraries, which are dependent on random customer arrivals over time. A simulation-based optimization problem formulation is presented to characterize this problem.

3.1. SDIRQ Model Formulation

Given the complexity of the shared-taxi operations, the location model is formulated as a bi-level problem. Bi-level programming problems split the decisions of the system planner (leader) and the system users (followers) into two levels so that the sub-problems are solvable and an iterative approach can be used to arrive at a local optimum. The upper-level problem is defined by

$$\min_u F[u, v(u)]$$

Subject to

$$G[u, v(u)] \leq 0$$

Where $v(u)$ is implicitly determined in the lower-level problem

$$\min_v f[u, v]$$

Subject to

$$g[u, v] \leq 0$$

where F and G are the objective function and constraint set of the upper-level; f and g are for the lower-level problem. The decision variables u and v are defined for the upper level and lower level, respectively.

In the case of the stochastic dynamic itinerary-interception refueling location problem with queue delay (SDIRQ), the master problem is split into an 1) upper level server allocation problem with queue delay, and a 2) lower level stochastic dynamic itinerary simulation.

3.1.1. Upper Level Formulation

The upper level problem is based on Berman and Drezner's (2007) multiple server allocation problem. The objective is to minimize the sum of the travel times and the average queue delays at each server node for all taxis. A complete graph $G(N, L)$ with a set N of n nodes and a set L of $m = n(n-1)$ links is considered. At each node $i \in N$, demand is generated at a rate of λ_i per time unit (e.g., an hour). This demand is obtained from the output of the lower level problem. The service rate of each server is μ customers per time unit with a maximum of p servers on the network. The set S of nodes which are assigned servers (server nodes) has a cardinality of $q \ll p$, which means that some of the nodes may not have servers. $C_i(S)$ is defined as the set of closest nodes in N to a node i in S, resulting in Equation (1) to determine the arrival rate at a service node $i \in S$.

$$\lambda_i(S) = \sum_{j \in C_i(S)} \lambda_j \tag{1}$$

The total travel times from demand nodes to all server nodes are modeled as a p-median objective. $W_k(\lambda, \mu)$ is defined as the expected time spent at a server node with k servers when the arrival rate is λ and the service rate is μ, where k is greater than or equal to one. The vector $K = \{k_1, \dots, k_q\}$ is the assignment of servers to the set S. The queue delay at each location can be represented as an M/M/k_i queueing system. The upper level objective function is then expressed as Equation (2), as shown in Berman and Drezner (2007).

$$\min_{S,K}\left\{F(S,K) = \sum_{i \in S}\left[\sum_{j \in C_i(S)} \lambda_j d_{ij} + \lambda_j(S) W_{k_i}(\lambda_i(S), \mu)\right]\right\} \tag{2}$$

subject to

$$\sum_{i \in S} k_i = p \tag{3}$$

Unlike Berman and Drezner's study, we assume that all demand nodes can be server nodes. Since $C_i(S)$ is not predefined in our study, the solution algorithm shown in Section 3.2 includes a consolidation process in which servers at a lower demand node can be consolidated with another node if it reduces the total queue delay. We set up an additional constraint for the maximum number of servers, $p_{i,max}$, at node i due to the geographical limitations of EV charging locations in an urban area.

$$0 \le k_i \le p_{i,max} \tag{4}$$

As pointed out by Pasternack and Drezner (1998), higher values of utilization ($\rho = \lambda / k\mu$) may cause small P_0 (the probability of no customers in the system) when calculating W (the average waiting time) in the standard way. Lower values of P_0 can cause computational overflows for subsequent calculations of L_q (the average queue length) and W. The same recursive method is used for the queue delay. For a given λ, μ, and k_i, the queue delay time can be obtained in a recursive manner shown in Equations (5) and (6).

$$a_1 = 1;\ a_i = 1 + \frac{\mu}{\lambda}(\mathrm{i} - 1) a_{i-1} \tag{5}$$

and

$$W_k(\lambda,\mu) = \frac{\lambda}{(k\mu-\lambda)^2\left(a_k + \frac{\lambda}{(k\mu-\lambda)}\right)} + \frac{1}{\mu} \quad (6)$$

3.1.2. Lower Level Problem

The lower level problem, $v(u)$ is replaced with a simulation similar to the recharging schemes previously designed for HCPPT by Jung and Jayakrishnan (2012). Because this problem requires great flexibility of implementing various types of vehicle routing algorithms and vehicle controls for door-to-door passenger services, which are not available in most commercial transportation simulation tools. However, there is a significant difference that needs to be considered. An EV taxi is assumed to visit a charging station only after completion of passenger delivery (similar to the assumption made by Kameda and Mukai (2011) for shared-buses in Tokyo). EV taxi charging is added as event at the end of vehicle schedule within the current driving range, but the event can be considered at any time step in the vehicle's movement. We assume that the dispatch algorithm is provided with the locations of charging stations so that it can determine when a connected taxi should be replenished and which station the taxi should visit. New pickup and delivery events are prohibited if the vehicle is already headed to a charging station or if the vehicle has higher risk of being discharged when the new schedule is performed given the remaining range. Further details can be found in Jung *et al.* (2013).

A simulation rule is used to insert a charging station visit when a taxi reaches a threshold and no passenger is being served. For battery charging, a vehicle tries to find the closest charging locations to minimize the travel distance and risk of being discharged after finishing its final delivery. The lower-level simulation generates a unique charging profile and queue delay at each charging location as well as a system performance for shard-taxi operation. The simulation time is set to be long enough so that vehicles need to visit charging stations multiple times. Since the stochastic taxi demand is determined by a simulation seed number, multiple sets of simulations are conducted for each iteration to minimize the random effect.

The simulation framework is depicted in Figure 2. The simulation container reads the configuration file and then creates the data access library to import maps, taxi service requests, and EV charging locations including charger allocations. The simulation time is synchronized by a timer that controls passenger request generation, real-time vehicle controls, and vehicle queue at charging locations.

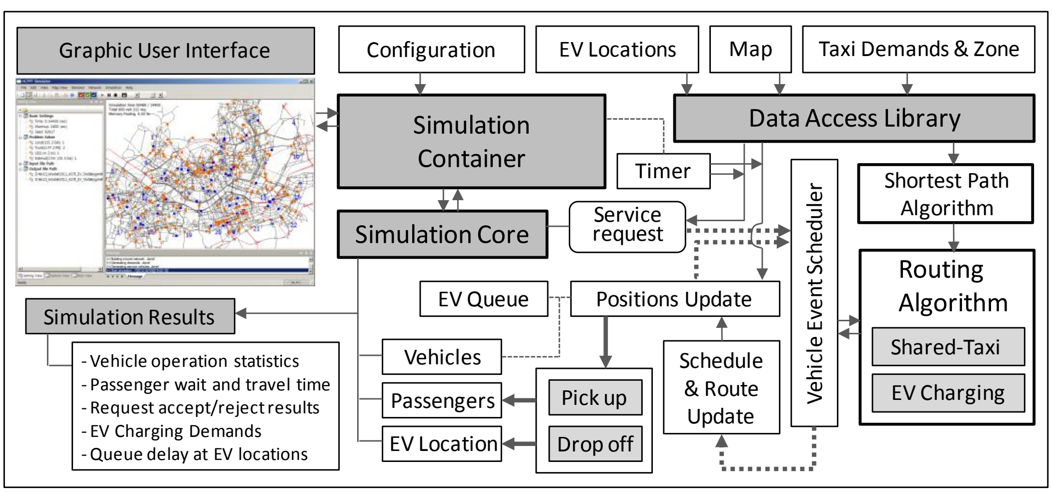


Fig. 2. Lower Level Simulation Framework.

In the preliminary simulation, we found that the arrivals at a majority of charging stations follow a Poisson process. However, some locations have lower sample sizes, depending on EV location allocations.

The lower samples are not enough to estimate the arrival rates because they can produce very noisy estimates. Second, the translated demand rate should always be less than the service rate according to the assumption made in the upper level (steady state intensity less than one). In other words, the basic inverse function shown in Equation (7) can fail in the upper level problem.

$$\lambda_i = W_{k_i}^{-1}(\mu, D_{q,i}) \tag{7}$$

The lower level simulation provides the number of servers k, service rate μ, and queue delay time D_q for each location. To address the intensity issue, an inverse interpolation method is used to approximate the corresponding demand rate so that the queue delay function in Equation (6) is continuously and monotonically increasing within an interval $\rho \in [0.0, 1.0]$ in which there is a single-valued inverse. A bisection method is applied to compute the inverse given the stable utilization interval.

3.1.3. Bi-Level Interactions

The iteration between the upper and lower level problems captures the cyclic effect that queue delay has on the dispatch of the taxi fleet, which in turn determine the demand patterns for the charging stations, resulting in the allocation of chargers that impact that queue delay. Figure 3 describes the iterative procedure for the simulation-optimization-based EV charger location problem. First, servers are evenly distributed over the candidate charging locations, and then the lower level simulation is called. The lower level simulation tracks all vehicle pickups, dropoffs, queueing, and charging events including which charging location is used. After the first run of lower level simulation, the charging profiles and queue delay are collected and transferred to the upper level problem. The upper level problem assigns servers optimally to minimize queue delay and travel time. Once the upper level problem is solved, the optimal allocation of chargers is compared with the previous allocation status. If there is no change, the procedure ends with the current charger allocation on candidate locations; otherwise, the lower level simulation is called with the updated charger allocation.

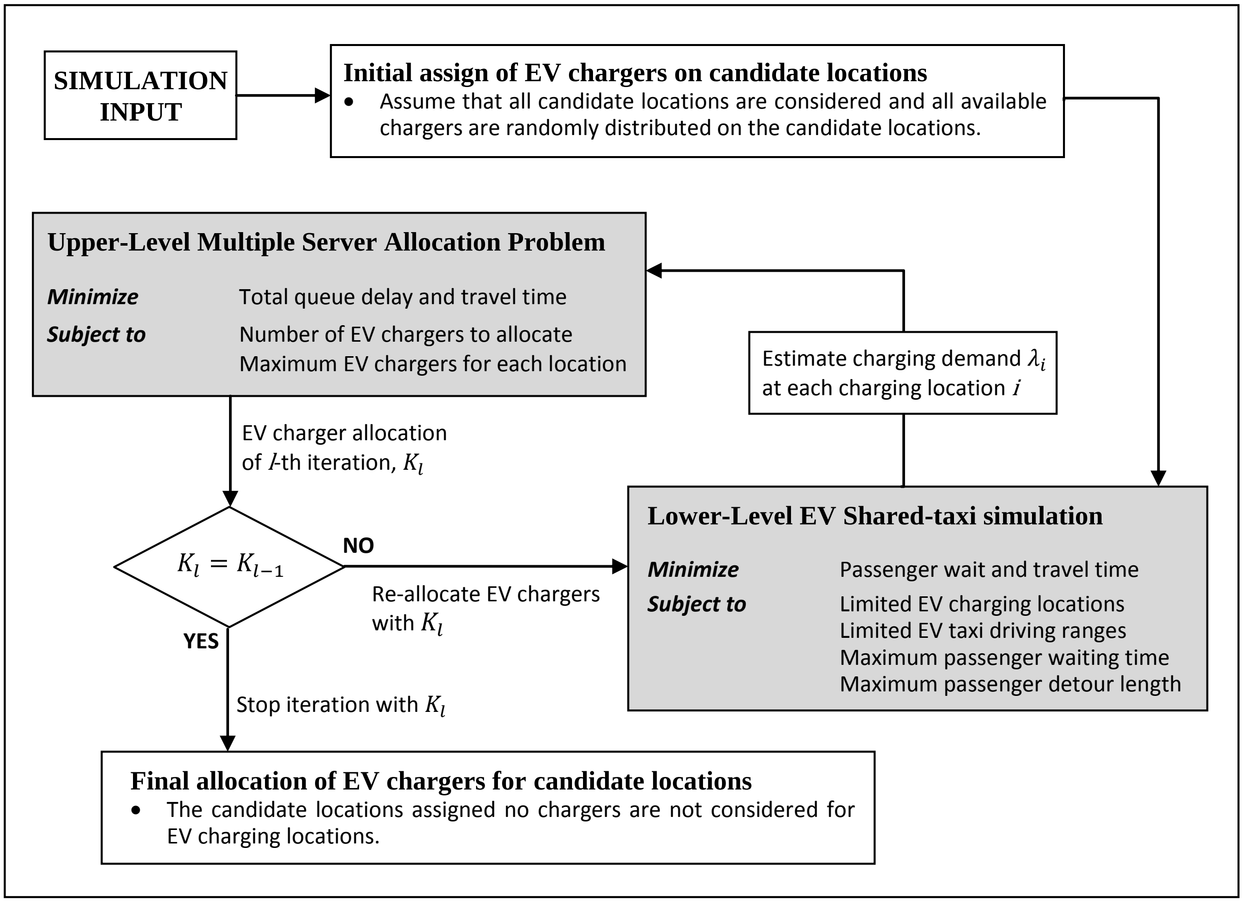


Fig. 3. Bi-Level Procedure for EV charger allocation.

A solution method is presented to solve this bi-level problem in Section 3.2. The method is based on a modification of Berman and Drezner's (2007) allocation algorithm. This solution method is compared to a benchmark single level solution method presented in Section 3.3. The benchmark method uses a genetic algorithm to assign chargers without considering any iterative interaction of queue delays in the upper level with the demand assignment in the lower level. It is used to validate the benefit of considering queue delay in the SDIRQ.

3.2. Proposed Solution Method – Bi-level

The solution of the upper level uses a modification of Berman and Drezner's (2007) Algorithm 1 that includes a consolidation procedure discussed in Section 3.1. At this level, the demand obtained directly from the lower level is assumed constant.

Proposed Upper Level Algorithm: (steps 1-4 are from Berman and Drezner (2007))

1. For each node $i \in S$, find the minimum number of servers necessary to service the demand. This minimum number is $l_i = \text{int}[\lambda_j(S)/\mu] + 1$.
2. If $\sum_{i=1}^{q} l_i > p$, then there is no feasible solution.
3. If $\sum_{i=1}^{q} l_i < p$, check the decrease in queue delay by changing l_i to $l_i + 1$. If $l_i < p_{i,max}$, increase the number of servers by one at the node with the maximum time decrease.
4. Repeat until $\sum_{i=1}^{q} l_i = p$.

5. For each node i, find the closest server node j with 1 or more servers, and compare the current objective value with relocating demand λ_i and servers l_i to the adjacent node j with $\lambda_j + \lambda_i$ and $l_j + l_i$ and considering travel cost d_{ij}. If $l_j + l_i \leq p_{j,max}$ and the objective value of relocating decreases, set node i and j as a candidate node pair.
6. Find a best node pair among the candidate node pairs to minimize the objective function, then consolidate node i to node j.
7. Repeat the step 5 and 6 until there is no improvement in the objective value.

We also propose an improvement to the dispatch algorithm from Jung and Jayakrishnan (2012). The original insertion heuristic starts by comparing all vehicles to find a best vehicle to minimize both passenger travel time and waiting time. However, the algorithm is computationally inefficient when both service area and fleet size are large.

The proposed algorithm consists of two steps. First, once a new passenger request, z_i ($i \in I$) comes in, the system selects available vehicles in the corresponding geographical zonal area to insert a new trip request. Each passenger trip is identified by its origin and destination. In this stage, available vehicles are filtered to prevent excessive computational burdens. At the second stage, the algorithm seeks the best vehicle to minimize service wait time and travel time of the new passengers as well as the previously assigned passengers. In this lower level simulation, the allocation of charging stations is obtained directly from the upper level and assumed fixed.

Proposed Lower Level Dispatch Algorithm

Stage 1: Identify the zonal area of the new passenger request, z_i, then prepare a vehicle set J based on z_i's trip points and time windows.

Stage 2: Find the best vehicle V_{min}, $V_{min} \in J$ satisfying the following objective function to insert l-th and m-th for the new request z_i in equation (5).

$$min\ IC_j = C_j(E_j \cup z_i) - C_j(E_j) \tag{5}$$

$$C_j(E_j) = \sum_{k \in K} \left[\mathrm{WT}(E_{j,k}) + \mathrm{TT}(E_{j,k})\right] \tag{6}$$

$$C_j(E_j \cup z_i) = min_{l,m} \sum_{k \in K+2} \left[\mathrm{WT}(E_{j,k}) + \mathrm{TT}(E_{j,k}) + \mathrm{WT}(z_{i,l}) + \mathrm{TT}(z_{i,m})\right] \tag{7}$$

where

z_i = new request from passenger i, $i \in I$
E_j= set of pickup and dropoff events in vehicle vehicle j's schedule
IC_j = incremental cost of vehicle j for inserting a new request z_i
$C_j(E_j)$ = current total cost of vehicle j's with schedule E_j
$\mathrm{WT}(E_{j,k})$ = waiting time (cost) associated with k-th event in E_j
$\mathrm{TT}(E_{j,k})$ = in-vehicle time (cost) associated with k-th event in E_j
$C_j(E_j \cup z_i)$ = total cost of vehicle j when adding a new request z_i
l, m = l-th pickup and m-th dropoff insertion positions for vehicle j's minimum cost with z_i

In stage 2, equations (6) and (7) denote the travel cost based on vehicle j's current itinerary and the updated cost by assuming that the new request is inserted into an optimal position among the existing pickup and dropoff events. When inserting a new request, all previously assigned events should be kept within the constrained time windows. If there are no more available vehicles to consider, the dispatch algorithm assigns the new request to the vehicle with minimum incremental cost, IC_j. Otherwise, the passenger request is rejected due to the constraint. The dispatch algorithm is then combined with the recharging scheme by Jung *et al.* (2013).

3.3. Single-level Solution Method

The single-level procedure is proposed as a baseline algorithm without the cyclic interaction of queue delay and demand between the upper and lower level. This method consists of simply evaluating simulation outputs of different EV charger allocation schemes and using Genetic Algorithm (GA) to improve the fitness of the population of solutions. Unlike the bi-level method, the direction finding is based purely on fitness function that includes queue delay, but no update of that cost is present in the simulation. This baseline algorithm is designed to show what happens if no queueing is considered by the taxi fleet dispatch even though the cost exists.

Single-level procedure

Step 1: EV charging demand allocation for each location is achieved from the lower level simulation by relaxing the queue delay constraint.

Step 2: The multiple server allocation problem is applied to allocate EV chargers.

In Step 1, it is assumed that enough chargers are installed at each location so that vehicles start battery charging without the queue delay effect when they visit charging locations. This approach provides a viable option in the sense that one can estimate the upper bound of potential EV charging demands when the deployment size is not determined yet (Jung and Jayakrishnan, 2011). However, such an upper bound can be seriously overestimated especially when the charging frequency of EV fleet operation is higher. As a result, the server allocation problem in Step 2 might end up with an infeasible solution that has a higher demand rate than the service rate.

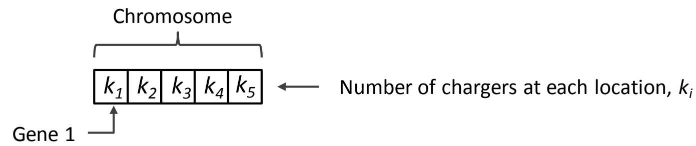
GAs have great flexibility to solve the combinatorial optimization problem with complex contratins. Although GAs do not guarantee optimal solutions, they are efficient in seeking approximate solutions in *NP*-hard problem. The solution procedure employs an initial population, selection, elitism, crossover, and mutation. Figure 4 provides an illustrative example of chromosome encoding and genetic operators designed for the single level procedure with five candidate locations. In figure 4(b), as the crossover operation may produce an infeasible solution – having more or less chargers in total, an additional modification procedure is required to keep the same total number of chargers. The modification procedure adds or subtracts additional chargers on a random position to keep the total number of chargers the same.

The EV charging demands from simulation results are translated into the demand rates because there is no queue delay measured in the simulation results with infinite numbers of servers. Instead, a finite number of servers are assumed to consider queue delay at each charging location. If the demand rate goes over the service rate given by the number of servers at a location, a penalty value, W_p, is applied with a linear relationship of queue utilization, ρ, as shown in Equations (8) and (9).

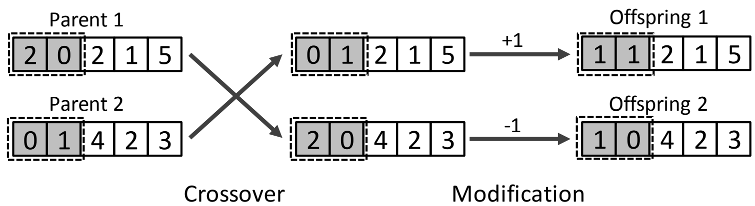
$$W_{k_i} = \begin{cases} W_k(\lambda_i(S), \mu \cdot k_i) & \text{if } \rho_i < 1.0 \\ W_p \cdot \rho_i & \text{otherwise} \end{cases} \tag{8}$$

where

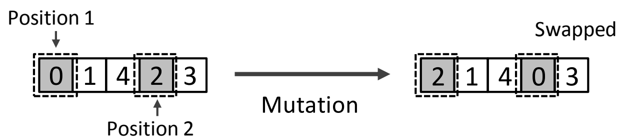
$$\rho_i = \lambda_i(S)/\mu \cdot k_i \tag{9}$$



(a) Chromosome Encoding



(b) Crossover



(c) Mutation

Fig. 4. Chromosome Encoding and Genetic Operators in GA.

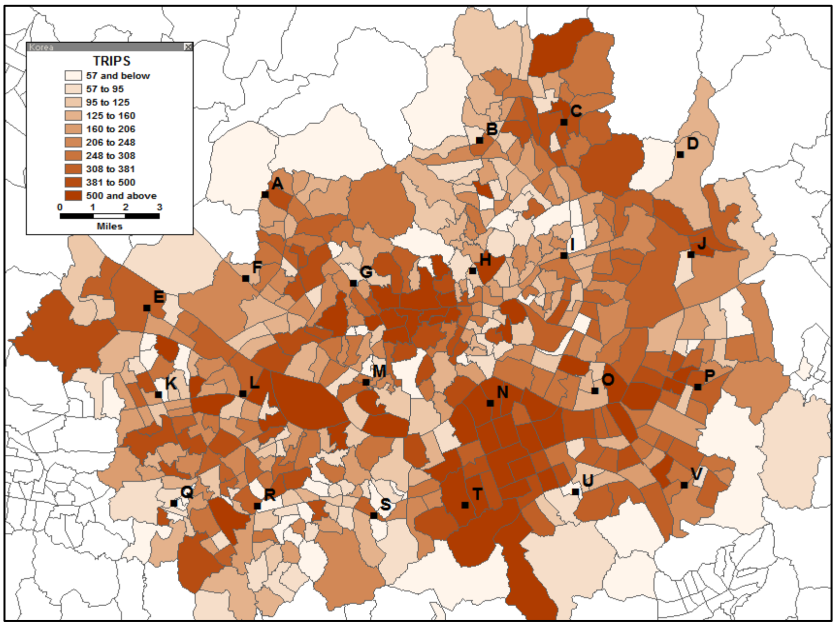
4. COMPUTATIONAL STUDY

The goal of the computational study is to compare the performance of the proposed bi-level solution algorithm against the benchmark "naïve" single-level GA approach. Because stochastic dynamic shared-taxi dispatch can only be evaluated as an itinerary, the earlier models of flow-interception or even node-based location models cannot be run without making crippling assumptions. In addition, the improvement of deterministic itinerary-interception models against node-based models has already been documented in Kang and Recker (2012). This comparison was to demonstrate the feasibility and the performance of the bi-level algorithm with and without the cyclic interaction between the demand and the allocation with queue delay that is present in the method.

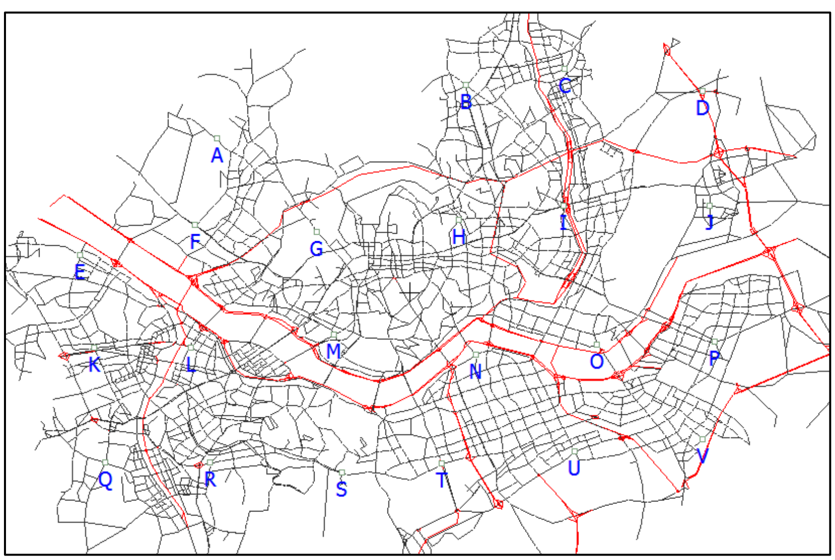
4.1. Study Setting and Parameters

In this study, the Renault-Samsung SM3 Z.E. is assumed, which is a model built by Renault-Samsung Motors for South Korea, based on the Renault Fluence Z.E. Renault-Samsung claims that the SM3 Z.E. is outfitted with a 24 kWh lithium-ion battery with a maximum range of 115 mile (184 km) measured on the NEDC (New European Driving Cycle) combined cycle, which is similar to Nissan LEAF. A "QuickDrop" system (known as battery swapping) allows the discharged battery to be replaced quickly with a fully charged one at a dedicated EV battery switch station. In addition, a fast charging system using a 32A 400V 3-phase supply enables the battery to be charged in 45 minutes. Since an official Korean range specification is not yet established, we assume conservative average ranges of 112 km (70 miles) on highways and 128 km (80 miles) on city traffic with a 16-km range of random variation for each individual vehicle.

Stochastic taxi demand is obtained from an EMME/2 transportation planning model developed at the Korea Transportation Institute (KOTI). As of 2011, trip demand consists of automobile, bus, subway, rail, taxi, and other types of demand, which cover Seoul with a total of 560 centroids over 603 km^2 (233 mi^2). Under the usual assumption of spatial uniformity of demand around a zone centroid, point-to-point taxi demands are randomly generated in accordance with destination probabilities of the taxi demand table in each centroid. The point-to-point demands are projected to the nearest road segment with its direction for door-to-door services except on limited-access roads. Real-time service requests arrive according to a Poisson process in a temporal manner. Figure 5(a) presents the spatial distribution of taxi demands including both origins and destinations used for stochastic trip requests for an 8-hour simulation with the minimum trip length of 1.0 km for taxi service. In figure 5(b), a total of 22 candidate charging stations on 5 km by 5 km reference grid cells are assumed over the road network in the EV taxi simulator. One hundred chargers are evenly distributed over the 22 candidate locations in the initial simulation setup.



(a) Distribution of Stochastic Taxi Demand



(b) EV Charging Locations and Transportation Network in Seoul

Fig. 5. Taxi Demands, EV Charging Locations on Transportation Network.

According to the statistics from Seoul Taxi Association, a taxi carries more than 35 passengers each day on average. The average travel distance is over 250 km/day and 5 km/trip, with a significant portion of the travel distance being non-revenue segments when the taxis travel empty for the pickups. A total of 72,000 taxi licenses are registered including owner-driver taxi and a total fleet of 40,000 vehicles due to the taxi rotation system to restrict the number of operating taxicabs operating as of 2011. Shared-taxis are prohibited by regulations in Korea, but it is publicly known that many taxicabs carry multiple passenger groups at the same time if they are traveling to the same destination because of limited supply. It is assumed that EV taxi service offers dynamic shared-ride for passengers in order to increase its system performance. According to 2014 Master Plan for Electric Vehicle by Seoul, Korea, the city plans to deploy around 30,000 commercial EVs including buses, taxis, and private vehicles. Since it is expected that EV taxi will be able to operate over a distance of 200 km to 400 km a day, Seoul seeks to initially establish replenishing stations with high-speed charging (SAE Level 3) or battery replace facilities.

In this study, 600 taxicabs are initially considered for electrification, which is equivalent to 1.5% of the total number of vehicles operated in Seoul. The initial positions of vehicles are randomly generated over the simulation area. The simulation time is set to 8 hours including 30-min as a warm-up period, which is assumed to include peak operating hours. An average of 1-min boarding and alighting times are assumed for each passenger, which is based on a random uniform distribution $U(0.5, 1.5)$. At each lower level simulation step, ten simulation sets are conducted by choosing ten different seed values.

In the shared-taxi scenario, we set maximum waiting time (900 sec) and maximum detour length (maximum 1.2 times of door-to-door travel distance) for passenger time windows. A taxicab can carry a maximum of four passenger groups at the same time, which provides a realistic scenario rather than having three or more groups. The charging time is assumed to be exponentially distributed with an average of 45 minutes. From our preliminary simulation, the majority of generated trip demands are within 10 km. The average trip length is 6.4 km and the expected door-to-door travel time 13.4 min under the assumption that vehicles can travel at 60~90% of the posted speeds on the network.

Two initial charging strategies are possible, one starting with fully charged EVs (FISOC: Full Initial Sates of Charge) and the other with randomly charged EVs (RISOC: Random Initial State of Charge) at the beginning of the simulation. Jung (2012) found that charging demands tend to be concentrated at certain times during the simulation when the eight hour period started with fully charged vehicles, so RISOC is assumed in this study. This study provides four different scenarios: (1) Single-level approach; (2) Bi-level approach; (3) Infinite chargers intended to relax charging delay factors; and (4) Non-EV operation with conventional ICE taxicabs.

Table 1
Scenarios for computational study.

EV Shared-taxi simulation	
Shared-taxi settings	
Service area (km^2)	605
Simulation time (hours) / Warm up (hours)	8.0 / 0.5
Number of service vehicles	600
Vehicle capacity (seats/vehicle)	4
Number of centroids (zonal areas)	560
Max. wait time (min) / Max. detour factor	15 / 1.2
Vehicle types	EV / ICE
EV settings	
Battery Capacity[1] (kWh) / Battery Charging Time (average min)	24 / 45
Ranges (km)	112 (Highway) / 128 (City Traffic)
Number of charging stations (/each candidate location)	Min. 0 – Max. 20 chargers
Charging schemes	Random Initial State of Charge (RISOC)
Candidate EV locations / Max. available chargers	22 / 100
EV Charging Operations	Single-level, Bi-level, Infinite Chargers, Non-EV

[1] Battery pack capacity (kWh): kilowatt-hour(s)

4.2. Results

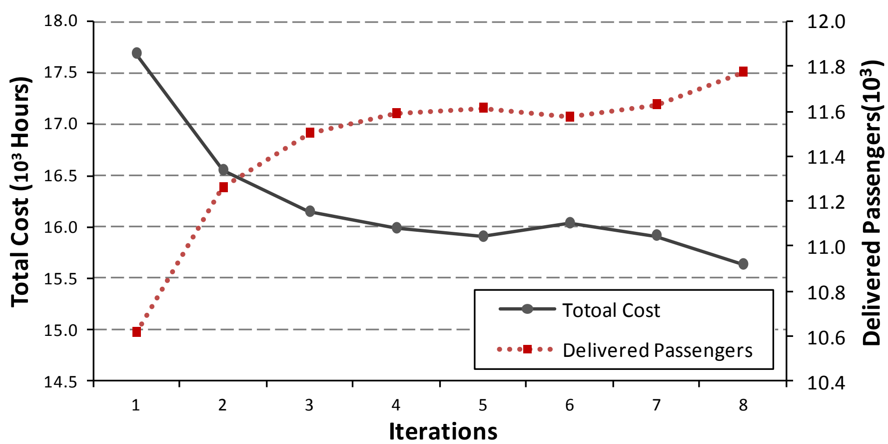
The system performance of the lower level simulation is calculated with passenger delivery and its associate wait and travel time. Since each simulation could have different number of delivered passengers and dropped requests, the total cost is proposed as the lower level objective value based on the number of delivered and rejected requests, average waiting time, and average travel time. A penalty value is applied for dropped passenger requests. When the cost is found this way, a lower value indicates better performance in Equation (10).

$$\text{Total Cost } (C_t) = T_p \cdot P^r + \sum_{i \in I} \mathrm{TT}(P_i^c) + \sum_{i \in I} \mathrm{WT}(P_i^c) \quad (10)$$

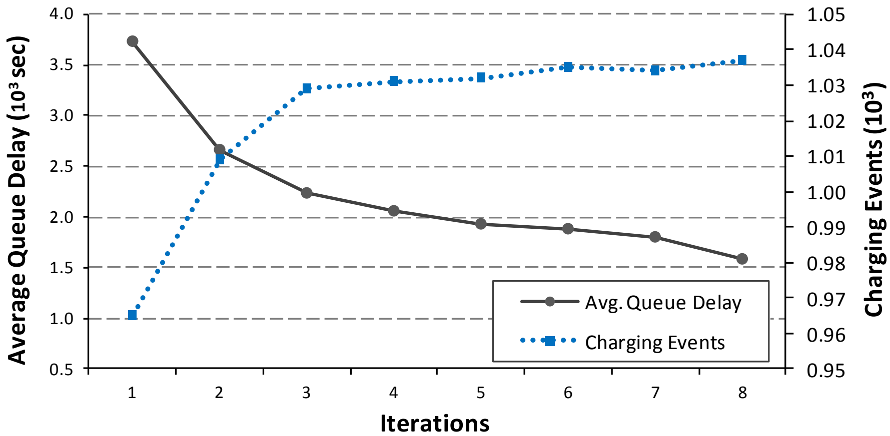
T_p: Penalty value for a dropped request, assumed to be 7200 sec
P^r: A set of rejected requests during the simulation
P^c: A set of completed requests during the simulation
$\mathrm{TT}(P_i^c)$: Travel time of passenger $P_i^c \in P^c$
$\mathrm{WT}(P_i^c)$: Waiting time of passenger $P_i^c \in P^c$

The number of charging events and queueing delay (D_q) at each charging location is measured as charging system performance since the upper level allocates EV chargers to minimize the queue delay. The lower level simulation model includes a queueing model for each charging location so that it creates a simulated queue delay given the charger allocations and the candidate location.

Figure 6 shows the overall SDIRQ performance by bi-level iterations in which the procedure converges at the eighth iteration. The total cost of lower level simulation significantly decreases during the first three iterations in figure 6(a). The minimum cost is achieved with 15,632 hours at the eighth iteration. At the same time, it shows that the number of delivered passengers increases from 10,615 to 11,773 passengers, which is consistent with the objective function to minimize passenger wait and travel time.



(a) Total cost and number of delivered passengers



(b) Average queue delay and number of charging events

Fig. 6. Overall SDIRQ Performance by Bi-level Iterations.

When comparing the average queue delay and the number of charging events in Figure 6(b), the proposed procedure not only decreases the average queue delay at charging locations, but also increases the number of charging events at the same time. With the initial setting, each EV taxicab spends an average queue delay of 3,730 sec/visit (62.2 min/visit) excluding battery replenishing time, which could be a significant loss of service time when considering a vehicle might visit charging stations with an average of 1.75 times during the 8-hour operation. As the iteration increases, the average queue delay drops to 1,580 sec/visit (26.3 min/visit), which enables vehicles travel longer distance for its operation. The average vehicle traveled distance increases from 130 km to 148.3 km as the procedure converges. In other words, given the same size of fleet, the decreased queue delay at charging locations can improve the EV fleet utilization, which causes more frequent charging behavior.

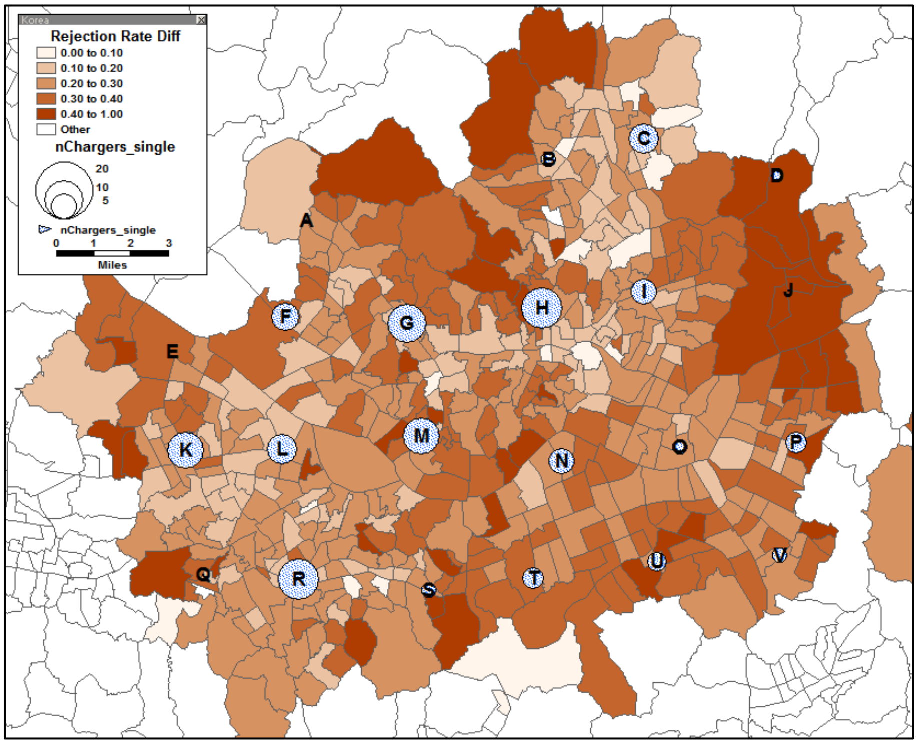
Table 2
System Performance.

System Performance	EV Operations			Non-EV
	Single-level	Bi-level	Infinite Chargers	
Total delivered passengers	11,273	11,773	12,940	15,738
Avg. wait time home (min)	12.71	12.63	12.47	11.76
Avg. passenger travel time (min)	14.73	14.77	14.77	14.75
Rejected passenger requests	5,700	5,128	3,872	947
Avg. vehicle load (pax/vehicle)	0.63	0.66	0.73	0.88
Avg. vehicle traveled distance (km)	140.05	148.29	166.47	204.71
Avg. operating hours except charging time (hour)	5.25	5.62	6.16	8.00
Total cost (hours)	16,557.66	15,632.20	13,621.81	8,849.24
Number of vehicle visits at EV locations, E_d	1,013	1,037	1,142	n/a
Number of charging events completed, E_c	875	971	1,142	n/a
Number of selected EV locations	18	12	22	n/a
Avg. queue delay at charging locations (sec)	2,509	1,580	0	n/a
Total queue length at charging locations (veh)	98.0	62.8	0.0	n/a
Avg. queue length at each location (veh)	5.4	5.2	0.0	n/a
Avg. distance to charging locations (km)	2.53	2.83	2.39	n/a

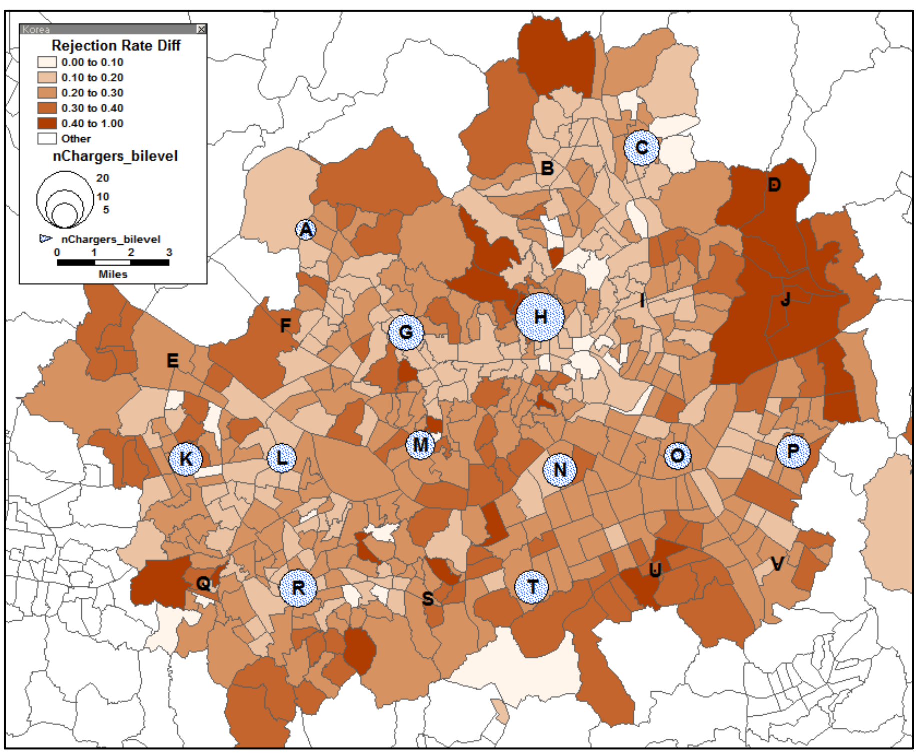
Table 2 reports the detailed simulation results comparing single-level, bi-level, infinite chargers, and non-EV. For non-EV condition, it is assumed that refueling ICE taxicabs is not considered in this condition because regular taxicabs in Korea use engines powered by Liquefied Petroleum Gas (LPG). Their refueling time is not considered in this simulation since there are more than 70 LPG refueling locations in that area. Note that the non-EV condition causes additional financial and environmental costs from EV taxi operations which are not considered in this study.

A higher average vehicle load indicates greater efficiency in taxi operation. Since the same vehicle routing algorithm is applied to all scenarios, the average vehicle load gets worse as vehicles stay longer at the charging locations. The single-level procedure ends up with a total of 18 locations selected for allocating 100 chargers while the bi-level result shows 12 locations with the same total number of chargers. A significant difference in the queue delays is measured between single-level and bi-level. The single-level result reports an average 2,509 sec/visit (41.8 min/visit) of queue delay at each location, which is about 15.5 min/visit more than the bi-level result. An average distance to charging location is defined as a weighted distance by the number of charging events representing from a vehicle’s last delivery position to the charging location where the vehicle is headed. If all candidate locations are used, the average distance is 2.39 km. The bi-level result shows the 0.3 km higher distance due to the smaller number of charging locations selected. Average actual taxi service hours with EV operations show less than 77% of Non-EV even without queue delay impact. Single-level shows 5.25 hours of service operation, which indicates average 2.75 hours are spent on moving to charging locations, queueing, and refueling.

Figure 7 reports the distribution of EV chargers on the geographical area. The colored polygons show the spatial distribution of increase in request rejection rate compared with non-EV scenario. The thicker color means there is a greater increase in passenger pickup request rejections in a zone. The size of circles represents the number of chargers allocated at each candidate location. The rejection rate increases around the EV charger allocated areas are less than the other areas. For example, the zonal areas around locations D, J, and Q show higher rejection rates in both Figure 7(a) and 7(b). This is because vehicles would choose nearby pickup requests when they finish EV charging. The bi-level solution clearly shows lower rejection rate than the single-level solution. Both solutions try to place more chargers in the center of Seoul. The spatial distribution of chargers in Figure 7(b) shows a similar pattern in comparison with the distribution of stochastic taxi demand in Figure 5(a).



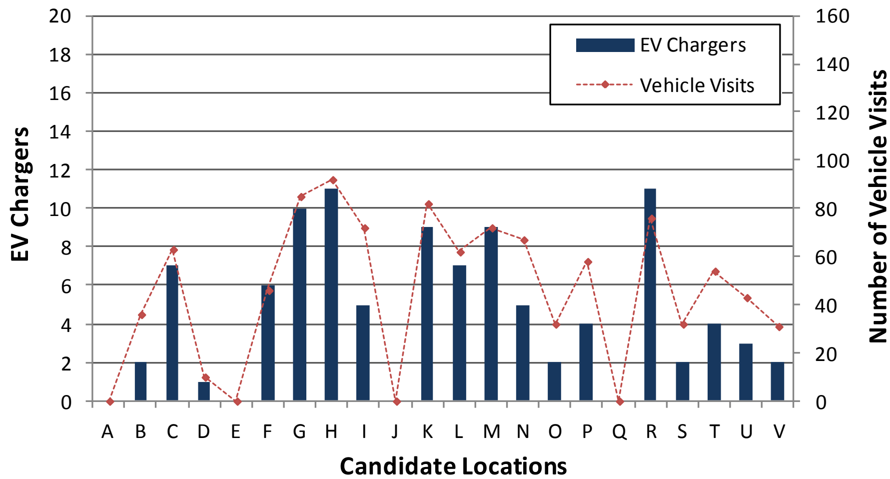
(a) Single-level



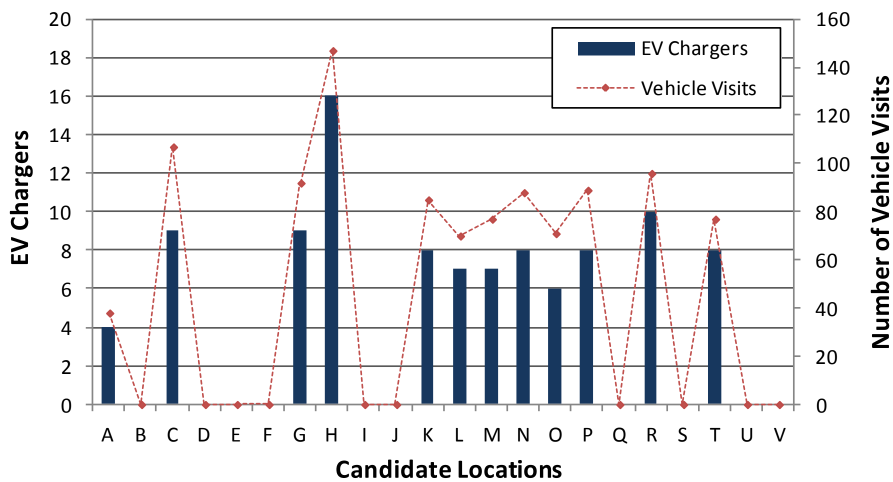
(b) Bi-level

Fig. 7. Changes in Passenger Rejection Rate of EV Taxi Operation.

Figure 8 shows the comparison between the number of vehicle visits and the number of chargers allocated at each location. The EV charger utilization (vehicle visits / charger) shows higher values with the bi-level results.



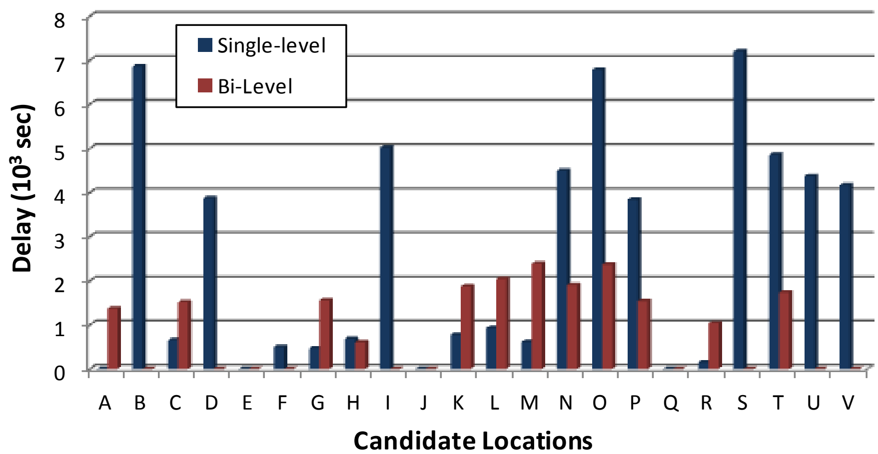
(a) Single-level



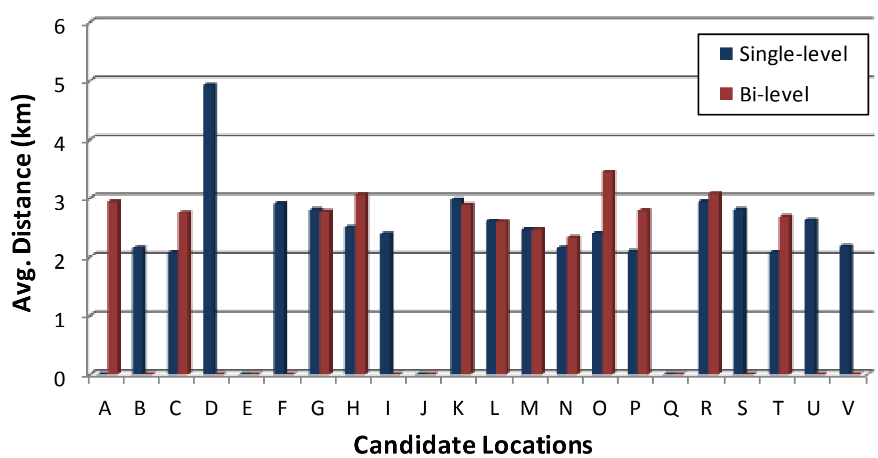
(b) Bi-level

Fig. 8. EV Charger Allocation and Simulated Charging Demands.

Figure 9 provides a side-by-side comparison for average queue delay at charging locations and average distance to charging locations. The average delays show significant difference between single-level and bi-level results. Locations such as C, G, K, L, and M in Figure 9(a) show lower average queue delays with single-level, but they also show extremely higher queue delays at locations with a fewer numbers of chargers. In contrast, the bi-level results show lower deviation in average queue delays on all selected locations. Figure 9(b) reports the average distance to charging locations. Most of them remain less than 3 km except the location D in single-level. These values can be used for re-calibration of the critical battery level considered in the recharging scheme. Note that these distances to charging locations are affected by the location distribution itself, in which vehicles tend to reject passengers whose destinations are far from any of charging locations when the battery level gets critical.



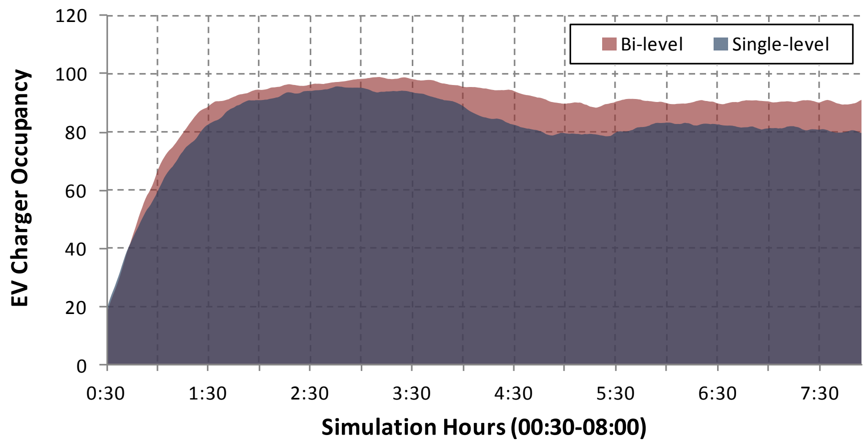
(a) Average Queue Delay at Charging Locations



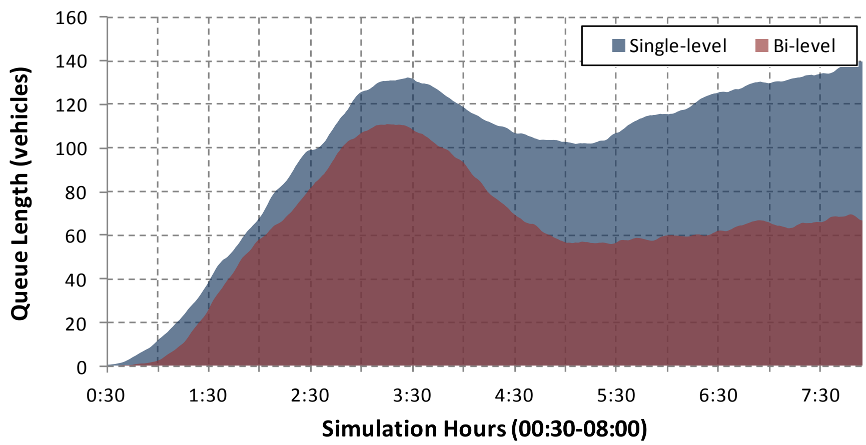
(b) Average Distance to Charging Locations

Fig. 9. Queue Delay at Locations and Average Distance to Charging Locations.

The temporal distribution of the aggregated EV charger occupancy and total queue length at charging locations are provided in Figure 10. The warm-up time period is not included. As the simulation starts, vehicles are visiting charging stations and saturate them after two hours in Figure 10(a). The bi-level solution shows an 8.2% higher occupancy level of EV chargers than the single-level. That indicates higher utilization of charging system with the bi-level solution given the same refueling time. Once the occupancy of chargers is saturated, the queue length starts decreasing in both solutions. Indeed, it is conceivable that the charging demands of EV fleet operations are not constant, but the demand itself can be influenced by the charging system performance. In other words, as charging locations hold more vehicles for longer in the queue, the subsequent charging demands decrease as well as the system delivery performance due to the changes in vehicle itinerary. This will be an important problem to be considered when available chargers are not enough to support a deployed fleet size.



(a) EV Charger Occupancy (chargers)

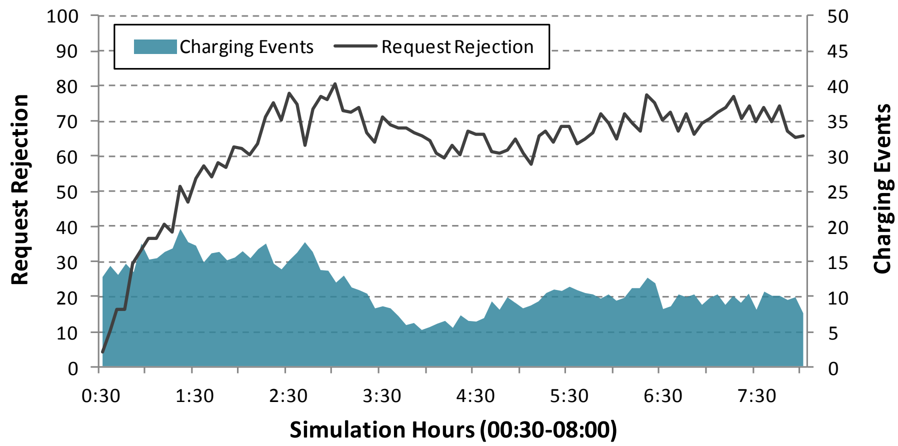


(b) Total Queue Length (vehicles)

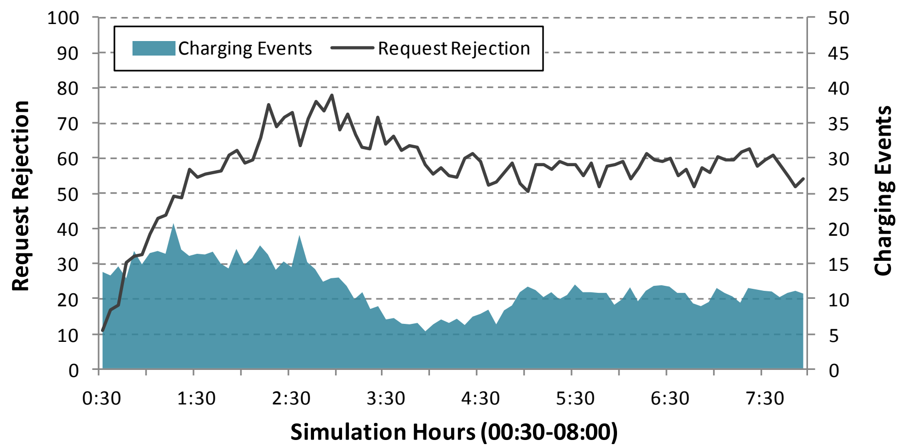
Fig. 10. EV charger Occupancy and Total Queue Length at Charging Locations

Regarding the queue length, both solutions reach their maximums around 3:30, and then start decreasing. The peak charging demand combined with the occupancy and the queue length is estimated around up to 3:00 after warm-up. The queue length of the single-level solution keeps increasing after 5:00. The overall average queue length of the single-level solution is 55% higher than the bi-level. Considering the bi-level solution generates more charging events than the single-level, the EV charger distribution and the location selection of the single-level is clearly worse than the bi-level solution.

Figure 11 compares the temporal relationship between the charging events generated and the passenger pickup request rejections. The request rejection curves are aggregated every 5 min. As mentioned, the peak charging demand is shown during the first three hours, and then the charging demand decreases. The passenger request rejection peaks around 3:00, which is consistent with holding more vehicles at charging locations as shown in Figure 10. The EV charging demands are derived from the taxi operation caused by passenger requests. The spatial and temporal characteristics of charging demand can be significantly influenced by the taxi demand, and neither of these characteristics can be analyzed using node-based or flow-based model location models.



(a) Single-level



(b) Bi-level

Fig. 11. Passenger Request Rejection and Charging Events

In summary, these results show that the proposed model is capable of locating charging stations with stochastic dynamic itinerary-interception and queue delay. The proposed bi-level solution method improves upon the benchmark algorithm without the interaction effects (allocation of fixed demand and demand realization with fixed allocation, but no feedback mechanism) in terms of realized queue delay (37%), total taxi service operating hours (6.7%), and reduced service request rejections (10%). Furthermore, we show how much additional benefit in level of service (17.7%) can be achieved if the current budget of 100 charging stations approaches infinity.

Obviously, having longer charging times will reduce the number of passengers served on average. Compared to the non-EV scenario for the 600 shared-taxis, having infinite number of chargers (queueing no longer becomes an issue), the reduction in number of delivered passengers is 17.8%. With only 100 charging stations, the number of delivered passengers is 25.2% of the non-EV scenario. These results, along with the others presented in Table 2, give decision-makers trade-off values to consider in their planning. For example, it is possible to evaluate the change in queue delay, allocation of charging stations, and number of passengers served if the charging time can be technologically improved by 50%. It is also possible to modify the simulation so to evaluate the presence of charging station queue informatics to the central dispatch. Currently the simulation considers taxicabs going to the closest server node, but a more efficient solution may be achieved if the dispatch is informed on average wait times at each node in real time. The difference in objective value would be the value of investing in that information. These trade-off values will be crucial for justifying the research and development investments in the battery technology, in equipping the charging stations and fleet with proper information flow, and in convincing taxi companies to switch to electric vehicles.

5. CONCLUSION

Implementation of EV FTS has received growing attention from policy-makers because mobile technologies make it easier to operate FTS and the clean energy technologies produce less GHG emissions in urban areas. This study focuses on the EV facility location problem for EV taxi service with FTS fleets in which vehicles operate with a stochastic dynamic itinerary. As opposed to the conventional refueling station problems with node-based static charging demands, the proposed SDIRQ is a simulation based bi-level optimization model in which EV refueling events are dynamically derived from the spatial and temporal passenger demand levels. The lower-level problem involves an EV shared-taxi simulation model to minimize both pickup requests and passenger travel times, while the upper-level objective offers the optimal refueling locations and the allocation of chargers.

For the baseline comparison, a single-level approach is introduced by relaxing the charging constraints to collect static charging demands in a conventional manner. Due to the complexity of involving the queueing model in the objective functions, a bi-level solution approach is proposed. The computational results report that SDIRQ with the bi-level approach shows a promising solution approach given the demonstrated EV FTS.

It is worthwhile to mention the lack of spatial-equity in the proposed solution. For example, it only takes into account the benefit distribution of EV locations and distributions for EV FTS operations. As mentioned in the simulation results, customers in the zones that are far from the EV locations may have lower chances to be served by EV taxicabs, in which case those customers are significantly worse off. On the other hand, the customers near charging stations will have much higher chance to be picked up. This might be a practical issue when determining the optimal EV charging facility locations. The spatial-equity issue will be studied further.

Another area of future research is the integration of the loading on the power grid or incorporating smart grid features, perhaps similar to He *et al.*'s (2013) study. Having an integrated approach can allow scheduling both the charging location and time to minimize delay on the fleet and temporal loading constraints on the grid. Time-of-day pricing should be considered in this case.

The location problem may also be viewed as a sequential, dynamic decision under uncertainty to introduce flexibility to adapt the investment strategy over time, similar to the forecast horizon introduced in the dynamic server relocation problem by Chow and Regan (2011). The recent work by Chung and Kwon (2012) shows a promising effort in that direction, introducing a multi-period aspect to the refueling location problem, although demand is kept deterministic.

ACKNOWLEDGMENTS

Dr. Jung and Dr. Jayakrishnan were funded by the Korea Transport Institute (KOTI), as part of "2011 Electric Vehicle Research on Business", initiated by the National Research Council for Economics, Humanities and Social Science, South Korea. Dr. Chow was partially supported by funding from the Canada Research Chairs program.

REFERENCES

[1] Alshalalfah, B., Shalaby, A., 2012. Feasibility of flex-route as a feeder transit service to rail stations in the suburbs: case study in Toronto. *Journal of Urban Planning and Development* 138(1), 90-100.
[2] Asiaone.com, "Taxi-sharing app to launch here end June", June 8, 2012, http://www.asiaone.com/Motoring/News/Story/A1Story20120608-351278.html, accessed April 15, 2013.
[3] Barth, M., Todd, M., 2001. User behavior evaluation of an intelligent shared electric vehicle system. *Transportation Research Record* 1760, 145-152.
[4] Berman, O., Larson, R.C., Fouska, N., 1992. Optimal location of discretionary service facilities. *Transportation Science* 26(3), 201-211.
[5] Berman, O., Krass, D., Xu, C.W., 1995. Locating discretionary service facilities based on probabilistic customer flows. *Transportation Science* 29(3), 276-290.
[6] Berman, O., Drezner, Z., 2007. The multiple server location problem. *The Journal of the Operational Research Society* 58(1), 91-99.
[7] Better Place, Tokyo Electric Taxi Project Overview. http://www.betterplace.com/global/progress. Accessed March 02, 2012.
[8] Boyaci, B., Geroliminis, N., 2011. Extended hypercube models for large scale spatial queueing systems. *STRC 2011*, 19p.
[9] Capar, I., Kuby, M., 2012. An efficient formulation of the flow refueling location model for alternative-fuel stations. *IIE Transactions* 44(8), 622-636.
[10] Cervero, R. Paratransit in America: Redefining Mass Transportation. Praeger Publishers, Westport, Conn., 1997.
[11] Cheng, R. and Gen, M. Evolution program for resource constrained project scheduling problem. Evolutionary Computation in IEEE World Congress, pp. 736-741, 1994.
[12] China.org.cn, 2012. "Beijing promotes taxi sharing to ease traffic jams". http://www.china.org.cn/video/2012-03/31/content_25040109.htm, accessed March 29, 2013.
[13] Chow, J.Y.J., Regan, A.C., 2011. Resource location and relocation models with rolling horizon forecasting for wildland fire planning. *INFOR* 49(1), 31-43.
[14] Chung, S.H., Kwon, C., 2012. Multi-period planning for electric-car charging station locations: a case of Korean expressways. Working paper, 31p., http://infohost.nmt.edu/~schung/chung2012.pdf.
[15] Church, R., ReVelle, C., 1974. Maximal covering location problem. *Papers in Regional Science* 32(1), 101–118.
[16] Cortés, C.E., Jayakrishnan, R., 2002. Design and operational concepts of high-coverage point-to-point transit system. *Transportation Research Record* 1783, 178-187.
[17] Crainic, T.G., Errico, F., Malucelli, F., Nonato, M., 2012. Designing the master schedule for demand-adaptive transit systems. *Annals of Operations Research* 194(1), 151-166.
[18] Dial, R. B., 1995. Autonomous dial-a-ride transit introductory overview. *Transportation Research* 3C(5), 261-275.
[19] Drezner, Z., Hamacker, H. (eds), 2002. *Facility Location: Applications and Theory*, Springer, New York.
[20] eGlobal Travel Media, 2011. "Seoul unveils '2014 Master Plan for Electric Vehicle'". http://www.eglobaltravelmedia.com.au/destinations/seoul-unveils-%E2%80%9C2014-master-plan-for-electric-vehicle%E2%80%9D.html, accessed March 24, 2013.
[21] FedEx. FedEx Cleaner Vehicles. http://about.van.fedex.com/article/cleaner-vehicles. Accessed March 20, 2012.
[22] Hadas, Y., Ceder, A., 2008. Multiagent approach for public transit system based on flexible routes. *Transportation Research Record* 2063, 89-96.
[23] He, F., Wu, D., Yin, Y., Guan, Y., 2013. Optimal deployment of public charging stations for plug-in hybrid electric vehicles. *Transportation Research Part B* 47, 87-101.

[24] Hodgson, M.J., 1990. A flow-capturing location-allocation model. *Geographical Analysis* 22(3), 270-279.
[25] Iglehart, D.L., 1965. Limiting diffusion approximations for the many server queue and the repairman problem. *Journal of Applied Probability* 2(2), 429-441.
[26] Jiang, Y., Zhang, Y., Zhang, C., Fan, J., 2012. The capacitated deviation-flow fueling location model for siting battery charging stations. In: Proc. *CICTP 2012*, ASCE, 2771-2778.
[27] Jung, J., 2012. Shared-ride passenger transportation systems with real-time routing. PhD Dissertation, University of California, Irvine, 186 pp.
[28] Jung, J., Jayakrishnan, R., 2012. High-coverage point-to-point transit: electric vehicle operations. *Transportation Research Record* 2287, 44-53.
[29] Jung, J., Jayakrishnan, R., Park, J.Y., 2013. Design and modeling of real-time shared-taxi dispatch algorithms. In: Proc. 92nd TRB Annual Meeting, Washington, DC.
[30] Jung, J., Jayakrishnan, R., Choi, K., 2013.Shared-Taxi Operations with Electric Vehicles. Working paper, UCI-ITS-WP-13-1, University of California, Irvine, 22p., http://www.its.uci.edu/its/publications/papers/ITS/UCI-ITS-WP-13-1.pdf
[31] Kameda, H., Mukai, N., 2011. Optimization of charging station placement by using taxi probe data for on-demand electrical bus system. In: Proc. KES 2011, Part III, LNAI 6883, 606-615.
[32] Kang, J.E., Recker, W.W., 2012. Strategic hydrogen refueling station locations with scheduling and routing considerations of individual vehicles. Working paper, UCI-ITS-WP-12-2, University of California, Irvine, 37p., http://www.its.uci.edu/its/publications/papers/ITS/UCI-ITS-WP-12-2.pdf
[33] Kang, J.E., Chow, J.Y.J., Recker, W.W., 2013. On activity-based network design problems. *Transportation Research Part B*, Special Issue on ISTTT 20, in press, doi: 10.1016/j.trb.2013.05.009.
[34] Kim, M., Schonfeld, P., 2012. Conventional, flexible, and variable-type bus services. *Journal of Transportation Engineering* 138(3), 263-273.
[35] Kim, J.G., Kuby, M., 2012. The deviation-flow refueling location model for optimizing a network of refueling stations. *International Journal of Hydrogen Energy* 37(6), 5406-5420.
[36] Kimura, T., 1983. Diffusion approximation for an M/G/m queue. *Operations Research* 31(2), 304-321.
[37] Koffman, D., 2004. *TCRP Synthesis 53: Operational Experiences with Flexible Transit Services*, Transit Cooperative Research Program, Transportation Research Board, Washington, DC, 71p.
[38] Kuby, M., Lim, S., 2005. The flow-refueling location problem for alternative-fuel vehicles. *Socio-Economic Planning Sciences* 39(2), 125-145.
[39] Kuby, M., Lim, S., 2007. Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs. *Networks and Spatial Economics* 7(2), 129-152.
[40] Laporte, G., Nobert, Y., Taillefer, S., 1988. Solving a family of multi-depot vehicle routing and location-routing problems. *Transportation Science* 22(3), 161-172.
[41] Larson, R., 1974. A hypercube queueing model for facility location and redistricting in urban emergency services. *Computers & Operations Research* 1(1), 67-95.
[42] Lau, J. K. S., Tham, C. K., & Luo, T. (2011). Participatory cyber physical system in public transport application. *4th IEEE International Conference on Utility and Cloud Computing*, 355-360.
[43] Lin, Z., Ogden, J., Fan, Y., Chen, C.W., 2008. The fuel-travel-back approach to hydrogen station siting. *International Journal of Hydrogen Energy* 33(12), 3096-3101.
[44] Lovell, D.J., Vlachou, K., Rabbani, T., Bayen, A., 2012. A diffusion approximation to a single airport queue. *Transportation Research Part C*, in press, doi: 10.1016/j.trc.2012.04.010.
[45] Mak, H.Y., Rong, Y., Shen, Z.J.M., 2012. Infrastructure planning for electric vehicles with battery swapping. Working paper, http://ssrn.com/abstract=2022651.
[46] Marianov, V., ReVelle, C., 1996. The queueing maximal availability location problem: a model for the siting of emergency vehicles. *European Journal of Operational Research* 93(1), 110-120.
[47] Marianov, V., Serra, D., 1998. Probabilistic, maximal covering location-allocation models for congested systems. *Journal of Regional Science* 38(3), 401-424.

[48] MirHassani, S.A., Ebrazi, R., 2012. A flexible reformulation of the refueling station location problem. *Transportation Science*, in press, doi: 10.1287/trsc.1120.0430.
[49] Mulley, C., Nelson, J.D., 2009. Flexible transport services: a new market opportunity for public transport. *Research in Transportation Economics* 25(1), 39-45.
[50] Nourbakhsh, S.M., Ouyang, Y., 2010. Optimal fueling strategies for locomotive fleets in railroad networks. *Transportation Research Part B* 44(8-9), 1104-1114.
[51] Nourbakhsh, S.M., Ouyang, Y., 2012. A structured flexible transit system for low demand areas. Transportation Research Part B 46(1), 204-216.
[52] Owen, S.H., Daskin, M.S., 1998. Strategic facility location: a review. *European Journal of Operational Research* 111(3), 423–447.
[53] Pasternack, B.A., Drezner, Z., 1998. A note on calculating steady state results for an M/M/k queueing system when the ratio of the arrival rate to the service rate is large. *Journal of Applied Mathematics & Decision Sciences* 2(2), 201-203.
[54] Perl, J., Daskin, M.S., 1985. A warehouse location-routing problem. *Transportation Research*, 19B(5), 381-396.
[55] Quadrifoglio, L., Li, X., 2009. A methodology to derive the critical demand density for designing and operating feeder transit services. *Transportation Research Part B* 43(10), 922-935.
[56] Recker, W.W. (1995). The household activity pattern problem: general formulation and solution. *Transportation Research* 29B(1), 61-77.
[57] Rolfe, A.J., 1971. A note on marginal allocation in multiple-server service systems. *Management Science* 17(9), 656-658.
[58] Snyder, L., 2006. Facility location under uncertainty: a review. *IIE Transactions* 38(7), 547–564.
[59] Stone, J. R., Nalevanko, A., Tsai, J., 1993. Assessment of software for computerized paratransit operations. *Transportation Research Record* 1378, 1-9.
[60] Tao, C.C., 2007. Dynamic taxi-sharing service using intelligent transportation system technologies. International Conference on Wireless Communications, Networking and Mobile Computing. Sept. 2007, 3209-3212.
[61] Tong, H.Y., Hung, W.T., Cheung, C.S., 2000. On-road motor vehicle emissions and fuel consumption in urban driving conditions. *Journal of the Air & Waste Management Association* 50(4), 543-554.
[62] Tsukada, N., Takada, K., 2005. Possibilities of the large-taxi dial-a-ride transit system utilizing GPS-AVM. *Journal of Eastern Asia Society for Transportation Studies*, EASTS, Bangkok, Thailand, 1-10.
[63] Upchurch, C., Kuby, M., Lim, S., 2009. A model for location of capacitated alternative-fuel stations. *Geographical Analysis* 41(1), 85-106.
[64] Upchurch, C., Kuby, M., 2010. Comparing the p-median and flow-refueling models for locating alternative-fuel stations. *Journal of Transport Geography* 18(6), 750-758.
[65] U.S. Environmental Protection Agency (2006). Greenhouse Gas Emissions from the U.S. Transportation Sector: 1990-2003. http://www.epa.gov/otaq/climate/420r06003.pdf.
[66] Wang, Y.W., Lin, C.C., 2009. Locating road-vehicle refueling stations. *Transportation Research Part E* 45(5), 821-829.
[67] Wang, Y.W., Wang, C.R., 2010. Locating passenger vehicle refueling stations. *Transportation Research Part E* 46(5), 791-801.
[68] Xi, X., Sioshansi, R., Marano, V., 2013. Simulation-optimization model for location of a public electric charging infrastructure. *Transportation Research Part D* 22, 60-69.
[69] Zhao, J., Dessouky, M., 2008. Service capacity design problems for mobility allowance shuttle transit systems. *Transportation Research Part B* 42(2), 135-146.