

Charging Network Optimization with Nonlinear Station Size Effects

Aaron Rabinowitz

1 Queuing at Stations

Because Battery Electric Vehicle (BEV) charging is relatively slow compared to Internal Combustion Engine Vehicle (ICEV) fueling, charging stations have less capacity per port than fueling stations. A fueling station dispensing gasoline at 7 gallons per minute energizes a vehicle at a rate of 14.15 [MW]. Current BEV can, often charge at a peak rate of several hundred [kW] given sufficient infrastructure but, even in this case, the rate is subject to exponential decay as State of Charge (SOC) increases meaning that the average rate is considerably lower. Consequently, queues are likely to form when demand is high. Queue waiting time is modeled using the M/M/c queuing formula. The expected waiting time in an M/M/c queue is computed as

$$W_q = f_q(\lambda, \mu, c) = \pi_0 \frac{\rho (c\rho)^c}{\lambda (1-\rho)^2 c!} \quad (1)$$

$$\pi_0 = \left[\left(\sum_{k=0}^{c-1} \frac{(c\rho)^k}{k!} \right) + \frac{(c\rho)^c}{c! (1-\rho)} \right] \quad (2)$$

$$\rho = \frac{\lambda}{c\mu} \quad (3)$$

where λ is the mean arrival frequency, μ is the mean service frequency, c is the number of homogeneous servers, ρ is the ratio of arrival frequency to composite maximum service completion frequency, and π_0 is the probability of an empty system. The M/M/c formulation assumes exponential distributions for λ and μ . One can think of ρ as equivalent to "utilization". Where ρ is low the station has excess capacity and where high the station approaches saturation. W_q approaches infinity as ρ approaches one.

Queue formation is a combinatorial effect. In order to maintain a stable queue length, the rate of vehicle arrivals must match the rate of vehicle departures. This is extremely unlikely over a short time interval and the length of the queue should fluctuate. Over a long enough period, a stable non-zero mean queue length can emerge even when the rate of arrivals is less than the theoretical maximum capacity of the station. This happens because the random sequence of arrivals and departures will, often, produce scenarios where all ports are occupied when the next vehicle arrives. There will be some times when there are no vehicles at the station at all and some times when a long queue exists. **Scenarios where all ports are occupied are more likely at stations with fewer ports even at identical utilization rates.** This effect is shown in Figure 1.

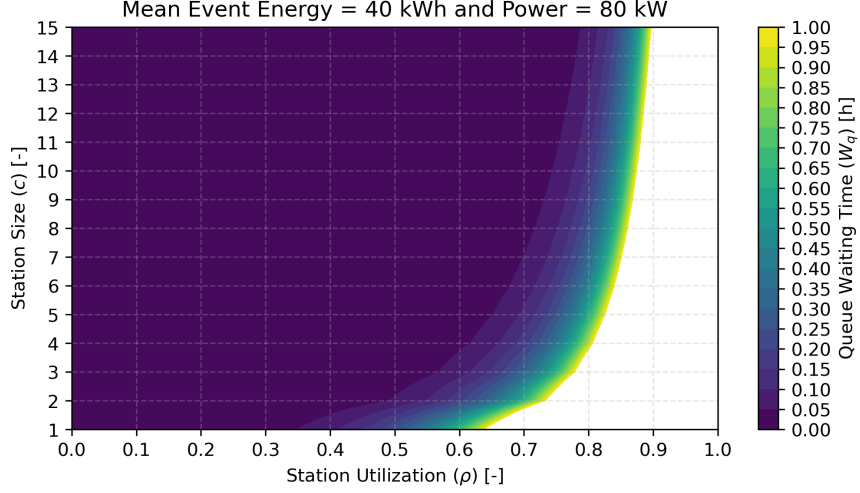


Figure 1: Queuing time with M/M/c queue model

Queuing dynamics mean that larger stations can handle higher utilization rates. As in Figure 1, a station with 15 ports can handle a rate of $\rho = 0.8$ before substantial queue formation equivalent to 24 vehicles per hour. A station of 4 chargers can handle a rate of $\rho = 0.6$ before substantial queue formation equivalent to 4.8 vehicles per hour. Thus, larger stations are more efficient on a capacity-per-charger basis with this effect being very substantial below 10 ports and leveling off after 15 ports given current values for mean energy dispensed and charging rate.

2 Supply Network Graphs

Consider a road network represented by a directed graph $G = \{V, E\}$ where V is a set of nodes and E is a set of edges. The set of nodes V contains places (cities and towns) at the nodes in $V_P \subseteq V$ and stations at the nodes in $V_S \subseteq V$. The set of edges E contains road links. Drivers may pass by places and stations and drivers who take the same road paths may not stop at the same stations. Thus, it is useful to transform the graph. Transformed graph $\hat{G} = \{V, \hat{E}\}$ contains the same nodes as G but the edges in \hat{E} are paths taken along the road network. \hat{E} contains edges for all O/D pairs $\langle i, j \rangle$ if the energy consumption required is less than the vehicle Energy Storage System (ESS) capacity. Routes along G will herein be called road-paths. Routes along \hat{G} will herein be called supply-paths. It is not possible for any edge $(i, j) \in \hat{E}$ to have negative cost. Thus, for routing purposes, only simple supply-paths need be considered. In order to facilitate in finding the k shortest simple supply-paths, for each origin $o \in V_P$, a level graph $\bar{G}_o = \{V, \bar{E}\}$ is constructed where $\bar{E} \subseteq \hat{E}$ and all edges $(i, j) \in \bar{E}$ lead away from o . Using \bar{G} the k shortest simple supply-paths are found using Yen's method.

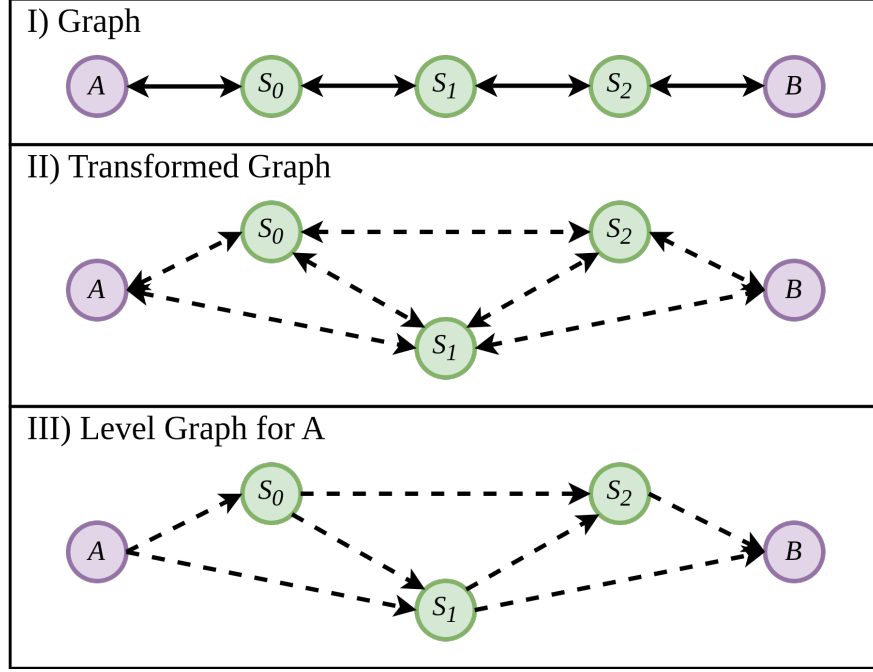


Figure 2: Graphs for simple vehicular transportation network. Panel I shows a graph G for a section of road from A to B with stations S_0 , S_1 , and S_2 . Panel II shows the transformed graph \hat{G} for a vehicle with sufficient range to reach S_1 from A and B . Panel III shows the level graph \bar{G}_A for the same vehicle leaving A . The simple supply-paths for pair $\langle A, B \rangle$ are $A - S_0 - S_1 - B$, $A - S_0 - S_2 - B$, $A - S_1 - B$, and $A - S_1 - S_2 - B$.

The time cost of edge $(i, j) \in \hat{E}$ is the time required to traverse the edge plus the trip time required to charge before traversing the edge. For a vehicle with a usable ESS capacity of β at node i which contains a charger whose maximum rate is α the time cost for edge (i, j) is

$$Y_{(i,j)}^T = \tau_{(i,j)} + \begin{cases} f_c(\varepsilon_{(i,j)}, \beta, \alpha) & \varepsilon_{(i,j)} \leq \beta \\ \infty & \varepsilon_{(i,j)} > \beta \end{cases} \quad (4)$$

where τ is edge traversal time, ε is edge energy consumption, and f_c is a function relating energy to charging time. Charging is modeled using a CC-CV relationship. The first part of charging is linear and the second part follows an exponential decay function. The time required for a given charge event is

$$f_c(\varepsilon, \beta, \alpha) = \begin{cases} \frac{\varepsilon}{\alpha} & \varepsilon \leq \eta\beta \\ \frac{\eta\beta}{\alpha} \left(1 - \ln \left(1 - \frac{\varepsilon - \eta\beta}{\beta(1-\eta)} \right) \right) & \varepsilon > \eta\beta \end{cases} \quad (5)$$

where η is the inflection point separating linear and nonlinear charging. A typical value for η will be in the range of 0.7 to 0.8. Charging past η will be substantially slower than below η .

3 Travel Time Minimization Formulation

The purpose of this formulation is to minimize total travel time in the system. Travel time minimization is accomplished by vehicle routing and charging station provision. For a given O/D pair there will, usually, be multiple viable charging paths of different lengths. As demand increases, chargers become increasingly congested leading to queuing time at stations. Queuing delays on shorter paths will push traffic to longer paths. Eventually, queuing will be sufficient to make the charging network no longer beneficial. This point is defined as when the marginal vehicle trip would take as much time using the network as it would take using level 1 charging. The goal is to place chargers and route vehicles to minimize total travel time. As such, each O/D pair has a "failure" flow which vehicles can be assigned to and the penalty assigned for this is equal to the travel time with level 1 charging.

Delay at stations is modeled using the outputs of the M/M/c queue formula as discussed in Section [REF]. Specifically, the outputs are linearized. Stations are initialized with a vector of m binary variables representing possible sizes (e.g. 1 charger, 3 chargers, 5 chargers). The vector of station size binary variables must sum to 1. For each size considered, the M/M/c model returns volumes and delays corresponding to a set of marginal utilization rates $R : |R| = n$. The utilization rates $\rho \in R$ are the bounds of a set of $n - 1$ utilization intervals. Thus, m by $n - 1$ matrices of marginal volumes and marginal delays are created. Additionally, a m by n matrix of unit-interval continuous variables are created and the sum of these variables multiplied by the corresponding marginal volumes must equal the flow passing through the station. the delay at the station is computed by summing the marginal utilization rates multiplied by the marginal delays.

- $G = \{V, E\}$: System graph containing nodes $v \in V$ and edges $(i, j) \in E$. Edge costs are defined by the following sets:
 - Y^T : The time required to traverse edge (i, j)
- $O \subseteq V$: Set of origin nodes
- $D \subseteq V$: Set of destination nodes
- $S \subseteq V$: Set of nodes with charging stations (or the possibility of a station). Stations provide energy to vehicle flows at a given rate. Depending on the utilization level of the station, vehicles may experience delay. The relationship between utilization and delay is linearized using the following sets:
 - C_s : Set of possible station sizes at station $s \in S$
 - $K_{s,c}$: Set of capacity intervals at station $s \in S$ for station size $c \in C_s$
 - Y^V : Set of volumes corresponding to each $c \in C_s$ and $k \in K_s$
 - Y^D : Set of delays corresponding to each $c \in C_s$ and $k \in K_s$
- \hat{C} : Maximum number of chargers which can be installed
- Q : Set of demand tuples of the form $\langle o, d, v, c, \hat{t} \rangle$ where o is the origin, d is the destination, v is the volume, c is the capacity of the ESS capacity of vehicles, and \hat{t} is the maximum travel time that is acceptable for the given demand.

- Y^Q : Set of time penalties for failing to accommodate flow. Set so that y_q^q is equal to \hat{t} in q .
- P : Set of paths corresponding to each demand $q \in Q$. Paths begin at $o \in O$ and end at $d \in D$. All intermediate nodes $i \in P \setminus \{o, d\}$ must be stations $s \in S$.
 - P^q : Paths that correspond to demand $q \in Q$
 - P^s : Paths that include station $s \in S$
- X : Set of continuous decision variables:
 - X^Q : Portion of demand flow not facilitated by the network
 - X^P : Flow volumes along paths
 - X^U : Portion of station capacity intervals utilized
 - X^V : Volume seen at station
 - X^D : Queuing delay seen at station
- U : Set of integer decision variables:
 - U^S : Booleans for station sizes corresponding to S and C

The objective of the optimization is

$$\min_{\bar{X}, \bar{U}} \underbrace{\sum_{q \in Q} x_q^q y_q^q}_{\text{Penalty Time}} + \underbrace{\sum_{q \in Q} \sum_{P^q \in P} \sum_{p \in P^q} \sum_{(i,j) \in p} x_p^p y_{(i,j)}^t}_{\text{Edge Traversal Time}} + \underbrace{\sum_{s \in S} \sum_{c \in C_s} \sum_{k \in K_{s,c}} u_{s,c}^s x_{s,c,k}^u y_{s,c,k}^d}_{\text{Queuing Time}} \quad (6)$$

subject to

$$Q[v] - x_v^q - \sum_{p \in P^q} x_p^p = 0 \quad \forall q \in Q \quad (7)$$

$$\sum_{p \in P^s} x_p^p - \sum_{c \in C_s} \sum_{k \in K_{s,c}} u_{s,c}^s x_{s,c,k}^u y_{s,c,k}^v = 0 \quad \forall s \in S \quad (8)$$

$$x_{s,c,k}^u - u_{s,c}^s \leq 0 \quad \forall s \in S, \forall c \in C_s, \forall k \in K_{s,c} \quad (9)$$

$$\sum_{c \in C_s} u_{s,c}^s - 1 = 0 \quad \forall s \in S \quad (10)$$

$$\sum_{s \in S} \sum_{c \in C_s} c u_{s,c}^s - \hat{C} \leq 0 \quad (11)$$

The objective function (6) minimizes total travel time in three terms. The first term is the time penalties accrued for failing to accommodate demand. The theory is that, without dedicated charging infrastructure, vehicles could, theoretically, complete the trip using level 1 charging but this would be very slow if the trip is beyond full-charge range. The second term is the time spent driving along edges and charging to drive along edges. The third term is the time spend queuing

for a charger. Constraint (7) forces the sum of flows and un-accommodated flows to be equal to total demand. Constraint (8) forces the sum of utilization intervals at a station to be equal to the sum of flows which pass through the station. Constraint (9) forces station utilization to only accrue for the selected station size. Constraint (10) forces only one station size to be selected per station and (11) limits the total number of chargers in the network.

4 Allowable Delay Formulation

The previous formulation adds complexity over common station location models by directly evaluating the time cost of queuing vs driving. This means that the objective function must be time with station provision as a constraint (and possibly outer-optimization objective). It is common in literature for the objective to be the number of cars or vehicle flows accommodated by the system. This problem is known as the Flow Refueling Problem (FRP). It is possible to integrate queuing dynamics into a FRP. In order to do this, one must define an acceptable amount of queuing delay at stations. The question then becomes: if a given level of queuing at stations is acceptable, where should chargers be placed in order to maximize accommodated flow? The relationship between station size and maximum allowable utilization rate for different levels of queue waiting time is shown in Figure 3.

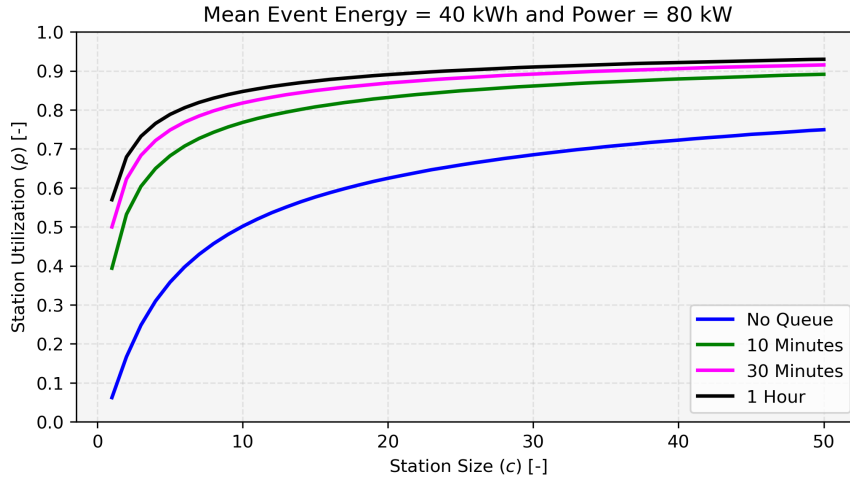


Figure 3: Maximum allowable utilization rates for stations of different sizes and different levels of allowable demand.

Substantial capacity gains are to be found in increasing the size of a station and in increasing the allowable delay at stations. However, in both cases, there are diminishing returns. Because time is not minimized in the FRP the selection of paths must be constrained. In this case, only paths that do not exceed a given multiple of shortest-path travel time are allowed. Path travel time is computed accounting for driving time, charging time, and assuming maximum allowable queuing delay at each station. Sets for the formulation are:

- $G = \{V, E\}$: System graph containing nodes $v \in V$ and edges $(i, j) \in E$.

- $O \subseteq V$: Set of origin nodes
- $D \subseteq V$: Set of destination nodes
- $S \subseteq V$: Set of nodes with charging stations (or the possibility of a station). Stations provide energy to vehicle flows at a given rate. The relationship between station size and allowable utilization is linearized using the following sets:
 - C_s : Set of possible station sizes at station $s \in S$
 - Y^p : Set of marginal allowable volume gains corresponding to each $c \in C_s$
- \hat{C} : Maximum number of chargers which can be installed
- Q : Set of demand tuples of the form $\langle o, d, v, c, \hat{t} \rangle$ where o is the origin, d is the destination, v is the volume, c is the capacity of the ESS capacity of vehicles, and \hat{t} is the maximum travel time that is acceptable for the given demand.
- P : Set of paths corresponding to each demand $q \in Q$. Paths begin at $o \in O$ and end at $d \in D$. All intermediate nodes $i \in P \setminus \{o, d\}$ must be stations $s \in S$.
 - P^q : Paths that correspond to demand $q \in Q$
 - P^s : Paths that include station $s \in S$
- X : Set of continuous decision variables:
 - X^Q : Portion of demand flow not facilitated by the network
 - X^P : Flow volumes along paths
 - X^V : Volume seen at station
- U : Set of integer decision variables:
 - U^S : Booleans for station sizes corresponding to S and C

$$\min_{\bar{X}, \bar{U}} \underbrace{\sum_{q \in Q} x_q^q}_{\text{Un-Accommodated Flow}} \quad (12)$$

subject to

$$Q[v] - x_q^q - \sum_{p \in P^q} x_p^p = 0 \quad \forall q \in Q \quad (13)$$

$$\sum_{p \in P^s} x_p^p - \sum_{c \in C_s} \sum_{k \in K_{s,c}} u_{s,c}^s y_{s,c}^p \leq 0 \quad \forall s \in S \quad (14)$$

$$\sum_{s \in S} \sum_{c \in C_s} u_{s,c}^s - \hat{C} \leq 0 \quad (15)$$

The objective function (12) minimizes un-accommodated. Constraint (13) forces the sum of flows and un-accommodated flows to be equal to total demand. Constraint (14) forces flow at all stations to be less than capacity. Constraint (15) limits the total number of chargers in the network. For large optimizations, an approximate solution can be obtained by relaxing the integrality constraint on U^S making the problem fully linear.