Installing Electric Vehicle Charging Stations City-Scale: How Many and Where?

Marjan Momtazpour, Mohammad C. Bozchalui, Naren Ramakrishnan and Ratnesh Sharma

Abstract Electric Vehicles (EVs) are touted as the sustainable alternative to reduce our over-reliance on fossil fuels and stem our excessive carbon emissions. As the use of EVs becomes more widespread, planners in large metropolitan areas have begun thinking about the design and installation of charging stations city-wide. Unlike gas-based vehicles, EV charging requires a significant amount of time and must be done more periodically, after relatively shorter distances. We describe a KDD framework to plan the design and deployment of EV charging stations over a city. In particular, we study this problem from the economic viewpoint of the EV charging station owners. Our framework integrates user route trajectories, owner characteristics, electricity load patterns, and economic imperatives in a coordinated clustering framework to optimize the locations of stations and assignment of user trajectories to (nearby) stations. Using a dataset involving over a million individual movement patterns, we illustrate how our framework can answer many important questions about EV charging station deployment and profitability.

M. Momtazpour (✉) · N. Ramakrishnan
Department of Computer Science, Virginia Tech, Blacksburg, VA 24060, USA
e-mail: marjan@cs.vt.edu

N. Ramakrishnan
e-mail: naren@cs.vt.edu

M.C. Bozchalui · R. Sharma
NEC Laboratories America, Inc, Cupertino, CA 95014, USA
e-mail: mohammad@nec-labs.com

R. Sharma
e-mail: ratnesh@nec-labs.com

J. Lässig et al. (eds.), *Computational Sustainability*,
Studies in Computational Intelligence 645,
DOI 10.1007/978-3-319-31858-5_8

1 Introduction

In the last decade, electric vehicles (EVs) have been considered a promising solution for some environmental and economical issues. Fast decline of fossil fuels and global warming have increased the interest of policy makers in developed countries to use sustainable approaches to energy production, distribution, and consumption [1]; EVs have been touted for their potential to dramatically reduce fossil fuel consumption and CO_2 emissions [2].

To operationalize and encourage EV usage, charging stations should be installed in multiple areas of a city. In large metropolitan areas with a significant number of EVs, charging stations must be installed in carefully selected locations. As a matter of fact, charging an EV is different from refueling a traditional gas-based car: EV charging takes much longer and places a significant amount of load on the electric grid [2]. Furthermore, compared to traditional cars, EVs must be recharged after relatively shorter distances. Proper placement of charging stations can result in optimal distribution of electricity load, maximization of revenue of service providers, and lead to increased availability of charging stations, and reduced range anxiety.

While charging station placement is an important task for EV deployment in urban areas, there is a relatively small number of prior research in this area (e.g., see [3–5], and our own work [6]) and all aim to locate charging stations to maximize the meeting of demands. In a comprehensive planning effort, however, it is crucial to consider economic factors in design of charging infrastructure for EVs to ensure financial feasibility as well as long-term economic growth. Various business models can be considered for EV charging station infrastructure, and in fact, EV charging infrastructure installation will be driven by models that reflect the economic benefits on top of policy objectives.

In our previous work [6], we propose a solution for charging station placement problem without specific assignment of EVs to charging stations. In [6], we developed a coordinated clustering formulation to identify a set of locations that can be considered as the best candidates for charging stations. The locations were determined to be those that have a low electricity load, and where a significant number of EV owners spend a considerable duration of time. The drawback of our proposed method in [6] is that it did not consider a concrete economic model for charging station placement. In fact, in [6], charging stations are placed based on the stay points of EV owners and distribution of electricity consumption in the city. Therefore, that approach may result in placements which are economically sub-optimal. Furthermore, in [6], we did not consider the trajectory of EVs, which in turn results in unacceptable detours.

In this book chapter, we propose a new integrated framework where the centralized assignment of EVs is addressed simultaneously with the charging station placement problem. This integrated framework solves an optimization problem that simultaneously considers revenue of charging station owners and the trajectory of EV owners. In this work, an economic model is formulated that takes into account the costs and benefits of installing and operating charging stations from their owner's perspective. In this model, charging station owners provide infrastructure, and own and operate EV charging stations. No extra incentive is considered for the charging station owners and it is assumed that they will be charged the same rates as other mid-sized commercial customers for buying electricity from the utility company. Charging stations sell electricity to EV owners at a fixed, flat rate. Furthermore, we use trajectory mining to find routes that could host popular locations where EV owners might desire to recharge their cars. We applied trajectory clustering on this dataset which helps us to install charging stations proximal to high-traffic roads, in order to reduce possible detours to reach charging stations. The trajectory of each individual in a typical day is derived through the use of APIs such as Google Maps. The results of this step are integrated to our final optimization equation to situate charging stations near high-traffic roads in order to reduce possible detours to reach charging stations. Finally, the economical model, results of trajectory mining, and information about each individual driving path of EV owners are fed into an integrated optimization problem. This optimization problem attempts to maximize the revenue of charging station owners, minimize distances of charging stations to high-traffic routes, minimize distances of charging stations to stay points of EVs, and minimize number of failure to find an appropriate charging station for an EV. Furthermore, using KL-Divergence, the optimization problem tries to place charging stations in a way that results in a uniform distribution of charging assignments.

We outline a KDD framework, involving coordinated clustering, to design and deploy EV charging stations over a city. Our key contributions are:

1. An integration of diverse datasets, including synthetic populations (capturing over 1.5 million individuals), their profiles, and trajectories of driving, to inform the choice of locations that are most promising for EV charging station placement. We solve the 'How many?' and 'Where?' problem using a coordinated clustering framework that integrates multiple considerations. We focus on the modeling of downtown areas since previous studies have shown that public EV charging infrastructures should be focused on big urban centers [2]. We use trajectory mining to detect popular roads EV owners are likely to use when they need to recharge their vehicles, and integrate this information in charging station deployment. In particular, our framework situates charging stations near high-traffic roads in order to reduce possible detours to reach charging stations.
2. Unlike our prior work [6], we formulate the EV charging station placement problem in both economic and user terms: the financial benefits to an EV charging station owner and the convenience benefits to EV owners are integrated into our framework. Empirical results reveal key distinctions between taking economic factors into account versus otherwise.
3. We conduct extensive empirical investigations into the practical feasibility of EV charging station placement w.r.t. multiple considerations: e.g., how many users in a population are serviced, how effectively are stations utilized, differences among varying types of charging infrastructure, and the need for storage units in charging stations.

2 Related Work

Charging infrastructure design: Relevant prior work in this area include [3–7]. Frade et al. developed a maximum covering model to locate charging stations to maximize demand [3]. In [5], a two-step model is proposed to create demand clusters by hierarchical clustering, then a simple assignment strategy is used to assign charging stations to demand clusters. In [4], a game-theoretic approach is used to investigate interactions among availability of public charging and route choices of EVs. In our prior work [6], we developed a coordinated clustering formulation to identify a set of locations that can be considered as the best candidates for charging station placement. The locations were determined to be those that have a low existing load, and where a significant number of EV owners spend a considerable duration of time. In [7], behavioral models are developed to predict when and where vehicles are likely to be parked, and aims to reflect parking demands in the optimization assignment.

Interactions with the smart grid: In addition to the problem of charging station placement, EV penetration in urban areas has been explored with respect to interactions between grid infrastructure and urban populations. City behavior is simulated by agent-based systems in terms of agents with a view toward having decentralized systems and maximizing profits [1]. Swanson et al. in [8] investigated the use of linear discriminant analysis (LDA) in assessing the probable level of EV adoption. Energy storage systems, systems that are used when there is not enough power available from grid, are addressed in [9]. In [10], a solution is proposed to balance energy production against its consumption. In addition, authors in [11] try to design a general architecture in smart grid to have a significant gains in net cost/profit with particular emphasis on electric vehicles.

Mobility modeling: There are many studies that consider mobility of vehicles in urban areas and in most of the cases, GPS datasets have been used as a popular source for modeling and mining in urban computing contexts, e.g., [12–14]. Example applications include anomaly detection [12] and taxi recommender systems [14]. In taxi recommender systems in particular [14], the ultimate goal is to maximize taxi-driver profits and minimize passengers' waiting times. Mining mobility patterns of cars and people has been used to determine points of interest for tourists [15] and for routing and route recommendation [16]. In [13], Yuan et al. proposed a method to discover areas with different functionalities based on people movements. Finally, in [17], clusters of moving objects in a noisy stadium environment are detected using the DBSCAN algorithm [18].

To the best of our knowledge, the problem tackled in this paper is unique, and the methodology we propose integrates a variety of data sources with data mining/optimization techniques.

3 Methodology

The datasets utilized by our approach and the overall methodology are depicted in Fig. 1. As shown, one of the primary datasets we consider is a synthetic population dataset representing the city of Portland which contains details of 1,615,860 people and 243,423 locations out of which 1,779 are located in the downtown area. Detailed information about this dataset is available at [19]. Next, information about mobility of people is provided in terms of start and end points and time of travel. Using this information, we can determine the trajectory of every individual in a typical day through the use of APIs such as Google Maps. A total of 8,922,359 movements are available in this dataset. Finally, we have available electricity consumption data to determine the initial load of each building based on the number of residents of the building at a specified time (organized by NEC Labs, America).

The first step of our methodology is to discover location functionalities and to characterize electricity loads. As in our previous work [6], we utilize an information bottleneck type approach [20] to characterize locations and integrated the electricity load information to characterize usage patterns across locations. In this step, we cluster locations based on geographical proximity such that resulting clusters are highly informative of location functionalities. Then, we integrate information about electricity load profiles to characterize electricity usage patterns.

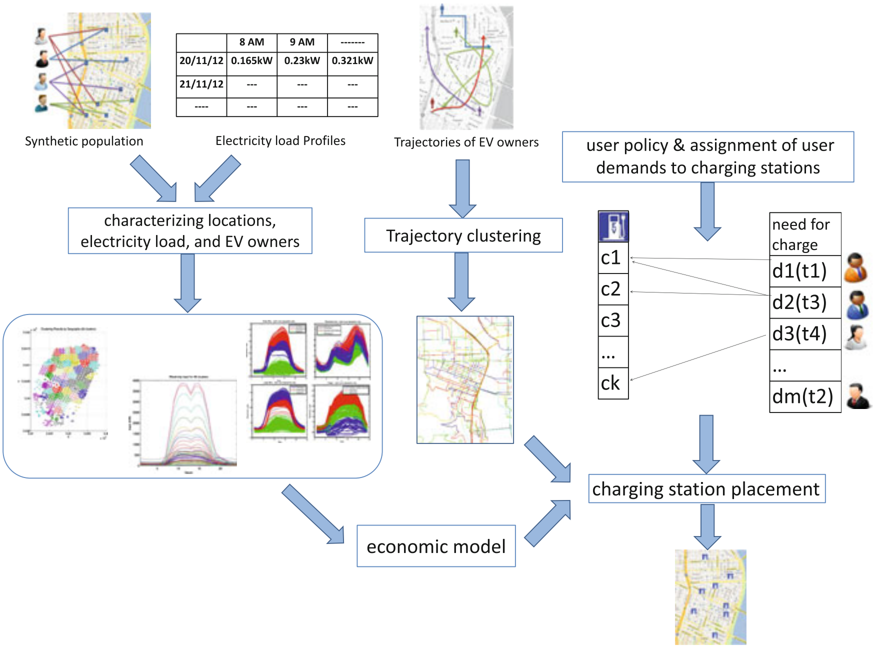


Fig. 1 Overview of our methodology

Second, we use trajectory clustering to find routes that could host popular locations where EV owners might desire to recharge their cars. To determine trajectories, we define a specific subset of people who are characterized using high-income attributes (as the likely owners of EVs). After locating the homes of these users, we can determine their trajectories and their start/stop locations. Based on this data, we can estimate their travel distances, and in turn estimate charging requirements of EVs, during a day. Since the maximum distance that a fully charged EV can travel is less than 100 km [2], it is highly likely that a significant number of them will need to be recharged en-route to their destinations. By clustering the trajectories, we can plan to install charging stations proximal to high-traffic roads, in order to reduce possible detours to reach charging stations.

Third, we develop an economic model that encapsulates costs of purchasing energy from the grid and other such expenses. Fourth, we identify demands based on our expectations about how users will behave. Finally, all this information provides the raw material for defining the charging station placement problem using clustering and optimization. Each of these stages are detailed next.

3.1 Characterizing Locations and EV Owners

The first step of the proposed data mining approach for EV infrastructure design is to characterize EV users and locations in the respective area. This step is similar to that in our previous work [6] and, hence, we provide an abridged summary of it.

Based on current trends, only a small percentage of people (6 % of people in the US) use EVs [21, 22]; in our study we explored a hypothetical scenario that considers a penetration for EVs in the Portland area to be 6.31 % of 329,218 people in our dataset. This assumption is realistic if various penetration scenarios in forecasted EV adoption between years 2012 and 2022 are to be believed [23], and can be easily modified.

From the synthetic population dataset, we can identify the locations a person visits, the duration of stay at each (stay points), and the purpose of the visit (e.g., work, leisure). We first begin by characterizing location with a view toward defining the specific purpose of the location. We focus on 1779 locations in the downtown Portland area whose attributes are given by a 9-length profile vector $P = [p_1, p_2, \ldots, p_9]$, where p_i is the number of travels incident on that location for the ith purpose. Specific purposes of each location (and cluster of nearby locations) can be used to determine electricity load distribution patterns. To uncover such patterns, we first cluster locations geographically and then characterize each of the discovered clusters using typical data available from public data sources such as the California End User Survey (CEUS). In addition to these kinds of patterns, we compute the electricity load leveraging the these patterns but w.r.t. our network model of the urban environment (by considering the average square footage occupied by one person in each specific location). Based on some exploratory data analysis, we selected a weekday (specifically, 18th March, 2011) and used the electricity load data of this day to map to the network model. More details are available at [6].

3.2 Trajectory Mining

The emergence of GPS-equipped devices has sprung a veritable cottage industry in the area of location and trajectory mining. One broad aim of trajectory mining (clustering) is to find similar routes in a dataset, but other applications have also been explored (e.g., see [24–26]). Most research in trajectory clustering is inspired by density-based clustering approaches such as DBSCAN and OPTICS. Leveraging such clustering methods, the authors in [27] propose a new framework (Traclus) for trajectory clustering which aims to discover common sub-trajectories. In applications where we have regions of special interest, finding common sub-trajectories is beneficial. In Traclus, each trajectory is partitioned into a set of line segments. Then, similar line segments are grouped together to form clusters of sub-trajectories [27]. This method has been proven to be effective in extracting similar portions of trajectories even when the whole trajectories are not similar. We employ this approach here to detect potential sub-routes where EV owners are more likely to travel and, thus, in need of charging.

3.3 Economic Model for Profit Maximization

The principal goal of this paper is to place charging stations in appropriate locations in order to maximize profits of charging station owners. Tran et al. [28] have studied cooperation of companies for profit maximization in dynamic systems. They have used regression and hierarchical agglomerative clustering to reveal optimal organizational substructures. Such approaches are not applicable here since they assume the locations (of markets) to be known and place restrictive assumptions on pricing schema.

The primary goal of charging station owners is to maximize revenue and profits by attracting enough customers during a day. The profit is defined as the difference between expense and income (revenue). Let us assume that R_i is the profit for charging station (location) i, which is the difference between the payments that the charging station owner will receive (S_i) and the costs spent on providing service to customers (C_i), as shown in Eq. 1. As Eq. 2 illustrates, C_i in turn consists of two elements: the static costs C_0, and the dynamic costs C_p.

$$R_i = S_i - C_i \tag{1}$$

$$C_i = C_p + C_0 \tag{2}$$

Static cost, C_0, is the initial cost for setting up a charging station, which includes the operational cost for installation and for storage units. Here, we calculate these costs for a single day, and thus assume an amortization function that estimates the installation cost for one day (e.g. if the installation cost will amortize in six years: $C_0 = \frac{InstallationCost}{6\times 365}$). Dynamic cost C_p, is the cost for the energy that the charging station will buy from the grid in order to service EV owners. The dynamic cost C_p, consists of two parameters: the cost of buying energy from the grid during the day (morning to evening), and the costs associated with recharging storage (during the night). These two parameters are denoted by C_b and C_r, respectively.

$$C_p = C_r + C_b \tag{3}$$

3.3.1 Calculating C_r

C_r is the payment that a charging station owner pays for charging the storage during the night (if needed). Typically, storage will be charged at night and used during the day and it should be sized to cover a day's net load.

$$C_r = P_{buy,night} \times StorageSize \tag{4}$$

where $P_{buy,night}$ is the price of off-peak hours that storage owner will pay to recharge the storage. $StorageSize$ is calculated through the following steps.

Suppose f is the load of building after considering EVs. Thus f is the Initial load of building $InitLoad_{i,t}$, and the load imposed by EVs. $Dload_{d,t}$ is the amount of electricity needed for user d at time t and n_i is the number of EVs receiving service by charging station i during a day.

$$f(t) = InitLoad_{i,t} + \sum_{d=1}^{n_i}(Dload_{d,t}) \tag{5}$$

In order to calculate the amount of storage for a particular charging station, we must calculate the number of EVs serviced by this charging station at each particular hour n_i. Here, we assume capacity of each building is constant and equals the maximum value of load of the building before introducing EVs:

$$capacity_i = \max_{0\le t\le 24} InitLoad_i \tag{6}$$

The size of required storage should be calculated from the area below the curve of new electricity load (that is f) (kW $\times$ h) and above the capacity (net peak load)(kW). X is the difference of load after EVs and capacity of building. Clearly, $StorageSize$ is a summation of X over time:

$$X(t) = \begin{cases} f(t) - capacity_i & \text{if } f(t) > capacity_i \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

$$StorageSize = \int_0^{24} (X(t))\, \mathrm{d}t \tag{8}$$

3.3.2 Calculating C_b

Calculation of C_b consists of three elements: Basic charges, Energy charges, and Demand charges [29].

$$C_b = C_{basic} + C_{energy} + C_{demand} \tag{9}$$

Basic charges is a constant charge ($240 per month[1]) [29] and energy charges is a multiplication of energy purchased at time t (kWh) in TOU rate at time t ($/kWh):

$$C_{energy} = \int_0^{24} (Y(t) - InitLoad_{i,t}) \times P_{buy,t}\, \mathrm{d}t \tag{10}$$

where $P_{buy,t}$ is determined based on time of the day (TOU rate) and Y is the amount of load of a building when storage is placed:

$$Y(t) = \begin{cases} capacity_i & \text{if } f(t) > capacity_i \\ f(t) & \text{Otherwise} \end{cases} \tag{11}$$

Demand charges involves facility capacity charges and on-peak demand charges:

$$C_{demand} = C_{FC}/30 + C_{OnPeak\ demand} \tag{12}$$

Facility capacity charges (C_{FC}) for one month is calculated in Eq. 13 [29]:

$$C_{FC} = \begin{cases} capacity_i \times 2.41 & \text{if } capacity_i \leq 200 \\ 482 + (capacity_i - 200) \times 2.14 & \text{otherwise} \end{cases} \tag{13}$$

On-peak demand charges is the maximum on-peak demand of the charging station (in kW) times per kW monthly on-peak demand rates ($/kW):

$$C_{OnPeak\,demand} = \max_{OnPeak\ t} Y(t) \times 2.67 \tag{14}$$

[1] In this paper all rates are in US dollar.

3.3.3 Calculating S_i

The income of charging station owner is calculated based on the summation of energies that he sells to EV owners over a day:

$$S_i = \sum_{d=1}^{n_i} \int_{t=0}^{24} (P_{sell,t} \times Dload_{d,t})\mathrm{d}t \tag{15}$$

where P_{sell} is price per kW.

The ultimate goal of charging station owner is to maximize his profit (maximize R_i).

3.4 Modeling Users for Demand Assignment

Before describing how we model users, it is necessary to review the types of charging stations since it is intricately connected to user behavior. The two basic types are level 2 chargers (240 V AC charging) and DC chargers (500 V). The former are more widespread (can even be installed in residential locations), whereas the latter are speedier to charge (and can be found in business and government buildings). We model users in the following manner: Let us assume that a user desires to travel from location A to location B and that he will stay for a certain time in each location. If during traveling from A to B, he runs out of charge, he will first seek an available charging station in the neighborhood of A. If he can find such a charging station, he will charge there, whether he stays at least 4 h (to charge with level 2) or less (DC). Otherwise, if he could not find any charging stations, or if charging stations are fully occupied at that time, we assume that he is aware of the availability of charging stations in neighborhood of B. This part is the same as before. If there is no charging station in A or in B, he has to charge his car by DC somewhere else along his route. The use of popular routes from trajectory clustering is helpful here where users know that there are charging stations along popular roads.

There are various strategies to the demand assignment problem. For example, [30] solved a task assignment problem with linear programming to maximize resource utilization in load balancing problem in multiple machines. Here, to assign users to charging stations, we use a typical first-in-first-out approach with the goal of uniform distribution of users over charging stations. For this purpose, we start from 1:00 AM to 12:00 AM and we assign each user to the least busy charging station which is located in its neighborhood. Algorithm 1 shows pseudo-code for assigning demands to charging stations for a particular hour.

Algorithm 1: Assignment of Users to Charging Stations

```
Input: Charging Stations, User Demands
Output: Assignment matrix
for each demand, d_i do
    for each charging station, CS_j do
        if distance(d_i, CS_j) ≤ r_0 then
            A_{i,j} = 1;
        end
    end
end
for each charging station, CS_j do
    Δ_j = AvailableSlots_{CS_j} − Σ_i (A_{i,j});
end
for k = 1 to K do
                /* K is number of charging stations */;
    m = arg max_j (Δ_j);
    for each demand d_i do
        if A_{i,m} = 1 and AvailableSlots_{CS_m} > 0 then
            Assign d_i to charging station m;
            AvailableSlots_{CS_m} = AvailableSlots_{CS_m} − 1;
        end
    end
    Δ_m = −∞;
end
```

The ultimate goal from the user's point of view is to maximize the number of assigned demands as well as reducing costs associated with recharging EVs. Our user policy attempts to reduce the number of failures, i.e., the number of times that EV owners run out of charge and need to switch to traditional gas-based fuel. Also, this policy reduces the cost of charging since charging with level 2 has a higher priority compared to DC charging.

3.5 Charging Station Placement Using Clustering and Optimization

In addition to maximizing charging station owners' profits, we aim to minimize the number of failed (unassigned) demands. To this end, we aim to place charging stations next to major arterial roads and nearby stay points to provide better service for future EVs. Furthermore, we aim to have similar schedules for all charging stations to reduce very crowded or very under-utilized stations. Based on these goals, we can formulate an optimization function as a linear combination of several measures:

$$
\begin{aligned}
F(X) = & - \alpha \times \sum_{i=1}^{K} R_i + \beta \times N_{fail} \\
& + \gamma \times \sum_{t} D_{KL}(\zeta_t \,||\, U(\frac{1}{K}))/24 \\
& + \eta \times \frac{1}{K} \sum_{i=1}^{K} \sum_{p \in \phi} Distance(CS_i, p) \\
& + \theta \times \frac{1}{K} \sum_{i=1}^{K} \sum_{r \in \tau} Distance(CS_i, r). \qquad (16)
\end{aligned}
$$

where D_{KL} Kullback Leibler distance and U is the uniform distribution. The goal is to uniformly distribute demands over charging stations. Here R_i is the profit for charging station i. N_{fail} is the total number of failed demands because either their distance from their nearest charging stations was more than r_0 or because the nearest charging stations were fully occupied. τ is a set of trajectory representatives and ϕ is set of stay points. $Distance$ calculates the distance of charging stations to popular roads and stay points. α, β, γ, η, and θ are constant coefficients. $\zeta_t = [\zeta_t(1), \ldots, \zeta_t(K)]$ captures the distribution of demands over charging stations at time t. $\zeta_t(i)$ is computed as follows:

$$
\zeta_t(i) = \frac{W_t(i)}{\sum_i W_t(i)} \qquad (17)
$$

where $W_t(i)$ is the number of assigned demands to charging station i at time t.

In our work, we further focus on a downtown modeling scenario and thus restrict charging station locations to be in such areas:

$$
\begin{gathered}
\text{Minimize } F(X) \\
s.t.\ X_i \in \text{Downtown}
\end{gathered} \qquad (18)
$$

where $X = \{X_1, X_2, \ldots, X_K\}$ contains coordinates of K charging stations.

This set of charging stations contains prototypes of K clusters such that each charging station will cover a certain area and also, distance between charging stations will be maximized. Furthermore, in each area, a charging station will be responsible for future demands in that vicinity.

To optimize the objective function, we first find initial prototypes (representing charging stations) using the k-means algorithm (with geographic coordination of locations as features). Next, we use a bound-constrained optimization (simulated annealing with a maximum iteration of 500) to identify the best prototypes that minimize the objective function and also satisfy the inequality constraints, i.e. points must fall into the downtown region. Simulated annealing is used here because the search space (set of building locations) is discrete. At each iteration of simulated annealing, assignment of users to current prototypes is done with respect to the specified parameters. Calculation of profit and other parameters is done at this step. After convergence, we calculate the profit, storage size, utilization, and assigned ratio of trajectories for this final solution.

4 Experiments

Our evaluation is focused on answering the following questions:

1. Which routes are popular among EV owners?
2. How many public charging stations are necessary to serve EV owner needs?
3. What are the load profiles of the designed charging stations?

Table 1 shows parameter settings for our experiments. The price of selling energy to customer is 49 cents per kilowatt hour [31]. Also, $P_{buy,night}$ and $P_{buy,t}$ are calculated based on [31]. Also, installation cost of storage is set to 100 dollars per kWh and the installation cost of chargers in charging station is 4000 dollars [32]. For the time period of amortization, we assume six years based on [2]. We assume that available slots in each charging station is at most 10 EVs at each hour. We also assume that people can charge their EVs if they can find a charging station 800 m (i.e. walking distance) away from their current location ($r_0 = 800$ m). Furthermore, in our experiments, EV owners are assumed to have chargers in their houses and, hence, are presumed to use public charging stations during the day (and recharge again during the night [2]). On-peak hours are determined based on the nature of our dataset, that is from 6 AM to 10 AM and from 5 PM to 8 PM.

Table 1 Parameter settings used in our experiments

Parameter	Value
Time of charging (level 2 (220 V))	4 h
Time of charging (DC)	1 h
C_0 for storage	100 $/kWh
C_0 for each charger	4000 $
C_{basic}	$8 per month
Life time of utility	6 years
Number of charger in charging station	10
Electricity load (level 2)	3.3 kW
Electricity load (DC)	50 kW
P_{sell}	0.49 $/kWh
$P_{buy,night}$	(5.420 + 0.277) cents
$P_{buy,t}$	6.454 × on-peak + 5.697 × off-peak cents
r_0	800 m
$\alpha, \beta, \gamma, \theta, \eta$	1

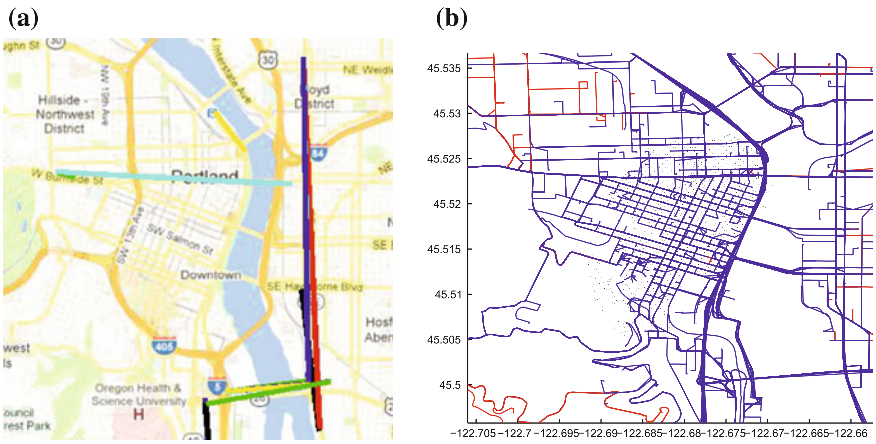


Fig. 2 **a** Trajectory representatives. **b** Schematic view of trajectories that assigned (*blue*) versus others (*red*)

4.1 Trajectory Clustering

We prune our dataset by removing users that do not enter or cross the downtown area of Portland. Next, we calculate those major routes that users take when they need to recharge their vehicles. Based on our previous work [6], a high proportion of users charge once during their daily travel and hence the probable routes for each user (when they needs to recharge their car) falls between two sequential stay points of the user. Our processing leads to 1259 trajectories. After extracting the actual trajectories from Google Maps, we use the Traclus algorithm [27] (with epsilon equal to 0.01 and minLns equal to 3) to cluster these trajectories. The result consists of 16 clusters. As Fig. 2a illustrates, representative trajectories mostly fall within the boundaries of downtown.

4.2 Ideal Number of Charging Stations

In order to compare the performance of charging stations suggested by our proposed method, we compare it with k-means clustering where only geographic coordinates of locations are considered. Figure 3 illustrates how the total profit of all charging stations changed by increasing number of charging stations. It appears that by deploying a certain number of charging stations, the total profit in our proposed method is much higher than the location-based algorithm ($5000 per day). Also, profit will begin to remain stable when the number of charging stations increases up to a certain threshold (25). It should be noted that, since the location-based k-means algorithm works solely on geographic coordinates, it will not consider the initial load values of buildings. This may cause randomness to the results.

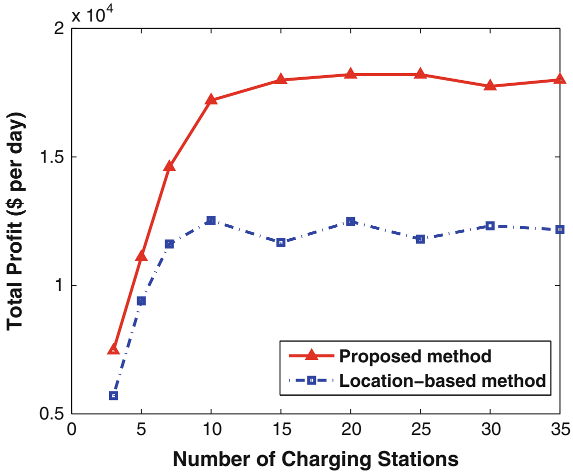


Fig. 3 Total profit of charging stations

The size of storage and utilization of charging stations (chargers and storages) are important issues in charging station and storage deployment. Utilization can be determined as $U_i = \frac{t_d}{24[h/d]}$ where $\frac{t_d}{[h/d]}$ is the daily time in use of the facility [2]. As Fig. 4a illustrates, time-based utilization of chargers is often less than 50 % but we assume satisfaction as long as the station profits exceed a certain threshold.

As the number of charging stations increases, we expect the total number of required storage units also increase. This expectation is shown in Fig. 5. Clearly, the proposed method works better than location-based k-means since the total profit and total utilization are higher. While the total storage is higher in our proposed method, time-based utilization of storages is higher than that of location-based (Fig. 4b). The time-based utilization measure depicts the percentage of the time storage units are used in charging stations.

The number of assigned users will not increase as the number of charging stations goes beyond a certain value (20). This is demonstrated in Fig. 4c. In our method, the ratio of assigned users is often more than 90 %. This ratio will vary if we change the radius of users' attention (r_0). As Fig. 4d illustrates, by increasing the allowed distance between the nearest charging station and users (r_0), the ratio of assigned users will increase. Here, number of charging stations is set to 15. This ratio in the proposed method is higher than location-based k-means because vicinity of charging stations to the common trajectories were considered in the optimization function.

To explore the profit of charging stations individually, we assess the number of charging stations with non-positive profit in each setting. As Fig. 6 illustrates, the number of such charging stations in location-based k-means is greater than in our proposed method.

Based on our results, the optimum number of charging stations which yields the highest profit and utilization is 15. For higher penetration rate, this method can be re-run to find a suitable number of charging stations. Here, we continue our experiments with 15 charging stations.

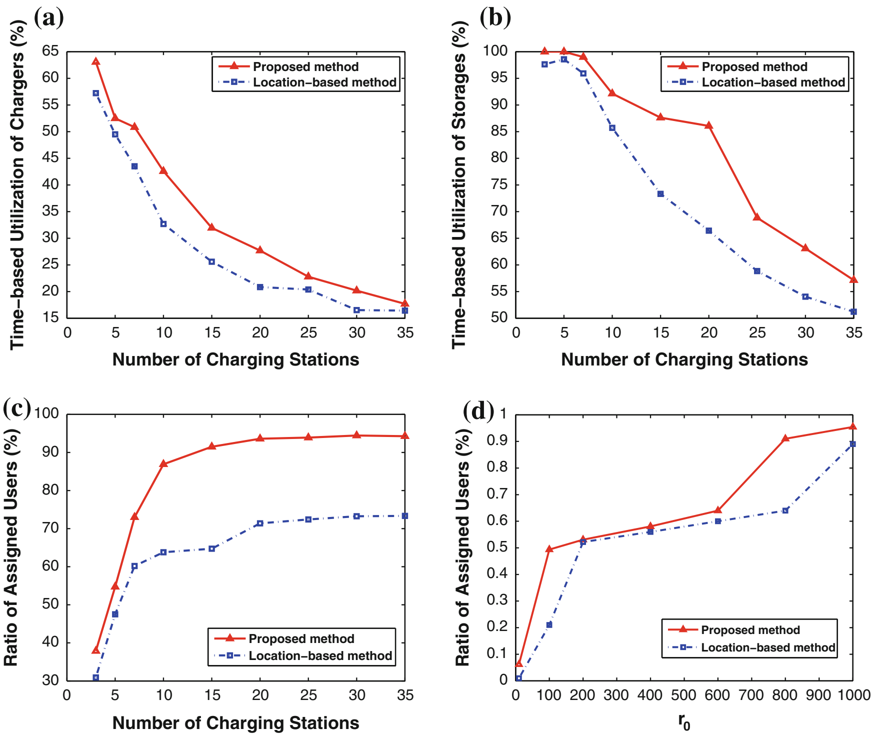
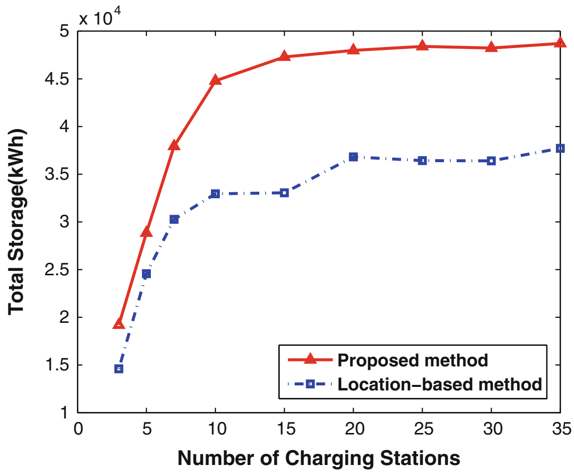


Fig. 4 Comparison of proposed method and location-based method: **a** Average time-based utilization of chargers. **b** Average time-based utilization of storages. **c** Total ratio of assigned users and **d** Total ratio of assigned users based on distance to charging stations

Fig. 5 Total storage size in charging stations



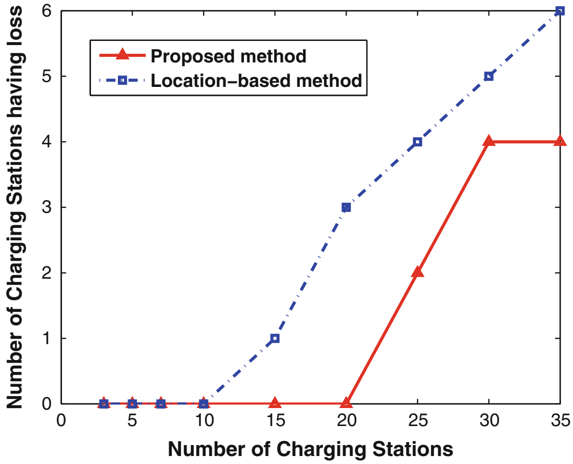


Fig. 6 Number of charging stations with loss

4.3 *Profile of Individual Charging Station*

After determining the charging stations, we can cluster other locations by considering charging stations to be the prototypes of location clusters. This strategy will be beneficial to understand which regions are covered by which charging station. Clusters of locations are shown in Fig. 7. One interesting result here is that charging stations 2 and 15 are deployed in the same locations, pointing to the potential of this location.

The daily profit, storage size, and utilization of chargers and storages in charging stations are shown in Fig. 8. A notable result here is that charging station 8 is not efficient as others since it has a low profit ($100) due to low storage and low utilization.

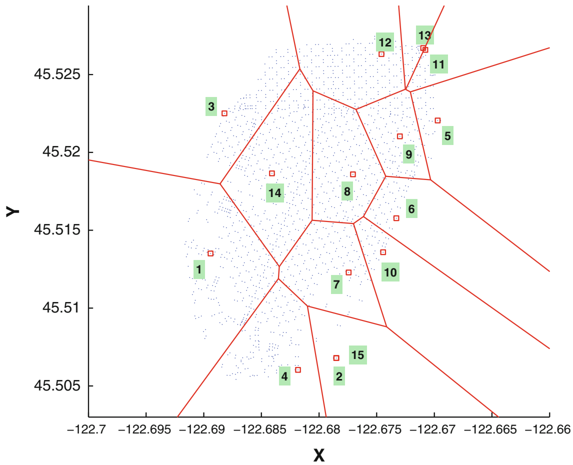


Fig. 7 Voronoi diagram of charging stations and their associated coverage area

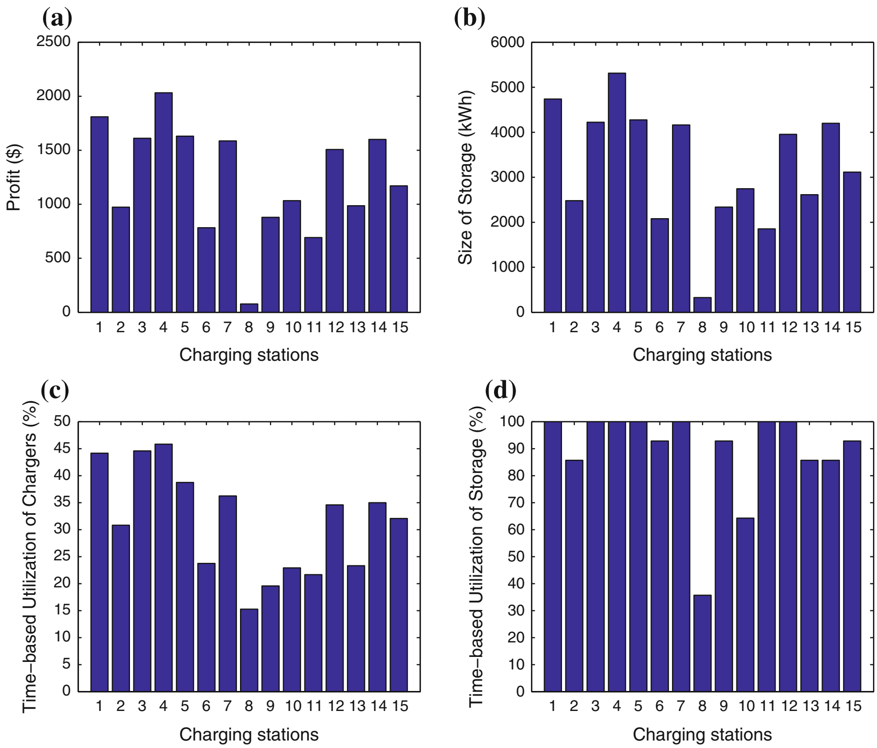


Fig. 8 Performance of charging stations: **a** Profit of charging Stations. **b** Storage of charging stations. **c** Time-based utilization of chargers and **d** Time-based utilization of storages

Most of the EVs need to be charged at peak hours (9 AM–8 PM). The number of EVs at each time slot in each charging station is shown in Fig. 9a. Since most downtown activities occur during afternoon and nights, most of the demands are concentrated between 12 PM to 8 PM. Also, the number of charging stations for each type of charging is shown in Fig. 9b, c, for level 2 and DC, respectively. Since those locations that people stay at least 4 h are outside of downtown, the demand for level 2 is lower than that for DC. Based on these results, we can determine the required number of chargers in each station. In our experiments, we assume that each charging station is able to have at most 10 chargers. As Fig. 9 illustrates, in charging station 1, we can organize it to have 2 chargers for level 2 and 9 chargers for DC. Conversely, for charging station 8, we do not need any level 2 chargers and only require 3 chargers for level 3 (DC).

Profiles of charging stations can be clustered with respect to their loads at different times. To this effect, we used the K-SC clustering approach originally proposed for time series data [33]. Here, the value of electricity load before adding EV, after adding EV, and after storage deployment during 24 h were considered as a sequence

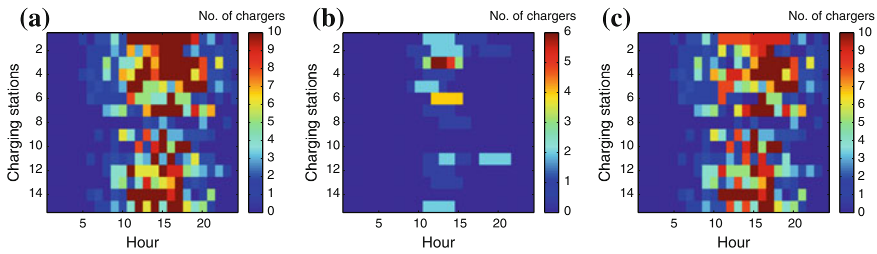


Fig. 9 **a** Number of EVs getting charged at each time slot. **b** Number of EVs getting charged at each time slot by level 2. **c** Number of EVs getting charged at each time slot by level 3 (DC)

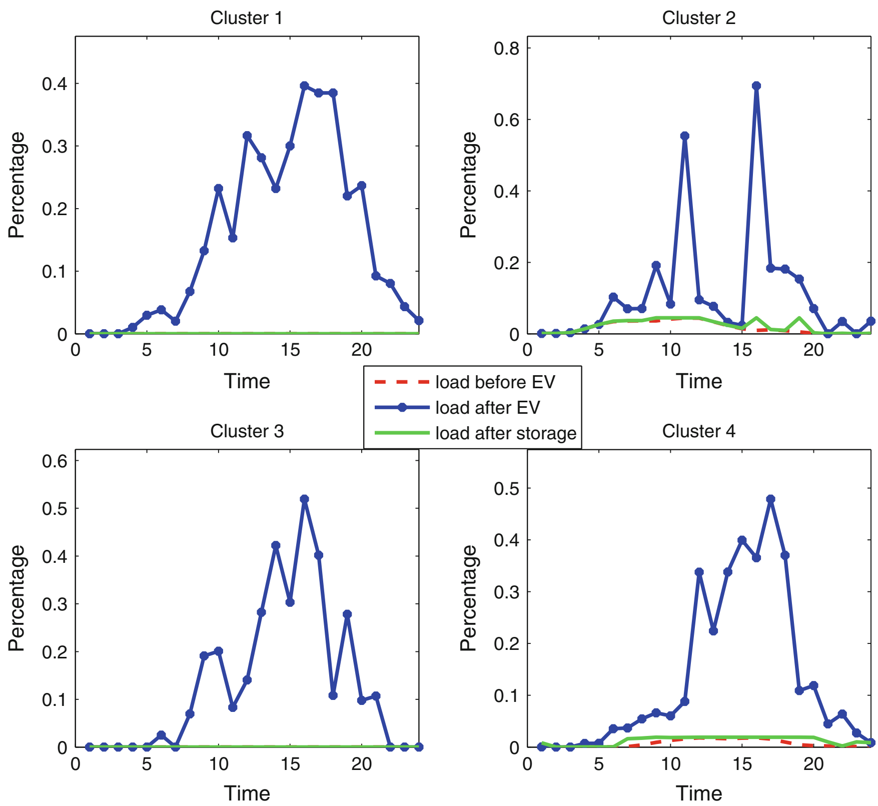


Fig. 10 Clustering of load profiles of charging stations

of 24 × 3 elements. Profiles of prototypes of four clusters are shown in Fig. 10. This figure is important in understanding the behavior of charging stations in order to make a decision between using a mobile storage versus a stationary one. Locations in cluster 1 and 3 are places where no one enters them (such as a parking lot).

Locations in cluster 2 show that the additional demand imposed by EVs lead to use of storage in 8–12 h and 16–20 h. Based on this profile, we can place mobile storage in locations where storage is needed during a specific time rather than an entire day (e.g., cluster 4).

5 Conclusion

Effective usage of the next generation of smart grids requires a comprehensive understanding of the interactions between networks of urban environments and electric systems. In this paper, we proposed a framework to design charging and storage infrastructure for electric vehicles in an urban environment. There is an inherent trade-off between user expectations and the expectations of charging stations owners, which is captured in our framework and aids in the selection of the number of charging stations along with their placement. More constraints such as availability of parking space, effects of charging stations on electrical substations, different pricing schema in charging stations are being considered for integration into our framework. Results of this research illustrate the efficiency of our approach in terms of profit maximization and energy usage. While we studied the effect of different parameters on the performance of charging station placement, there are other factors that can be considered in this problem. In this regard, the impact of different EV penetration rates and the use of probabilistic framework in assignment strategy of drivers to charging stations can be considered as future works.

References

1. Ramchurn, S.D., Vytelingum, P., Rogers, A., Jennings, N.R.: Putting the 'smarts' into the smart grid: a grand challenge for artificial intelligence. Commun. ACM **55**(4), 86–97 (2012)
2. Wirges, J., Linder, S., Kessler, A.: Modelling the development of a regional charging infrastructure for electric vehicles in time and space. Eur. J. Transp. Infrastruct. Res. **12**(4), 391–416 (2012)
3. Frade, I., Ribeiro, A., Goncalves, G., Antunes, A.: An optimization model for locating electric vehicle charging stations in central urban areas. In: Proceedings of the TRB 90th Annual Meeting, Transportation Research Board (2010)
4. He, F., Wu, D., Yin, Y., Guan, Y.: Optimal deployment of public charging stations for plug-in hybrid electric vehicles. Transp. Res. Part B Methodol. **47**, 87–101 (2013)
5. Ip, A., Fong, S., Liu, E.: Optimization for allocating BEV recharging stations in urban areas by using hierarchical clustering. In: Proceedings of the 6th International Conference on Advanced Information Management and Service, pp. 460–465 (2010)
6. Momtazpour, M., Butler, P., Hossain, M.S., Bozchalui, M.C., Ramakrishnan, N., Sharma, R.: Coordinated clustering algorithms to support charging infrastructure design for electric vehicles. In: Proceedings of the ACM SIGKDD International Workshop on Urban Computing, UrbComp '12, pp. 126–133 (2012)

7. Chen, T.D., Kockelman, K.M., Khan, M.: The electric vehicle charging station location problem: a parking-based assignment method for seattle. In: Proceedings of the Transportation Research Record (2013)
8. Swanson, J., Aslin, R., Yuccel, Z.: Electric vehicle penetration study using linear discriminant analysis. In: Proceedings of the Center for Research in Regulated Industries (CRRI), pp. 1–9 (2011)
9. Hoffman, M., Sadovsky, A., Kintner-Meyer, M., DeSteese, J.: Analysis tools for sizing and placement of energy storage in grid applications. Technical report PNNL-19703. Pacific Northwest National Laboratory Richland, Washington (2010)
10. Makarov, Y., Pengwei, D., Kintner-Meyer, M., Chunlian, J., Illian, H.: Sizing energy storage to accommodate high penetration of variable energy resources. IEEE Trans Sustain. Energ. **3**(1), 34–40 (2012)
11. Bayram, I., Michailidis, G., Devetsikiotis, M., Bhattacharya, S., Chakrabortty, A., Granelli, F.: Local energy storage sizing in plug-in hybrid electric vehicle charging stations under blocking probability constraints. In: Proceedings of the IEEE International Conference on Smart Grid Communications, pp. 78–83 (2011)
12. Liu, W., Zheng, Y., Chawla, S., Yuan, J., Xie, X.: Discovering spatio-temporal causal interactions in traffic data streams. In: KDD '11, August 2011
13. Yuan, J., Zheng, Y., Xie, X.: Discovering region of different functions in a city using human mobility and POI. In: KDD '12 (2012)
14. Yuan, J., Zheng, Y., Zhang, C., Xie, W., Xie, X., Huang, Y.: T-drive: driving directions based on taxi trajectories. In: Proceedings of the ACM SIGSPATIAL GIS 2010 (2010)
15. Zheng, Y., Zhang, L., Xie, X., Ma, W.: Mining correlation between locations using human location history. In: Proceedings of the ACM SIGSPATIAL GIS 2009 (2009)
16. Takahashi, R., Osogami, T., Morimura, T.: Large-scale nonparametric estimation of vehicle travel time distributions. In: SDM '12, pp. 12–23, April 2012
17. Rosswog, J., Ghose, K.: Detecting and tracking coordinated groups in dense, systematically moving, crowds. In: SDM '12, pp. 1–11 (2012)
18. Ester, M., Kriegel, H., Sander, J., Xu, X.: A density-based algorithm for discovering clusters in large spatial databases with noise. In: KDD '96, pp. 226–231 (1996)
19. Synthetic Data Products for Societal Infrastructures and Proto-Populations: data set 1.0. Technical report NDSSL-TR-06-006. Network Dynamics and Simulation Science Laboratory, Virginia Tech (2006)
20. Tishby, N., Pereira, F.C., Bialek, W.: The information bottleneck method. In: Proceedings of the 37th Annual Allerton Conference on Communication, Control and Computing, pp. 368–377 (1999)
21. Munro, N.: Obama hikes subsidy to wealthy electric car buyers. http://dailycaller.com/2012/02/13/obama-hikes-subsidy-to-wealthy-electric-car-buyers/ Accessed 16 May 2012
22. Simply Hired, Inc. Portland jobs. http://www.simplyhired.com/a/local-jobs/city/l-Portland,+OR Accessed 16 May 2012
23. KEMA, Inc. Distributed energy storage: serving national interests, advancing wide-scale DES in the United States. Technical report 20130065. National Alliance for Advanced Technology Batteries (2012)
24. Andrienko, G., Andrienko, N., Rinzivillo, S., Nanni, M., Pedreschi, D., Giannotti, F.: Interactive visual clustering of large collections of trajectories. In: Proceedings of the IEEE Symposium on Visual Analytics Science and Technology, pp. 3–10 (2009)
25. Monreale, A., Pinelli, F., Trasarti, R., Giannotti, F.: WhereNext : a location predictor on trajectory pattern mining. In KDD, pp. 637–645 (2009)
26. Zheng, Y., Zhang, L., Xie, X., Ma, W.: Mining interesting locations and travel sequences from gps trajectories. In: WWW **49** (2009)
27. Lee, J., Han, J., Kyu-Young, W.: Trajectory clustering : a partition-and-group framework. In: Proceedings of the 2007 ACM SIGMOD International Conference on Management of data, pp. 593–604. Beijing (2007)

28. Tran, N., Giraud-Carrier, C., Seppi, K., Warnick, S.: Cooperation-based clustering for profit-maximizing organizational design. In: Proceedings of the 2006 IEEE International Joint Conference on Neural Network Proceedings pp. 1813–1817. IEEE (2006)
29. Portland General Electric Company. Schedule 85 large nonresidential standard service (201–1,000 kW). http://www.portlandgeneral.com/our_company/corporate_info/regulatory_documents/pdfs/schedules/sched_085.pdf. Accessed 5 Feb 2013
30. Kumar, V., Saxena, P.C., Katti, C.P.: A clustering approach for task assignment problem. Int. J. Comput. Appl. **47**(7), 46–49 (2012)
31. Holly Yan. Public charging stations fuel desire for electric cars. http://www.cnn.com/2012/10/24/us/public-car-chargers/index.html. Accessed 12 Feb 2013
32. Kley, F., Dallinger, D., Wietschel, M.: Assessment of future EV charging infrastructure. In: Proceedings of the International Advanced Mobility Forum (2010)
33. Yang, J., Leskovec, J.: Patterns of temporal variation in online media. In: ACM WSDM'11, pp. 177–186 (2011)