A Heuristic Algorithm for allocation and sizing of charging stations using the Capacitated Flow-Refueling Location Model

Wang Yue¹, Shi Jianmai¹*, Liu Zhong¹
¹Science and Technology on Information Systems Engineering Laboratory, National University of Defense Technology, Changsha, China
(wangyy0115@hotmail.com, jshi1980@163.com, liuzhong@nudt.edu.cn)

Abstract – Electric vehicle is one of the most promising solutions to decrease the use of fossil fuels. However many customs are reluctant to purchase EVs for lack of charging infrastructure making this a chicken and egg problem. In this paper, we add capacity constraint into flow-refueling location model and provide a model that fit to solve the allocation and sizing problem of charging station. Formulate the dividing behavior of EV flows by placing some common senses of drivers. Finally develop a heuristic algorithm to solve this problem. Apply this approach to solve a real world case of Hebei highway network. Get some useful conclusions of location and sizing of charging station.

Keywords - charging station; location problem; maximum flow; mixed-integer programming

1 Introduction

For the last few decades, fossil fuels are playing more and more necessary role in transportation sector, which brings serious environmental and political problems. Frequent appeared in Beijing and Hebei province, fog and haze has become a hotspot attracting widespread attention. Calls to reduce exhaust and protect the environment are on the rise. One major culprits of air pollution is vehicle emission, which brings about 30% of the atmospheric pollutant in China. As the number of fuel vehicles growing fast, the situation China faced is severe.

Battery Electric Vehicles(EVs) are regarded as one of the most promising solution of this problem. EVs use electricity, which produces less noise and emission than fuel vehicles. Meanwhile, the electricity can improve the efficiency of energy. However, the promotion of EVs also faces many obstacles. First of all, range anxiety. The driving range of a domestic EV should be 200-300km once fully charged and this number could be 1000km for a fuel car. That means EVs are not able to achieve long-term trip unless there are sufficient charging stations along the route. Meanwhile, there is no backup battery if that they are out of power in the halfway. For practicality, few people will buy a car just to meet the need of commute. That is why most EVs in China are used as taxis.

Besides, charging rate is another intractable issue. There are two ways to charge an EV, using 220V household power or high voltage piles. The former takes about 10 hours while the letter, about half an hour. Both will take a long time that cannot be ignored, which makes the capacity of a charging pile very limited. In other words, consumers will tend to choose EVs if there are sufficient facilities making them get charged conveniently and avoid wasting time in queue. However, charging pile is of high prices, while the maintenance will also consume much manpower and resources. This huge cost need to charge from the users, which makes it a chicken-and-egg problem. Given this, the location and size of charging station is very important and must be well optimized. The capacity of a charging station depends on the number of piles. So this problem can be changed into decision of where to locate the stations and how many piles there are.

This should be a difficult work because of two points. First, location and sizing must be considered together. Compared with conventional refuel station, the capacity of charging station is limited because of the long charging time which cannot be ignored. When discussing which demand should be satisfied, we must consider how much of them can be satisfied at the same time. The other is about the behavior of drivers. For long-term trip, there may be various chooses of stations to charge the whole way. We cannot force the drivers travel the route as we want, also the charging stations cannot refuse to serve the cars not in plan. How to describe the rules driver follows, that will be a hard work.

In this paper, we construct the behavior model of charging station and driver. Illustrated the rule they follow in the charging process. Give out the mathematical model of the location and sizing of charging station. Develop a heuristic algorithm to find an optimal solution of this problem. Finally, we apply this approach to solve a realistic case of Hebei highway network and find some useful conclusion to guide the construction of charging station.

2 Literature review

Facility location-allocation problem has been well studied in last century. The key goal of this problem is to maximize served demand or minimize cost. For different facilities, demand appears in various forms. Traditional facility location models like p-median, p-center, set cover and maximal cover focus on node-based demand. They all studied in last century. The key goal of this problem is to maximize served demand or minimize cost. For different facilities, demand appears in various forms. Traditional facility location models like p-median, p-center, set cover and maximal cover focus on node-based demand. They all consider service request can be aggregated in nodes of the network. The objectives of these models are to minimize the total transportation and construction cost, or to maximize the number of served demand points with finite budget. Transportation cost could be the weighted distance between demand nodes and the nearest facility, while construction cost include more aspects as number, size and location of facilities.

Lin et al. address an approach named “fuel-travel-back” when solving refueling station location problem. Driven

by the notion that “where you drive more is where you more likely need refueling”, they suppose the origin of a trip depending on the probability quantified by distribution of vehicle miles traveled (VMT). In this way, they concentrate the demand of a road segment into a node which makes it a p-median model. Albert proposes four solution methods to tackle node-based charging station location problem and also compare the complexity and effectiveness of the optimal solution.

However, charging demand is not always expressed at nodes in a road network. As for long-term trips, driver needs transport from one city to another and recharge several times during the whole journey. Hodgson (1990) offer a strategy to express this demand as traffic flow from origin to destination (O-D). He emphasizes that flow passes along links and through nodes in a network, so a facility is capturing instead of covering demand. With this perspective, he develops a new kind of model named as flow capturing location model (FCLM). Although FCLM treats flow as objective, no passing limit makes it similar to maximum cover problem. Kuby and Lim (2005) make an extension, they consider driving range restriction in FCLM, fixes valid flow only if one or more facility combinations can charge the whole trip, which avoids to double count flows captured by different facilities along the same path. This model they call it flow refueling location model (FRLM). As FCLM has already been NP hard problem, another constraint makes it more complex. First they pick out all the facility combinations of O-D pairs and remove supersets. Then develop a mixed-integer linear programming (MILP) formulation to determine which combination should be chosen. Kuby and Lim (2009) use this approach to solve realistic cases of Florida and Orlando. They found that the most difficult part of this approach is the number of combinations grows exponentially with the number of nodes in the network. As for Florida case with 302 candidate nodes, 74 O-D nodes, the computer (Pentium4, 3.2GHz, 1G RAM) took 13 hours only to generate the valid combinations for the first 39 paths of the total 2701 shortest paths. In order to make FRLM efficiently, Lim and Kuby (2010) developed some heuristic algorithms, greedy-adding, greedy-adding with substitution and genetic algorithm. The heuristic algorithms take about 10%-50% time of genetic takes and effectiveness of the optimal solution.

Wang et al. (2009) introduced another flow-based model using concept of set covering. The majority of assumption is the same as FRLM, but the objective is to minimize the cost while refueling all the flow demand. Given this point, O-D flow demand can be cut into arc in the network.

Upchurch and Kuby (2009) pointed that original FRLM assumes the refueling station is sufficient to serve all flows passing through, regardless of their volume. Actually the service ability of a pile is capacitated as the number of piles in a station is not fixed. So they develop a modified model called capacitated FRLM (CFRLM). CFRLM is not simply 0-1 MILP, the number of piles in a station, one of the decision variable, could be any nonnegative integer. As the complexity enhanced, the case study only included 25 O-Ds and 25 candidate nodes.

3 Model formulation

The basic hypothesis is same with FRLM, the road network consist of OD nodes, candidate nodes and arcs. First we try to describe how the flow distributes in network. Figure 1 illustrate a simple instance with five nodes, in which s, t are origin-destination (OD) nodes and A, B, C are candidate nodes of charging station. It assumes that vehicles with driving range of $R=100km$ want to travel from s to t. To avoid running out of power, there are two schemes of station combination, \{A, C\}, \{B, C\}. Of course no one will choose combination \{A, B, C\} for extra waiting time without necessary.

Fig. 1 A simple example of road

1) Behavior of drivers

First, a hypothesis about the power last in battery at OD is placed. As ODs in highway network always come as cities and towns, drivers can find 220V household power everywhere. If there is a long-term trip next day, they will always get well prepared for that. That means the battery is fully charged in OD when set off. Given this, we won’t consider round-trip constrictions as FRLM.

The amount of flow depends on two factors. One is the historical flow of fuel vehicles which can be reference to estimate the max demand of EVs. Also we can turn to population of ODs and distance between ODs instead if historical data is difficult to collect. The max flow can be fully or partly satisfied hinges on the service capacity of charging station it goes through. The actual demand of an OD pair is the sum of all combinations can charge.

For combinations of the same OD pair, just as \{A, C\}, \{B, C\} in Fig. 1. Drivers may choose any station they want, so that every station shares equal opportunity. If one combination is overloaded, the remaining flow spread to other combinations until all the flow are charged or no station of the combination has free capacity, then the remaining flow will be cut down.

2) Flow distribution

As we mentioned before, stations cannot refuse to serve customs no matter where they come. Drivers always tend to attend the station that is not too busy. Given this perspective, we describe the flow distributing process with following rules. In order to illustrate clearly, here we indicate a flow with pair like ((O, D), combination, Amount), the unit of amount is (cars/day).

Rule 1, all the flows charged by one station will meet same discount rate if overload occurs. When there are many flows charged in one station, more potential demand means getting more charging source. But as the drivers have the same opportunity to get charged, the proportion of discount will be similar too if overload. Figure 2 is a simple instance with two flows, ((s, t), \{A\},
300) and ((u, v), {A}, 200). The capability of A is 400. There are 100 cars overloaded, so the actual charged amounts of two flows will be 240 and 160 because they both suffer 20% discount.

Rule 2, how many cars of flow will get through the path depends on the highest discount in the combination. Vehicles will get through the path only if all the stations in the combination are available. So the capability of bottleneck point will decide the passing amount. Figure 3 shows an extended example with two stations A, B whose capability are 400 and 150. The combination of (s, t) becomes {A, B}. The discount rates caused by A is 20%, but station B is more congested which is 50%. So ((s, t), {A, B}, 300) only gets 150 passed while ((u, v), {A}, 200) gets totally charged as there comes more free capacity of station A.

3) Main model

Take the rules above into consideration, we can get the main model of this problem. The main model is based on the FRLM with some basic changes. The formulation of the main model is defined as follows:

Objective:

Max \[ Z = \sum_{q \in Q} \sum_{h \in H_{qh}} f_{qh} w_{qh} \] (1)

Subject to:

\[ \sum_{q \in Q} \sum_{h \in H_{qh}} f_{qh} v_{h} \leq l x_{k} \quad \forall k \in K \] (2)

\[ a_{bh} y_{h} \geq v_{h} \quad \forall h \in H_{h} \quad k \quad a_{bh} = 1 \] (3)

\[ \sum_{k \in K} (m_{k} x_{k} + n_{k} y_{k}) \leq C \] (4)

\[ E y_{h} \geq x_{k} \quad \forall k \in K \] (5)

\[ v_{h}, y_{h} \in \{0,1\} \quad \forall h, q \] (6)

\[ x_{h} \in \{\text{nonnegative integer}\} \] (7)

Where the following notation is the same as in the FRLM:

\( q \) = index of O-D pairs
\( Q \) = set of all O-D pairs
\( f_{q} \) = flow volume on the shortest path between O-D pair \( q \)
\( k \) = potential station location
\( K \) = set of all potential station locations
\( h \) = index of combinations of stations
\( H \) = set of all potential station combinations
\( a_{bh} \) = a coefficient equal to 1 if station \( k \) is in combination \( h \) and 0 otherwise
\( b_{qh} \) = a coefficient equal to 1 if station combination \( h \) can refuel O-D pair \( q \) and 0 otherwise
\( v_{h} \) = 1 if all stations in combination \( h \) are open, 0 otherwise

and the remaining notation is modified in our main model:

\( x_{k} \) = the number of piles located at \( k \)
\( y_{k} \) = 1 if a station is located at \( k \), 0 otherwise
\( w_{qh} \) = real proportion of \( f_{q} \) charged by station combination \( h \)
\( l \) = charging ability of a pile
\( C \) = budget
\( E \) = a sufficiently large integer
\( n_{k} \) = construction fee of infrastructure in node \( k \)
\( m_{k} \) = cost of a charging pile in node \( k \)

The objective function (1) is to maximize the valid flow with limited budget. Constraint (2) is about service capacity, here we measure the capacity of a station by the number of piles. Constraint (3) implies only if all the stations of combination are available that can make flow go through. Constraint (4) is cost limit, the cost is separated into construction and infrastructure. Construction cost for a station will not change too much with the size while the other part is more sensitive. Here we simplify the dilatation cost grow with the number of piles. Constraint (5) ensures that piles must located in stations. Constraints (6) and (7) are the integrality requirements for the variables.

Different with CFRLM, \( w_{qh} \) is not a decision variable. Its value varies with the change of location and size of
stations, i.e., once the scheme is confirmed, the $w_{ij}$ is fixed. That makes the model a non-classical optimization model.

4 Algorithm

The objective of this problem is to maximize flow charged by stations. For fixed budget, optimal solution often take the most use of money, because the rest budget can cover extra demand and makes it a better solution. Therefore, the objective can be transited into maximizing the volume-cost ratio for the total budget. The overall optimization is broken down into seeking set of high volume-cost ratio combination of stations.

There are two common senses to achieve high volume-cost ratio: 1) short-term trip takes less charging source than long ones; 2) larger station decreases the construction proportion which cannot be transformed into charging capacity. Given this, solutions with fewer stations and more small-scale combinations will be more optimal. So we develop a dynamic process based on deep-search to seek such solutions.

The main idea of the algorithm is to combine the combinations with a higher volume-cost ratio into larger combinations, until there is some combination that can consume entire budget. First we list all the 1-node combinations and put them into alternative set, sort them with the volume-cost rate. This is to guarantee that every node will be considered. Then use the combinations in alternative set to construct 2-node combinations. Add first N 2-node combinations with higher ratio in alternative set and sort again. Repeat this process to construct larger combinations. This approach can be described in 2 steps below:

Step 1: find the termination size of searching. The termination condition is closely related to the time and the final outcome. Large size will cause unacceptable calculating time, while small size may miss better solutions. Here we set the max size by finding an initial feasible solution. We sort all the candidate nodes by ratio. Each time add one node from the first into the temporary set, until the temporary combination can consume the entire budget. Then the temporary combination is the initial solution we want, and the size of the combination is the max searching size. The initial combination is an intuitively optimal solution. However there is always some repeated demand in the first few nodes. The combination with more nodes will get more construction fees and long-term demand, which means lower ratio. Besides some extreme situations, the initial solution should be a good one, but not the best one. Set the max size in this way ensure we can find solutions better than the initial one. The pseudo-code of this process is showed below:

---

**Step 1:** Find the initial feasible solution and max searching size

**Input:** candidate point set $P$; total budget $B$;
Alternative set $A$;
for ($i=1;i<=|P|;i++)$
  add(combination($p[i]$)); //add 1-node combination into alternative set
sort(A); //sort the combinations in alternative set with the volume-cost ratio
for ($i=1;i<=|P|;i++)$
  tempcombination.addnode($p[i]$); // add node into temple combination
if(tempcombination.cost == budget){
  maxsize = i;
  break;
}
return maxsize;

---

Step 2: construct larger combinations and broaden the alternative set. The major idea of this approach is trying to combine high ratio combinations into feasible solutions. However the way of deep search is very sensitive to the size, the complexity will suffer combination explosion. For large-scale network, we cannot even get all the 5-node combinations in acceptable time. We add a pruning to avoid redundant search. Record the minimal ratio of each size combination in the alternative set. For every step of combine, compare current ratio with the corresponding minimal ratio. If the current ratio is less, stop further combine. This pruning will of course lose some possibility of better solution, but it will decrease searching time effectively. The pseudo-code of this process is showed below:

**Step 2:** Construct larger combination

**Input:** alternative set $A$;
  target combination size $s$;
  step adding size $N$;
  minimal ratio of $i$-node combinations in alternative set $T[i]$;
depsearch(cursize, curloc) {} //cursize presents the current size of temp, curloc presents the location in the alternative set
if(cursize == s){
  tempA.add(temp); //add new generated combination into temple set
} else{
  if(tempcombination.ratio < T[temp.size])
    return;
  for(i=curloc+1;i<=|A|;i++)
    tempcombination.addnode(A[i]); // keep the first cursize nodes and add all the non-repeating nodes in the A[i]
  if(tempcombination.size <= s){
    if(tempcombination.ratio >= maxflow)
      return;
  }
}

---
The pseudo-code of entire algorithm is as follows:

**Entire Heuristic Algorithm**

**Input:** step adding size $N$;
- alternative set $A$;
- target combination size $s$;
- candidate node set $P$;

```c
for (i = 1; i <= |P|; i++) {  // add all the 1-node combinations into the alternative set
    new combination.addnode(P[i], 0);
    A.add(combination);
}
sort(A);
for (i = 1; i <= maxsize; i++) {  
    s = i;
    depsearch(0, 0);  // construct i-node combinations
    sort(tempA);
    for (j = 1; j <= N; j++) {  
        A.add(tempA[i]);  // add first N i-node combinations with highest ratio into A
    }
    sort(A);
    tempA.clear();  // remove all the combinations in tempA
}
return A.maxcombination();
```

5 Case study

This case study uses a simplified version of hebei highway network shown in Fig 4. The network consists of 58 entrances and 116 rest areas which can be seen as Origin-Destinations and candidate nodes of charging station. The entrances are chosen with large flow demand of fuel vehicles judged by historical traffic situation. The rest areas are the large ones which are appropriate to construct charging stations. The flow demand is predicted with the real fuel vehicle flow data in a week. Because of a lack of data, all the external flows are counted on the edge nodes. This change won’t bring much difference as only the internal portion is valid in this problem.

The driving range of the EV is assumed to be 150 km while most products have more than 200 km technical range. This is a prudent limit in case of traffic jams, upslope, fatigue of battery and any other factor that may impact the range. So the flow with distance below 150km won’t be considered as there is no need to recharge in the half way.

Unlike Upchurch’s research, the candidate nodes are real existed. However, the candidate nodes are not well distributed. In general, two adjacent rest areas will be 50 to 70 km apart, some are very closed to entrance, even at same place. These may be some large cities or scenic spots that drivers need to recharge when they get off the highway.

The capacity of charging pile is supposed to be 40 vehicle per day, as one vehicle will take about half an hour to be fully charged. The flow demands are unidirectional, charging station can serve all the vehicles go through it if not overloaded.

Fig 5 compares the optimum solutions and computing time of genetic and heuristic algorithms with different budget. The y axis presents the met proportion of total demand achieved by the best solution of each algorithm. The heuristic algorithm preforms nearly as well as genetic algorithm, even better in budget of 300 million. As for any scale of budget, heuristic algorithm takes less than 2 seconds to get the solution. It is about 1 percent of the time genetic takes, so the curve of heuristic appears like a straight line on the x axis. The detail times heuristic algorithm used are shown in Fig 6.
We can find in Fig 6 that the solving time does not increase linearly with budget. The search band is depend on the rank of initial feasible solution. The ranks of initial feasible solution are equal with each other for budget of 250 to 400 million. So there is not much difference in the solving time.

The Fig 7 shows the total volumes of demand flow go through each point. Then compare it with the optimal solution of 550 million. There are some nodes share nearly little or even no flow, which means no long-term flows go through but the nodes are wasted. As shown in Fig , the solution of genetic algorithm can cover above 99% of the total demand. The total flow only expresses the how much flow goes throw the nodes and there is much overlap in these demands. When the demands of one node are fitted, the other nodes related with these demands will decrease these volumes in their total demands. That is why the solution seems only covering little part of the whole. The optimal solution does not appear in the nodes with largest volume of flow demand, so it is not so apparent that can be easily achieved by greedily choose.

Fig 8 shows the optimal solution with budget of 550 million. The nodes chosen are separated well, which is much similar with the real situation. The EV flow from any direction can find charging station before it is out of power. There is no station in the north-east part of the network. As we mentioned, the network is about the highway charged by Hebei province. There is another part from Beijing going through north-east part of Hebei, which is charged by Beijing. That part is the arterial road sharing the majority of the vehicle flow. So the north-east part of Hebei highway network shares a little flow, especially long-term flow.

6 Conclusion

As shown in result, heuristic algorithm can get nearly the same result as genetic algorithm. The optima seems to be valid as it is appeared in network. In the future work, we will test the efficiency and results in random road network of different scale. After that, we will add some reasonable assumption of drivers’ behavior, considering queuing and gaming, to make the model more realistic.

REFERENCES


