Transportation
DOI 10.1007/s11116-013-9486-1

The return on investment for taxi companies transitioning to electric vehicles

A case study in San Francisco

Tommy Carpenter · Andrew R. Curtis · S. Keshav

Abstract We study whether taxi companies can simultaneously save petroleum and money by transitioning to electric vehicles. We propose a process to compute the return on investment of transitioning a taxi corporation's fleet to electric vehicles. We use Bayesian data analysis to infer the revenue changes associated with the transition. We do not make any assumptions about the vehicles' mobility patterns; instead, we use a time-series of GPS coordinates of the company's existing petroleum-based vehicles to derive our conclusions. As a case study, we apply our process to a major taxi corporation, Yellow Cab San Francisco (YCSF). Using current prices, we find that transitioning their fleet to battery electric vehicles and plug-in hybrid electric vehicles is profitable for the company. Furthermore, given that gasoline prices in San Francisco are only 5.4 % higher than the rest of the United States, but electricity prices are 75 % higher; taxi companies with similar practices and mobility patterns in other cities are likely to profit more than YCSF by transitioning to electric vehicles.

Keywords Electric vehicles · Bayesian networks · Public transportation · Taxis

Introduction

Replacing standard petroleum taxis with electric vehicles (EVs) can save significant amounts of petroleum. For triple-shift taxis (those that are driven 24 h a day) we estimate savings of approximately 15,000 l each year per taxi, and 10,000 l for double-shift (16 h a

T. Carpenter (✉) · A. R. Curtis · S. Keshav
School of Computer Science, University of Waterloo, 200 University Avenue West,
Waterloo, ON N2L 3G1, Canada
e-mail: tcarpent@uwaterloo.ca

A. R. Curtis
e-mail: a2curtis@uwaterloo.ca

S. Keshav
e-mail: keshav@uwaterloo.ca

Published online: 21 June 2013

day) taxis. However, taxi operators will only invest in EVs if it is economically viable. Therefore, we design a process to determine a taxi company's ROI in transitioning to electric vehicles. We first build a Bayesian model of taxi fleet mobility. We then show how to use the model to determine the return on investment (ROI).

Existing research on transitioning petroleum vehicles to EVs has focused on bus transit and personally owned vehicles. Taxis do not travel on fixed routes like transit buses and are not parked a majority of the time like personally owned vehicles. As a consequence of these different mobility patterns, we were not able to apply any existing model to study the cost of transitioning a taxi fleet to EVs.

Our main contributions are:

1. *We propose a data-oriented process to compute a given taxi company's ROI in transitioning to electric vehicles*. The process uses time-series of GPS and passenger data collected from instrumented petroleum taxis. We use Bayesian data analysis to compute the costs of operating a taxi fleet consisting of petroleum vehicles, plug-in hybrid vehicles (PHEVs) or battery electric vehicles (BEVs).
2. *As a case study, we analyze adoption of PHEVs or BEVs by a taxi company with over 530 vehicles, Yellow Cab San Francisco*. We study different infrastructure scenarios, including battery switching and roadside charging. For each scenario, we quantify the revenue losses or gains, quantify the investment payback period, and extrapolate the analysis to a wide array of electricity and petroleum prices.
3. *We formulate the problem of locating battery switching stations that serve the taxi fleet as an optimization problem, and present a framework to compute this placement for a given city*. Using our algorithm, we find only three battery switching stations are needed for Yellow Cab San Francisco for BEVs to be profitable.

Our case study shows that both PHEVs and BEVs have a positive ROI given (2011) vehicle and fuel prices in San Francisco.

The rest of this paper is organized as follows. "Background and related work" are given in this section. We present our ROI analysis methodology in "Data Oriented Process to Estimating ROI". The results from our case study can be found in "San Francisco Case Study". Future work and limitations are discussed in "Limitations and future work", and finally we present our "Conclusions" in that section. A brief background on Bayesian networks is given in Appendix A and the details of our optimization algorithm are presented in Appendix B.

Background and related work

This section presents an overview of electric vehicles, existing taxi corporation practices, and related work.

Electric vehicle types

We study two types of electric vehicles:

1. *Plug-in Hybrid Electric Vehicles (PHEVs)*. PHEVs have a grid-chargeable battery and an internal combustion engine (ICE). PHEVs are powered completely from the battery for first portion of a trip, without using the ICE. The standard notation to describe a PHEV is "PHEVxxm", where xx refers to the distance in kilometers (km) the PHEV is

expected to drive using only the battery. Once the battery has been nearly depleted, the ICE is used to propel the vehicle for the remainder of the trip and the battery may continue to power electronics onboard the vehicle.

2. *Battery Electric Vehicles (BEV)*. BEVs are fully powered by batteries and do not have an ICE; these vehicles are not reliant on petroleum for transportation.

Terminology

We now define a few terms used throughout the paper.

- *SV (standard vehicle)*. SV is an abbreviation for fully ICE-powered petroleum vehicles.
- *Kilowatt-hour*. A kilowatt-hour (kWh) is a unit of energy. One kWh is the amount of energy used by a device consuming power at a rate of 1 kW for 1 h, equal to 3.6×10^6 J of energy.
- *Capacity*. The capacity of an EV is the amount of energy (usually stated in kWh) that can be stored in its battery.
- *Discharge Rate*. The discharge rate of an EV is the rate at which the EV consumes energy. It is analogous to the km per liter of an SV.
- *Battery Charging and Battery Switching*. There are two ways to refuel an EV. *Battery charging* is when the car is plugged into a charging unit with a connection to the electrical grid. These units transfer power from the grid into the battery at a rate depending on the type of charging unit and the availability of power in that location. This process takes several hours depending on the type of charging unit. The other option is *battery switching*, where a battery switching station physically changes the batteries in an EV. Here, a user comes to the station with a nearly-depleted battery and the battery is replaced with a fully charged battery. The depleted batteries are then charged at the switching station. This type of refueling is similar to petroleum vehicles, where the vehicle is refueled in minutes instead of hours. The first major manufacturer of switching station infrastructure states that the entire switching process takes only 80 s (Galbraith 2009). We note that as of 2012, battery switching is only available for BEVs, and BEVs with this capability are only starting to enter the market. The Nissan Leaf, for example, does not have a switchable battery. However, companies like Better Place have now deployed switching stations and delivered switchable BEVs to some customers (Bloomfield 2012).
- *Charge Rate*. The charge rate of an EV is the rate at which the vehicle draws power from the grid into the battery at a charging station. This rate is not limited by the battery, but by the amount of power that can be supplied by a charging unit (Kempton and Tomic 2005) (which depends on the type of unit and the electrical connection the charger has to the grid). There are three STANDARD levels of charging. Level 1 charging is the slowest form and uses a 110 V connection found in any standard electrical outlet in North America. Level 2 charging uses a 220 V connection, which most homes and businesses also have for large appliances. Level 3 is known as “quick charging,” and uses a 480 V connection (Boulanger et al. 2011; Canizares et al. 2010). However, quick charging has not yet been widely deployed.
- *Range*. The range of a vehicle is the expected distance it can travel given normal driving conditions when fully fueled (using petroleum for SVs, electricity for BEVs, and both for PHEVs). PHEV manufacturers additionally state the range that can be driven using electric power before the ICE is used.

Taxi operation overview

We briefly describe the operation of a taxi company as it relates to our work and how its operating practices may change if it transitions its fleet to EVs.

Our model applies to taxi companies in *fare-regulated* taxi markets where the local government or local Taxi Commission controls the pricing structure of taxis within the region. Taxi companies cannot unilaterally change their pricing structure. Moreover, we assume that employee-drivers operate company-owned vehicles in *shifts* that typically last 8 h. At the end of the shift, drivers return the vehicle to the company premises where the vehicles are re-fueled and handed over to another driver for another shift. Drivers only refuel the vehicles while they are not carrying passengers.

Fares may be pre-arranged by calling the company to schedule a pickup, or they can be arranged on-the-fly by signaling taxis as they drive past. We use the term *fare* to refer to a contract between a driver and a passenger to transport the passenger to a desired destination for some price. Each taxi company has its own pricing model, that is, how it charges for fares. It is usually a function of time, distance, and other incidental charges. A driver's goal is to complete as many fares as possible during their shift, as this is the sole source of revenue for the company.

To maximize the likelihood of fares, taxi drivers may continuously drive around looking for passengers or may wait at busy locations such as airports and city centers. This behavior is not fuel efficient, but the revenue from additional fares usually compensates for the cost of wasted fuel.

We now note how this existing operation may change if the taxi company were to convert their fleet to BEVs or PHEVs. Such a change can impact the frequency of refueling, the potential introduction of refueling delays between shifts, and driver behavior between fares.

- *BEVs*. The range of a BEV is about a third of an SV. Thus, BEVs must be refueled about three times more often than SVs. Consequently, either drivers must refuel more often between fares or turn down more fares. Note that installing battery switching stations allows BEVs to be refueled as quickly as SVs. Therefore, there is no additional delay at the company premises between shifts.

Many equations in the following sections are dependent on a variable τ, which represents the battery charge threshold below which taxi drivers switch their battery if they are at a location with a switching station. We use the notation $\cdot(\tau)$ to represent a variable's value assuming the switching threshold is τ. We assume in our analysis that a BEV driver switches their battery whenever:

1. Their battery's charge level is less than τ,
2. The driver is at a location with a switching station—we assume drivers never modify their trajectories to switch batteries.

We discuss computing the optimal value of τ in "Optimal switching threshold".

- *PHEVs*. PHEVs do not have to be refueled more often than SVs. However, the primary gain from switching to PHEVs is to reduce fuel costs by driving the taxis primarily using the battery. This reduction is possible only if taxis rarely switch to their ICE mode, which requires their batteries to be fully charged after the end of a shift. PHEV batteries cannot be switched in today's models. This introduces a large delay between shifts while the vehicles are charged. To avoid this delay, the taxi company could purchase additional PHEVs to ensure vehicle availability for the next shift. This issue is discussed in detail in "Battery and extra vehicle costs".

- In both cases, a driver's practice of opportunistically attracting fares by driving around would be affected. Drivers need to trade off the benefit from additional fares for the cost of battery depletion.

Related literature

Our work concerns three major areas; the integration of EVs into public transportation, revenue analyses for EV integration, and infrastructure location. Here we present relevant work in these areas.

- *Integration Into Public Transit*. Several papers have addressed the feasibility of hybrid and electric bus transit (Darovsky et al. 2010; Tzeng et al. 2005; Wirasingha et al. 2008). Gao and Kitirattragarn (2008) study the likely hybrid EV (HEV) penetration level in New York City and the potential environmental benefits of transitioning. We are unaware of any economic study of a taxi company transitioning to EVs. We note that Better Place, a manufacturer of EV infrastructure, has recently completed a feasibility study of EVs in Tokyo, Japan, but the results from this study have yet to be published. They have also stated they will be conducting experiments in San Francisco (Better Place 2010), the location of our case study.
- *Profitability Analyses*. Several papers address the issue of cost revenue analysis for individual customers purchasing EVs (Aecom 2009; Becker et al. 2009; Delucchi and Lipman 2001; EPRI 2001; García and Miguel 2012; Kliesch and Langer 2006; National Research Council 2010; Prud'homme 2010; Simpson 2006). These studies target individual drivers, who have different mobility patterns than that of taxi drivers. Moreover, the major limitation of these studies is that their analysis is based on nationwide average values for drivers' mobility patterns, and these generally come from national government surveys. This is because mobility data is hard to obtain from individual drivers. Our work is data oriented; it assumes the taxi company has monitored their taxis with GPS devices (which most taxi companies already do in case their vehicles are hijacked or stolen), and our model is based on analyzing these GPS traces. We make no assumptions about the taxis' mobility patterns.
- *Infrastructure Location*. Papers regarding locating alternative refueling infrastructure are divided into two classes: those that assume drivers are not willing to deviate from their paths to stop at a refueling station, and those that assume drivers would deviate a small distance from their paths. Traditionally, refueling facilities are considered *discretionary facilities*—facilities that are not traveled to as a destination themselves but rather facilities where drivers stop on the way to other destinations (Berman et al. 1990, 1992). We also make this assumption in this work.

Papers in this space present *facility location* algorithms to place facilities beside roads with high vehicle flows. Shukla et al. (2011) give a simple refueling model based on the flow interception facility location model (FIFLM) (Berman et al. 1990, 1992). Their objective is to place refueling facilities to maximize facility *interceptions*, where an interception refers to a facility being on a driver's route between an origin and destination. Kuby and Lim (2005) later adapt this model and introduce the flow-refueling location model (FRLM) which aims to maximize the total refueled flow volume given the range constraints of vehicles. This original FRLM formulation limits the set of potential facility sites, known as *candidate sites*, to a set of network *nodes*, points at which various roads cross. Kuby and Lim (2006) later improved their original candidate site selection algorithm to include “smarter” sites in addition to nodes, showing that restricting all facilities to nodes is suboptimal. Upchurch et al. (2009) further refine the FRLM by allowing for facility capacities—they introduce the *capacitated FRLM* allowing bounds on the number of passengers serviceable at a single facility per unit time. Finally, Kuby et al. (2009) demonstrate their FRLM on a detailed case study; siting a network of Hydrogen refueling facilities in Florida for the Florida Hydrogen Initiative.

Our work differs than these studies because we are not interested in maximizing flow volume or total refueled miles, we are interested in maximizing profit. If battery switching stations are located for public use, the former is more appropriate to maximize drivers' aggregate utility. However, for a private company, the number of flows through a region may not be the only factor in determining whether a switching station should be placed there, as we assume taxi drivers do not stop while on a fare to switch their batteries. Under this assumption, locations where taxis park often may be more appropriate.

Data oriented process to estimating ROI

Our goal is to calculate the company's ROI in transitioning a fraction of their taxi fleet to EVs. To do this, we describe an analytical technique that allows us to use measurements of the company's existing SVs to study whether they should adopt EVs. This section is outlined as follows. In “Model assumptions” we discuss the assumptions we make in our model. In “Inputs”, we describe the required inputs necessary to use our model. In “Outputs” we describe the model outputs. We describe how we estimate the company's SVs as EVs in “Estimating charge levels”. In “Using the model to infer costs” we discuss how the model is used to infer the company's ROI as a result of transitioning a portion of their fleet to EVs.

Model assumptions

We make the following simplifying assumptions. Many of these are further discussed as future work (“Limitations and future work”).

- We assume that if a taxi depletes its battery during a shift, all revenue the taxi would have generated during the remainder of that shift is lost. This is a conservative assumption given that in practice a taxi could potentially drive to a switching station, but this would change the trajectory of the vehicle. Since we are using records of past trips, we would not know whether the fares after this trip to a switching station would be valid.
- Our model takes as input the lifetime L of the company's taxis; this is further discussed in “Inputs”. A common range for L is three to five years, based on surveying taxi companies in different cities. We assume the company replaces every SV, BEV, PHEV, and battery after L years of use; that is, we extend L to be the replacement rate of all SVs, BEVs, and PHEVs, and batteries, regardless of their usage. It would be straightforward to add an additional model parameter separately modeling the lifetime of batteries if desired.
- As stated in “Taxi operation overview”, our model only applies to fare-regulated taxi markets. We cannot estimate the revenue derived from taxis or the relative costs of

owning an SV versus an EV in an unregulated market, as taxi drivers could make arbitrary changes to their fare structure in response to changes in gasoline and electricity prices, making our projections inaccurate. From surveying a representative sample of cities in North America however, we conclude most taxi markets are regulated by the local Taxi Commission or local government.

- Our model is data oriented—it relies on measured data rather than a survey of taxi operations. Any dataset will have values "built in" because the values are measured from an already existing system, such as the average number of fares completed per taxi, the average utilization of the company's taxis, and the average revenue brought in per taxi. We assume the company does not make a large change to their fleet size so these values over the next L years will resemble the values from the dataset.
- We assume that the price of gasoline and electricity do not go below some lower bounds during L. That is, our model computes a worst-case ROI—it assumes the price of gasoline and electricity are at least as high as the input parameters. Removing this assumption would require us to introduce stochastic processes for the price of gasoline and electricity, greatly complicating the model.
- Our model computes the ROI for the taxi company over L years assuming a fraction of their fleet is replaced with EVs at the *start of L*. That is, we assume EVs are introduced into the fleet at the start of the analysis period. It is straightforward to extend this analysis to the case where EVs are introduced in batches, as each ROI analysis is independent, and the total ROI would be the sum of each batch.
- We assume batteries charge at a constant rate and the battery capacity does not change over its lifetime.
- We do not allow taxis to deviate from the routes in the data set, that is, we do not modify the dataset in any way. Taxis are only allowed to switch batteries if they are at a location with a switching station and do not have a fare. In practice, taxis would monitor their battery levels and may deviate to a switching facility in between fares.
- While we do model the number of batteries needed at each battery switching station using queueing theory, we do not model queuing delays at switching stations due to multiple taxis switching at the same time. This is left as future work.
- Our model ignores overheads that remain fixed whether the company uses EVs or SVs, such as driver salaries and dispatch expenses.
- We assume maintenance costs for a fleet of EVs is the same as a fleet of SVs. This is a conservative assumption because EVs, with fewer moving parts, require less maintenance than SVs.
- We do not model swapping station maintenance costs.
- We assume the taxis do not charge with the air conditioning (AC) on. We model the effect of AC usage on the battery capacity, but make this assumption so that at any time the taxi is either consuming or gaining energy, but not both.

These assumptions essentially state the taxi market remains relatively constant over L years. Removing these assumptions requires a complicated stochastic model using stochastic prices, fleet size etc. Building such a model is left as future work.

Inputs

Our process for determining the changes in revenue for the company as a result of switching to EVs requires the following inputs:

1. *Mobility Data*. A critical input to our taxi model is mobility data from the existing SV fleet. We require the periodic collection, from each taxi, of its geographical location and fare status, for a period of several weeks. This could be obtained by collecting a log file for each SV, where each record of the log file has a time stamp, the GPS location of the SV, and whether there is a paying passenger currently in the vehicle. The input dataset must be a set of *shift files*, where a shift file represents data for one drivers working shift as defined in ”Taxi operation overview“. We require a set of shift files for each driver and each taxi.
2. *Reduced Coordinate Space*. A second input to our model is a reduced coordinate space that minimizes model dimensionality without overly affecting its correctness. We overlay the taxi company's geographical operating region with the set of points specified in the reduced space and we map GPS data to its closest grid coordinate using Euclidean distances.
3. *Fare Pricing Model*. Every taxi company has their own pricing function they use to charge for fares, and this needs to be given as input. Let r_{FARE} be the cost of one fare. Most taxi companies use a function of the following form:

$$r_{\mathrm{FARE}}(C_{\mathrm{I}}, d, C_{\mathrm{D}}, p, C_{\mathrm{T}}, M) = C_{\mathrm{I}} + d \cdot C_{\mathrm{D}} + p \cdot C_{\mathrm{T}} + M \tag{1}$$

where C_{I} is an initial cost, d is the distance traveled during the fare, C_{D} is a cost per kilometer, p is the time parked at traffic lights, C_{T} is the cost per minute of waiting at lights, and M is miscellaneous fees.

4. *Operating Costs*. Gasoline, electricity, and vehicle prices vary between different cities around the world.
5. *Vehicle Specifications*. We require specification of EV parameters such as battery size, range, and charging rates.
6. *Vehicle Replacement Rate*. We require the lifetime L of the company's vehicles. Some taxi commissions require vehicles be replaced based on their age in years, while others require taxis be replaced after having driven a certain distance. Because this is a data-oriented model, for the latter case L can be estimated from the mean distance driven by the taxis each year.

Outputs

The process produces the following outputs:

1. The company's ROI over L years, based on the fraction of the fleet transitioned BEVs or PHEVs, and the average lifetime L of the company's vehicles.
2. Assuming a PHEV transition, the number of additional vehicles that must be purchased so that each driver can begin their shift with a fully-charged vehicle.
3. Assuming a BEV transition, the number of extra batteries the company must purchase.
4. Assuming a BEV transition, the number and location of needed battery switching stations.

Estimating charge levels

A necessary intermediate step in our analysis is to estimate the charge level of an EV battery based on its mobility pattern and initial state of charge. In this section we describe how we estimate SVs as though they were EVs.

Estimating EV battery charge level

We develop a Bayesian model to infer an EV's charge level at any time t given the time-series of GPS coordinates from the corresponding SV. We found Bayesian networks are a natural fit for inferring the hidden charge-level variable. The problem of estimating battery charge levels can be modeled as a *causal* graphical model, as explained in the following sections, and Bayesian networks are designed to infer variables' values in casual models. We refer the reader to Appendix A for further information on the type of Bayesian network we are using, and to Koller and Friedman (2009) for detailed background information. Here, we present our specific model.

A *dynamic Bayesian network* tracks variables that change over time by observing them at discrete *timeslices*. A *timeslice* is an instantaneous point in time that we observe the network, and the kth timeslice is denoted t_k. These timeslices are spaced by a *timestep*, a period of time between two timeslices, which can be constant or variable. The choice of the timestep duration is difficult in many applications. In our case however, we have a natural solution to this problem—we associate one timestep for every GPS measurement. This can be thought of as observing constantly changing variables at "random" points in time; random because two GPS measurements can have an arbitrary length of time between them, so any two timeslices can have an arbitrary timestep between them.

It is important not to confuse the relationship between continuous variables and discrete timeslices. Continuous variables in dynamic networks are real-valued but are observed at discrete timeslices. For example *time* is a continuous variable in our network even though we observe this variable at discrete timeslices.

We define $P(X(t_k)|Pa(X(t_0,\ldots,t_{k-1},t_k)))$ as the conditional distribution of X at time t_k, given the values of its parents at times $t_0,\ldots,t_{k-1},t_k$. For some models, computing this distribution may be extremely complex, as the set $Pa(X(t_0,\ldots,t_{k-1},t_k))$ may be large. We assume that our model follows the *Markov assumption*, i.e.,

$$P(X(t_k)|Pa(X(t_0,\ldots,t_{k-1},t_k))) = P(X(t_k)|Pa(X(t_k,t_{k-1})))$$

for discrete timestep Bayesian networks. This assumption greatly reduces the computational cost of querying the network. Charge level is the variable we ultimately estimate from the network, and the charge level at timestep t_{k+1} is independent of $t_0,\ldots,t_{k-1}$; it is only dependent upon its state at t_k and the energy used or gained between t_k and t_{k+1}. Furthermore, because fares may be requested by random passengers to and from anywhere, it not unreasonable to assume the location of a taxi is independent of its prior locations. (In reality, the location is actually dependent; for example taxis that travel from the downtown portion of a city out to the airport may be more likely to return downtown than to serve fares near the airport.)

There are many varieties of Bayesian networks. We use a dynamic conditional linear Gaussian network (DCLGN). DCLGNs allow inference (querying) of both continuous and discrete variables, and are designed for querying *expectations of variables* instead of probabilities. We are interested in the query "what is the expected current charge level of the taxi". Standard Bayesian networks can only answer probability queries, e.g., "what is the probability the charge level is currently X?". In the latter case, our only option is to compute $P(X|Pa(X))$ for all values of $Pa(X)$ and then find $E(X)$. With DCLGNs, we already have $E(X)$ and can easily query its value. Therefore, even though DCLGNs are more theoretically complex, they are far more efficient for querying means as opposed to probabilities; further details are presented in Appendix A.

Inferring charge level

We now discuss querying the network for the battery charge level. We first note a simplifying assumption. Charge level is a hidden variable and is never observed. We maintain the mean of the charge level which is assumed to be a Gaussian distribution. We do not maintain the variance in our model, but we explain this limitation in Appendix A.

Our goal is to find the mean of the charge level Gaussian at every timestep (every GPS data point). Figure 1 shows the graphical model that is used to estimate the battery charge levels over time. The dotted arrows represent variables that have an effect on the next timeslice called *persistence edges*. The solid lines represent *inter-time edges* that do not affect variables at the next timeslice.

In Bayesian networks, the complexity of querying a variable X grows exponentially with $|Pa(X)|$. Therefore we introduce *helper variables* that reduce the number of parents of variables we are interested in querying. Table 1 shows the variables in our network, and whether they are observed, helper variables, hidden, discrete or continuous.

We now explain the three most important variables in the network.

- *Energy Used*. This variable represents the energy the taxi consumes between two timesteps. Let D be the discharge rate of the EV (kWh/km), and e represent the most significant energy usage of an EV other than propelling the vehicle: air conditioning (AC). Then,

$$u(t_{k-1}, t_k) = d(t_{k-1}, t_k) \cdot D + e \tag{2}$$

The US National Renewable Energy Laboratory states "Air conditioning loads can reduce EV range and HEV fuel economy by nearly 40 % depending on the size of air conditioner and driving cycle". We discuss our AC assumptions in San Francisco Case Study, and note that AC usage by taxi companies is different depending on climate in their regions.

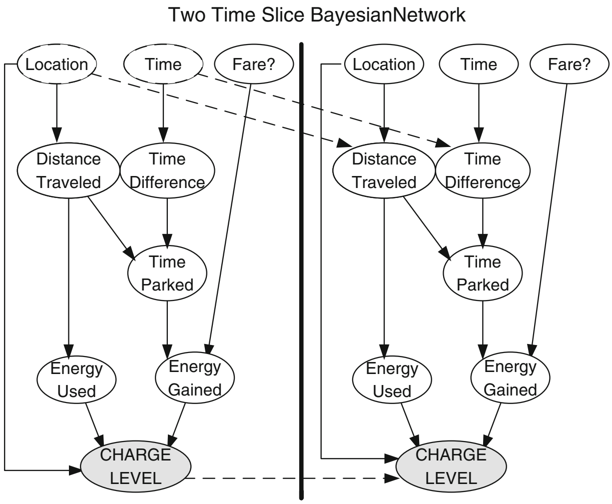

Fig. 1 Bayesian network to estimate the battery charge level

Table 1 Table of variables in the Bayesian network

Variables			
Name	Symbol in Eq.	Meaning	Type
Location	$L(t_k)$	Taxi's current location	O, D
Fare	$F(t_k)$ (0/1)	Whether a passenger is in the vehicle at t_k	O, D
Time		Value of time	O, C
Distance traveled	$d(t_{k\text{-}1},t_k)$	Km traveled between last two timesteps	H, C
Time difference		Time elapsed between last two timesteps	H, C
Time parked	$p(t_{k\text{-}1},t)$	Seconds parked between last two timesteps	H, C
Energy used	$u(t_{k\text{-}1},t_k)$	kWh used between last two timesteps	H, C
Energy gained	$g_x(t_{k\text{-}1},k)$	Energy gained between last two timesteps	H, C
Charge level	$CL(t_k)$	Current charge level	Hd, C

O Observeable, *H* Helper, *Hd* Hidden, *D* Discrete, *C* Continuous

- *Energy Gained*. This variable is only used when studying the effect of roadside charging on battery charge level. If a taxi is parked between two timesteps, we assume the taxi could have been charged during this time. Let $g_x(t_{k\text{-}1},t_k)$ be the energy gained between two timesteps assuming level x charging (kWh), and B_{gx} is the amount of energy gained per second (kWh/s) assuming level x charging (this depends on the BEV or PHEV model). Because we assume drivers never charge with passengers in the vehicle, we model Level 1 and charging and use the formulas:

$$g_1(t_{k-1}, t_k) = F(t_k) \cdot p(t_{k-1}, t_k) \cdot B_{g1} \tag{3}$$

$$g_2(t_{k-1}, t_k) = F(t_k) \cdot p(t_{k-1}, t_k) \cdot B_{g2} \tag{4}$$

- *Charge Level*. Charge level is the variable our network is designed to query. This variable has four parents: the charge level from the previous state, the current location, the energy used, and the energy gained. The edge between the two variables *charge level* and *location* is because the *charge level* is dependent upon *location* because of battery switching. Let $CL(t_k)$ be the charge level at timestep t_k, then[1]

$$CL(t_k) = \begin{cases} full^* \\ \quad \text{if } L(t_k) \text{ has a switching station} \\ \quad \text{and } CL(t_k) \quad < \tau \\ \quad \text{and } F(t_k) = 0 \\ CL(t_{k-1}) + u(t_{k-1}, t_k) + g_x(t_{k-1}, k) \\ \quad \text{otherwise} \end{cases} \tag{5}$$

Note that one of $u(t_{k\text{-}1}, t_k)$, $g_x(t_{k\text{-}1}, t_k)$ will always be zero—either the taxi parked and acquired energy or the taxi traveled and used energy.

[1] *This equation only applies when studying BEVs with switching stations. The different scenarios we study are given in Case study EV scenarios

Using the model to infer costs

We now describe the process to use this Bayesian model to determine company's ROI in transitioning their fleet to EVs.

Notation

In the following equations, we use i to index a specific taxi, w to index a specific shift, f to index a specific fare, and j to index a specific battery switching station. I represents the total number of SV taxis the company owns.

Methodology

For BEVs and PHEVs respectively, our process is to compute

$$r_B(\tau) = r_E - r_L(\tau) + s_B(\tau) - C_B - c_{EB}(\tau) \tag{6}$$

$$r_P = r_E + s_P - C_P - c_{EP} \tag{7}$$

$$\Delta_B(\tau) = r_B(\tau) \cdot x - c_{BSS}(\tau) \tag{8}$$

Table 2 Variables used in determining ROI

Name	Description
I	Total number of *SV* taxis the company owns
L	The replacement time (years) at which all SVs, BEVs, PHEVs, and batteries are replaced
$\Delta_B(\tau)$	Total ROI in transitioning x SVs to BEVs
Δ_P	Total ROI in transitioning x SVs to PHEVs
$r_B(\tau)$	Average ROI per BEV over L
r_P	Average ROI per PHEV over its L
x	No. of vehicles the company transitions to EVs
r_E	Average revenue generated by the company's existing SV taxis
r_E^i	Revenue of the company's SV taxi i
$r_L(\tau)$	Average revenue lost per taxi from missed fares
$r_L^i(\tau)$	Revenue lost by taxi i from missed fares
$r_T^i(w, \tau)$	Revenue generated during shift w of SV taxi i
$r_S^i(w, \tau)$	Revenue generated during shift w of taxi i *before battery depletion* assuming it was a BEV
$s_B(\tau)$	Average fuel savings per taxi from using BEVs instead of SVs
$s_B^i(\tau)$	Fuel savings from taxi i generated by using a BEV instead of an SV
s_P	Average fuel savings per taxi from using PHEVs instead of SVs
s_P^i	Fuel savings from taxi i generated by using a PHEV instead of an SV
$c_{EB}(\tau)$	Average cost of BEV batteries needed per taxi
c_{EP}	Average cost of extra PHEVs,per taxi, needed so drivers start their shift with full batteries
$c_{BSS}(\tau)$	Cost of all battery switching stations needed
C_B	Cost(BEV-SV); incremental BEV cost
C_P	Cost(PHEV-SV); incremental PHEV cost
r_{FARE}	Revenue generated from one fare (see Eq. 1)

$$\Delta_{\mathrm{P}} = r_{\mathrm{P}} \cdot x \tag{9}$$

where the terms in this equation are defined in Table 2.

We now explain how we compute each of these costs.

Determining existing taxi revenue

We compute the company's existing revenue r_{E} using the fare data and the companies pricing model. Let F_i be the set of all fares completed by taxi i. Using the input r_{FARE} from "Inputs", the pricing function the company uses for a fare,

$$r_{\mathrm{E}}^{\mathrm{i}} = \sum_{f \in F_i} r_{\mathrm{FARE}}(C_{\mathrm{I}}, d_f^i, C_{\mathrm{D}}, p_f^i, C_{\mathrm{T}}, M) \tag{10}$$

$$r_{\mathrm{E}} = \frac{\sum_{i=1}^{I} r_{\mathrm{E}}^{i}}{I} \tag{11}$$

where d_f^i, p_f^i come from fare f (see Table 3) and C_{I}, C_{D}, C_{T}, M come from the input pricing function for a fare.

Revenue loss from lost fares

We now show how to compute the revenue loss due to transitioning to BEVs, $r_{\mathrm{L}}(\tau)$ (PHEVs do not have revenue losses as they use the ICE after battery depletion). We assume that if a taxi depletes its battery during a shift, all revenue the taxi would have generated during the remainder of that shift is lost. This upper bounds revenue losses because we assume drivers can only switch batteries if they are in a location with a switching station and they do not modify their paths to drive to a switching station. As a result of this worst-case restriction, the drivers may deplete their battery on their shift under our model. Let δ_i be the set of all shifts completed by taxi i. We determine the revenue loss by:

Table 3 Variables used in computing fuel savings

Name	Description
G	Price of gas per liter
E	Price of electricity per kWh
V_{E}	Efficiency of an SV (mpg)
B_{E}	Efficiency of the BEV (mpkWh)
P_{E}	Efficiency of the PHEV (mpkWh)
P_{G}	Efficiency of the PHEV after battery depletion (mpg)
p_f^i	Time the company's taxi i is parked at lights while on fair f
d_{S}^i	The total distance driven by the company's SV taxi i
d_f^i	Distance driven by the company's SV taxi i during fare f
$d_{\mathrm{B}}^i(\tau)$	Total distance driven by taxi i assuming it is a BEV: sum of distance driven before depletion over all shifts
d_{E}^i	Total distance driven by taxi i assuming it is a PHEV on electricity
d_{G}^i	Total distance driven by taxi i assuming it is a PHEV on petroleum

$$r_{\mathrm{L}}^{i}(\tau)=\sum_{w\in\delta_i} r_{\mathrm{T}}^{i}(w,\tau)-r_{\mathrm{S}}^{i}(w,\tau) \quad (12)$$

$$r_{\mathrm{L}}(\tau)=\frac{\sum_{i=1}^{I} r_{\mathrm{L}}^{i}(\tau)}{I} \quad (13)$$

Fuel cost reduction

Transitioning to EVs has the one primary financial benefit: electricity is a cheaper form of fuel than petroleum so it costs less to fuel EVs than SVs. Note that PHEVs usually have a higher gasoline efficiency after battery depletion compared to SVs due to regenerative breaking (which is taken into account in our model).

The fuel savings are computed as follows,

$$s_{\mathrm{B}}^{i}(\tau)=G\cdot\left(\frac{d_{\mathrm{S}}^{i}}{V_{\mathrm{E}}}\right)-E\cdot\left(\frac{d_{\mathrm{B}}^{i}}{B_{\mathrm{E}}}\right) \quad (14)$$

$$s_{\mathrm{P}}^{i}=G\cdot\left(\frac{d_{\mathrm{S}}^{i}}{V_{\mathrm{E}}}\right)-\left(E\cdot\left(\frac{d_{\mathrm{E}}^{i}}{P_{\mathrm{E}}}\right)+G\cdot\left(\frac{d_{\mathrm{G}}^{i}}{P_{G}}\right)\right) \quad (15)$$

$$s_{\mathrm{B}}(\tau)=\frac{\sum_{i=1}^{I} s_{\mathrm{B}}^{i}(\tau)}{I} \quad (16)$$

$$s_{\mathrm{P}}=\frac{\sum_{i=1}^{I} s_{\mathrm{P}}^{i}}{I} \quad (17)$$

where the variables are defined in Table 3. Variables $d_{\mathrm{B}}^{i}(\tau)$, d_{E}^{i} and d_{G}^{i} come from the Bayesian network. We start the analysis with the first datapoint of the first shift for each taxi and assume the charge level of the vehicle is full. At each datapoint (GPS reading), we update the total distance driven by the taxi so far, and query the Bayesian network for the charge level of the vehicle. Assuming we are analyzing PHEVs, if the charge level ever reaches zero, then d_{E}^{i} is the distance driven to that point and d_{G}^{i} is the distance driven throughout the remainder of the shift. If we are analyzing BEVs, if the charge level reaches zero, $d_{\mathrm{B}}^{i}(\tau)$ is the distance driven to that point (then Eqs. 12, 13 must be used to compute the revenue losses).

Switching station infrastructure

Due to BEV range limitations, battery switching is necessary for BEVs to be feasible for use by taxi companies. Battery switching allows drivers to have a fully charged battery within minutes. This mitigates the range limitations of BEVs, assuming there are enough switching stations to service the taxi fleet. Switching stations have a large upfront cost—the infrastructure cost is estimated to be \$500,000 by Better Place, a manufacturer of EV switching infrastructure (Galbraith 2009; Yarow 2009). This does not take into account the cost of real estate in a given area.

To provide an adequate coverage area, the fleet may need to be served by several switching stations spread across a city. Given the expense of switching stations, we want to find the minimal number and optimal location of stations to supply the fleet *without wasting money on buying unnecessary stations*. This problem can be stated as an optimization problem: given a set of taxis and the mobility data, find the optimal location(s) for switching stations such that the taxi company's profits are maximized. We formally define this problem and present an algorithm to find $c_{\mathrm{BSS}}(\tau)$ in Appendix B.

The switching station location problem is a generalization of the well-studied facility location problem [see, e.g., (Korte and Vygen 2006; Drezner and Hamacher 2002)]. The facility location problem is NP-hard, so theoreticians believe it to be a computationally difficult problem, and by extension, the switching station location problem is also computationally difficult. Details of the NP-hardness reduction are given in Appendix B.

Battery and extra vehicle costs

This section presents the computation of the cost of batteries ($c_{\mathrm{EB}}(\tau)$) and additional PHEVs (c_{EP}). Taxis should start each shift with a fully charged battery. This requires purchasing extra batteries to be kept at each switching station (BEVs) or storing extra PHEVs at the headquarters (PHEVs).

Using Little's law (Kleinrock 1975), $c_{\mathrm{EB}}(\tau)$ and c_{EP} can be computed on a per taxi basis as follows:

$$c_{\mathrm{EB}}(\tau) = \sum_{j=1}^{n} q_j(\tau) \cdot B_{\mathrm{C}} \tag{18}$$

$$q_j(\tau) = \lambda_j(\tau) \cdot B_{\mathrm{F}x} \cdot \left(\frac{C_{\mathrm{B}} - l_j(\tau)}{C_{\mathrm{B}}} \right) \tag{19}$$

$$c_{\mathrm{EP}} = \lambda_H \cdot R_{px} \cdot P_{\mathrm{C}} \tag{20}$$

where the terms are defined in Table 4. Equation (18) multiplies the number of batteries needed at switching station i (per taxi) by the cost of each battery, and sums over all needed switching stations.

Note that $l_j(\tau)$ and $\lambda_j(\tau)$ are proportional to τ. If τ increases, batteries are switched with higher remaining capacity and take less time to charge. As τ decreases, batteries are switched with lower remaining capacity and take more time to charge.

We assume additional PHEVs are kept only at the headquarters and drivers only switch PHEVs at the end of their shifts. We are not considering storing and charging PHEVs at the

Table 4 Variables used in calculating battery and additional vehicle costs

Name	Description
n	The number of switching stations needed
$q_j(\tau)$	Average no. of batteries needed at j per taxi
$\lambda_j(\tau)$	Average no. of battery switches at j per taxi per day
$l_j(\tau)$	Average remaining charge level (kWh) of batteries switched at j
λ_H	Rate at which PHEVs return to headquarters (PHEVs/day)
$B_{\mathrm{F}}x$	Time (days) it takes to charge a fully depleted battery assuming level x charging
C_{B}	The capacity of each battery (kWh)
$R_{\mathrm{P}}x$	PHEV charging rate (days/PHEV) assuming level x charging
B_{C}	The cost of one battery (dollars)
P_{C}	Full cost of a PHEV (dollars)

BEV battery switching stations. This is because it is less expensive to store batteries than vehicles—batteries can be stacked and stored in the same building but vehicles require expensive real estate for parking.

Optimal switching threshold

The optimal value of τ is unknown. We cannot find the optimal value of τ by optimizing $c_{\mathrm{EB}}(\tau)$, $s_{\mathrm{B}}(\tau)$, or $r_{\mathrm{L}}(\tau)$ alone because this will not globally maximize Δ_{B}. Therefore we numerically evaluate $\Delta_{\mathrm{B}}(\tau)$ for each value of τ in the set $\{10, 20, \ldots, 100\ \%\}$ and choose the value of τ that maximizes $\Delta_{\mathrm{B}}(\tau)$.

San Francisco case study

We applied our process to a data set collected by Yellow Cab San Francisco (YCSF) as part of the Cabspotting project (Crawdad 2009; The Exploratorium, 2008). San Francisco is a regulated taxi market. The San Francisco Municipal Transportation Agency sets the fare price for all taxi companies in the region. In addition to controlling fares, San Francisco is a *medallion based market*—the agency also controls when and how many new taxis can be integrated into the region. When this agency votes to allow new taxis, medallions are sold to companies allowing them to operate one new taxi; these medallions are now (as of October 2012) sold for $300,000 (Bay City News 2012). Interestingly, in 2011 two medallions were issued specifically to operate battery switched electric taxis (CBS San Francisco 2011).

This section is outlined as follows. In "Dataset and preprocessing" we discuss our dataset and our preprocessing of the data. In "Clustering locations" we discuss clustering GPS coordinates into a finite set of locations. We discuss assumptions specific to our case study in "Assumptions For BEV/PHEV Revenue Analysis". In "Clustering locations" we detail what EV adoption scenarios we examine. "Existing Taxi Revenue" gives the fare price structure ($r_{\{\mathrm{FARE}\}}$) for YCSF. The BEV and PHEV revenue analyses are discussed in "Revenue Analysis for BEVs" and "Revenue analysis for PHEVs" respectively. These two revenue analyses are compared in "PHEV versus BEV comparison". Finally, we discuss how relevant results from YCSF would be to other taxi companies in "Sensitivity analysis".

Dataset and preprocessing

The dataset includes the following information for 536 YCSF taxis during May 17, 2008–June 10, 2008. Each measurement includes:

- Latitude and longitude to five decimal places
- Whether a paying passenger is inside the vehicle
- The current time of the data point

The average timestep between each data point is 60–90 s. As discussed in "Inputs", we split the data from each taxi into shifts.

GPS devices sometimes report erroneous data, so we preprocessed the dataset to remove inconsistencies. For example, in some cases a taxi's position would be incorrectly reported between two correct readings. We noticed data from a taxi was either >99.5 % correct or very erroneous due to a faulty GPS device in that taxi. For the taxis that only had few erroneous points we simply removed those points, whereas the taxis with many problems

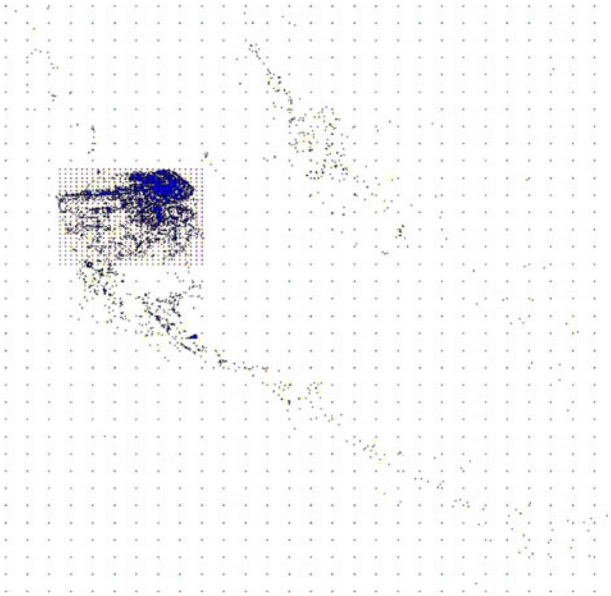

Fig. 2 *Points* represent taxi mobility data. The *grid* shows the reduced coordinate space. We used a *denser grid* in downtown San Francisco due to the large number of data points in this region

were simply discarded and excluded from all results. We discarded all data from seven out of 536 taxis.

Clustering locations

We clustered the GPS coordinates in our data using the reduced coordinate space displayed in Fig. 2. The clustering locations are spaced 4 km apart with the exception of downtown San Francisco. For the downtown area, we used a denser grid (1 km × 1 km) because of the higher density of data within this region. After collapsing each GPS datapoint into its closest grid point, the GPS data was discarded.

Assumptions for BEV/PHEV revenue analysis

Here we state the assumptions we made while performing the revenue analysis.

- EVs may use up to 40 % of their battery for AC while it is on (Farrington and Rugh, 2000). We assume that ACs doubled in efficiency since 2000 and taxi drivers in San Francisco use AC 50 % of the time—hence we assume 10 % of the battery is used for AC.
- We earlier quoted the price of a Better Place battery switching station to be $500,000 (Yarow 2009). We also use their current estimates for battery prices. The company executives state "EV batteries are approaching $500 a kWh" and "[Better Place] is now purchasing batteries for cars at $400/kWh for delivery in early 2012" (Hensley et al. 2010). We assume Nissan Leaf batteries (to be kept at switching stations) cost $450/kWh for a total of $11,000 based on these estimates. We also note several sources

indicate battery prices are likely to continue decreasing (Boulanger et al. 2011; Hensley et al. 2009; Kamat 2009).

- We assume the companies existing taxis have an efficiency of 25 mpg (Scientific American 2009).
- Our data set indicates that each taxi is driven 112,000–177,000 kms per year. Based on this figure, we assume the company replaces each SV, BEV, PHEV, and battery after 4 years of use ($L = 4$). (This does not include replacing parts over the 4 years). Companies do not disclose their vehicle replacement rates, which makes estimating this figure difficult. However, the The Taxi and Limousine Commission of New York City states "Cars brought into service as taxicabs must be brand new vehicles and generally must be replaced five years after being placed into service. (Schaller Consulting 2006).
- We assume the company currently spends $15,000 for a new SV when replacing an old SV. We use this figure to compute the incremental cost to purchase an EV instead of an SV. YCSF does not provide this figure thus we estimate it based off the vehicles listed on their website. We do not know how much the company actually pays for their vehicles due to bulk discounts.
- On May 12th 2011, the average gas price in San Francisco was $1.08/l, and electricity was $.22/kWh (GasBuddy 2011a; US Bureau of Labor Statistics 2011b). Although we extend our revenue analysis to a wide array of fuel prices, our discussion is based on these prices.
- We assume that switching infrastructure lasts 15 years.

Case study EV scenarios

We studied ten different scenarios which are as follows:

- (1–2) BEVs with Level 1 and 2 roadside charging only
- (3–4) BEVs with Llevel 1 and 2 roadside charging and battery switching
- (5) BEVs with only battery switching
- (6–7) PHEVs with Level 1 and 2 roadside charging only
- (8–9) PHEVs with Level 1 and 2 roadside charging and PHEV switching at YCSF headquarters
- (10) PHEVs with PHEV switching at YCSF headquarters only

However, we found that only scenarios five and ten were interesting. Taxis in our case study were parked only 12 % of the time and were constantly driving the other 88 % (we note this may be atypical, but this is indicated by our dataset). Even when we assume level two roadside charging is available *everywhere* in San Francisco (an extremely unrealistic assumption), the results changed by less than 15 % for both PHEVs and BEVs. Thus, for BEVs we only show results for scenario 5, and for PHEVs we show results for only scenario 10.

Existing taxi revenue

As discussed in "Inputs", each company has their own fare pricing model. For YCSF, $r_{\mathrm{FARE}}(f)$ is given on their website (Yellow Cab San Francisco 2011):

$$r_{\mathrm{FARE}}(f) = 3.10 + .45 \cdot (p + (d - .2)) + 2\delta \tag{21}$$

where d is the distance of the trip in miles, p is the time the taxi was parked (at traffic lights) during the fare, and δ is one if the passengers' destination was the airport and zero otherwise (the company charges an airport surcharge fee).

Revenue analysis for BEVs

We now compute $\Delta_B(\tau)$ given current prices and vehicle specifications using Eqs. 8 and 9. First, we show how we compute the revenue losses, r_E, in "Existing Taxi Revenue". The cost of the BEV we study, C_B, is derived in "Nissan leaf specifications". In "Switching station location and distribution over locations" we derive the cost of the battery switching stations, $c_{BSS}(\tau)$, and show its relationship to $r_L(\tau)$. In "Switching threshold analysis" we measure the relationship between the threshold τ, $r_L(\tau)$ and the cost of extra batteries needed, $c_{EB}(\tau)$. Roadside charging is briefly discussed in "Case study EV scenarios". We incorporate $r_L(\tau)$, the fuel savings $s_B(\tau)$, and $c_{EB}(\tau)$ into "Overall BEV transition cost which shows the overall return on investment we are after, $\Delta_B(\tau)$.

Nissan leaf specifications

For our BEV experiments, we study the Nissan Leaf, a consumer available BEV. We note that the the Nissan Leaf does not have a switchable battery, but we expect vehicles with similar attributes with switchable batteries to be sold soon (Bloomfield 2012). The price of the Nissan Leaf in California is \$33,720 (Nissan 2011a). However, current federal tax rebates for the purchase of electric vehicles provide a tax credit of \$7,500 for the Leaf in California(Cornell University Law School 2010). Therefore, each Leaf can be purchased for \$26,220, which is consistent with the post-tax credit price listed at (Nissan 2011a). Using our assumption that the company replaces their SVs for \$15,000, C_B = \$11,220.

Even though manufactures list the full capacity of a battery, the full capacity is not actually used—the battery is not fully charged or discharged to preserve the life of the battery (Boulanger et al. 2011). However, manufactures list the expected range based on the *usable* portion of the battery—the figure we are interested in. Table 5 gives the values of the constants needed for our revenue analysis for the Nissan Leaf. These figures were derived from the specifications given on their website (Nissan 2011b).

Switching station location and distribution over locations

We now calculate the cost of battery switching stations, $c_{BSS}(\tau)$, and show its relationship to the revenue losses incurred, $r_L(\tau)$. We find the locations of switching stations by applying the algorithm presented in Appendix B. We were able to find a global optimal solution using brute force because there are fewer than 500 locations in our data set.

With the optimal value of τ (discussed in the next section), we find that three switching stations are optimal, so $c_{BSS} = \$1,500,000$.

Table 5 Nissan leaf specifications (Nissan 2011b)

Constant	Value for Nissan leaf
D_r in Eq. 2 (kWh/km)	.15
B_{g1} in Eq. 3 (kWh/s)	.00033
B_{g2} in Eq. 3 (kWh/s)	.001
B_{f1} in Eq. 19 (days/battery)	.83
B_{f2} in Eq. 19 (days/battery)	.29
C_B in Eq. 19 (kWh)	24
C_B in Table 2	\$11,220

The relationship between $r_L(\tau)$ and $c_{BSS}(\tau)$ is shown in Table 6. Without battery switching, even if we assume charging infrastructure is available everywhere (i.e., whenever a taxi is stopped, its battery charges while it is parked), a third of all fares are lost. However, with additional switching stations at the San Francisco airport and Yellow Cab headquarters, only 3 % of fares are lost. Adding additional stations to these three has negligible impact on $r_L(\tau)$ but greatly drives up $c_{BSS}(\tau)$; three stations represents the optimal value for YCSF.

We find the distribution over all locations where fares began and ended to gain intuition as to why three stations are adequate. Figure 3 shows this distribution. We find approximately 90 % of all pick-ups and drop-offs occur in only 20 % of the locations. This explains why a small number of switching station locations suffice; switching stations near these locations will be heavily used.

Switching threshold analysis

We find the ROI $\Delta_B(\tau)$as a function of τ as discussed in "Optimal switching threshold". This threshold is an optimization between $r_L(\tau)$, $s_B(\tau)$, and $c_{EB}(\tau)$ as follows. Consider Figs. 4 and 5 which show $r_B(\tau)$ versus τ. Note the single peaked distribution. Increasing the threshold increases the taxis' average charge levels, which increases $s_B(\tau)$ and decreases $r_L(\tau)$, but at some point the threshold is too high and c_{EB} is large enough to offset these benefits. Therefore an optimal value of τ exists and we find its value by computing the ROI for all values of τ.

We emphasize that we do not allow taxis to deviate from the routes in the data set. They only switch batteries if they are at a location with a switching station and their threshold is less than τ. In practice, taxi drivers would be likely to actively monitor their battery charge level and travel to switching stations when needed to avoid depletion. Thus, our analysis is conservative.

Overall BEV transition cost

Figure 6 shows the cost to transition each SV to a BEV (r_B) for a wide array of gas and electricity prices. We see that at current prices, BEVs are more profitable than SVs in San

Table 6 Percentage of fares lost in different BEV scenarios

No charging or switching	41.5 %
L2 roadside charging only	37 %
Union Square BSS (no charging)	15 %
YC, Union Square, Airport BSS	3 %

BSS battery switching station(s)

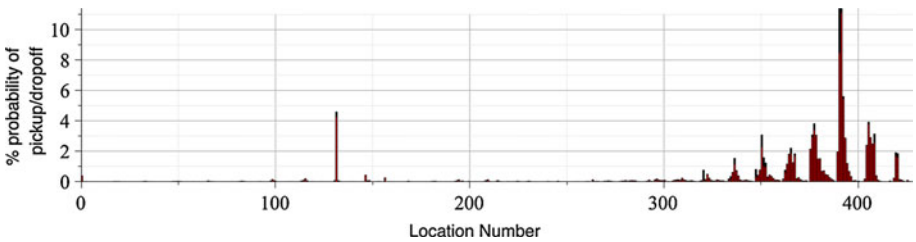

Fig. 3 Distribution over pick-up and drop-off locations as distribution. *Black* shows the probability (percentage) of a pickup, *red* shows the probability of a dropoff, but the two distributions are nearly identical. (Color figure online)

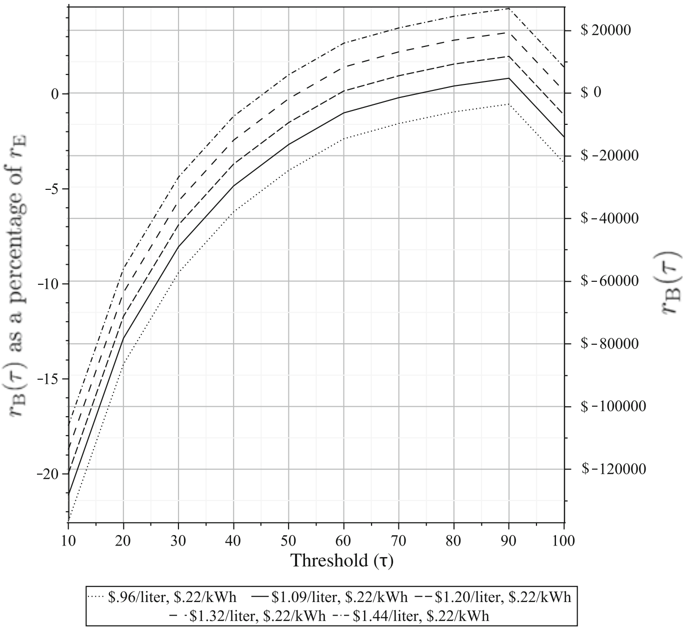

Fig. 4 Switching threshold τ versus $r_B(\tau)$ as a percentage of r_E, $r_B(\tau)$ for varying gasoline prices, fixed electricity price of $.22/kWh

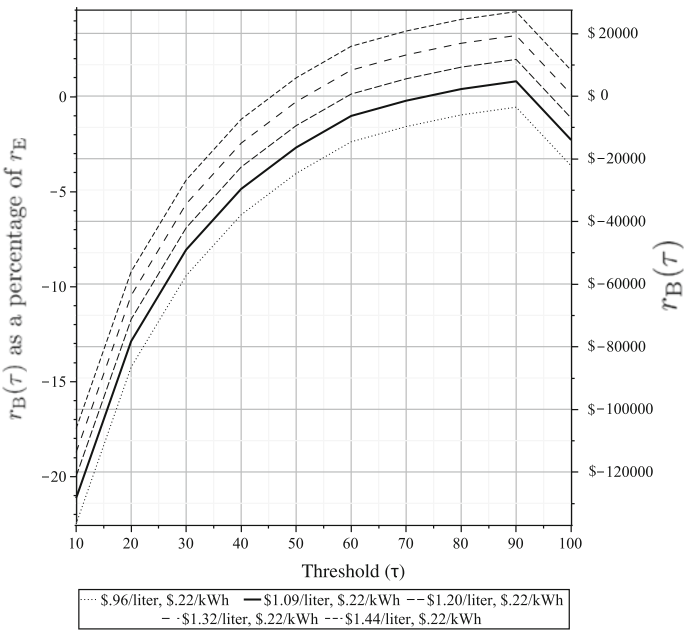

Fig. 5 Switching threshold τ versus $r_B(\tau)$ as a percentage of r_E, $r_B(\tau)$ for varying electricity prices, fixed gas price of $1.08/l

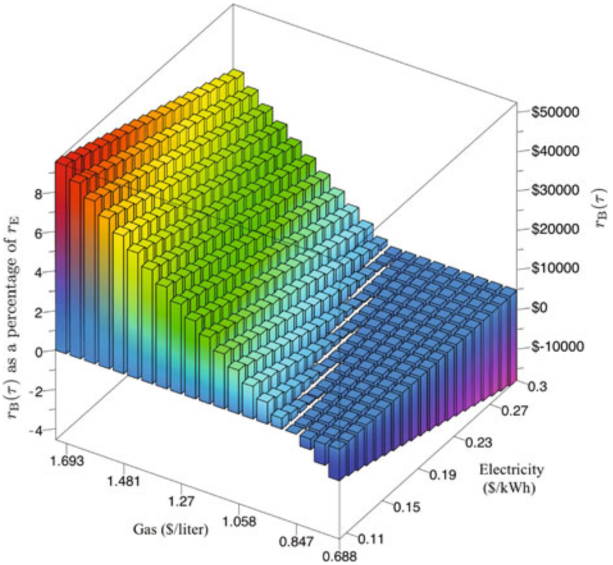

Fig. 6 Gas and electricity prices versus $r_B(\tau)$ as a function of r_E, $r_B(\tau)$. Three switching stations

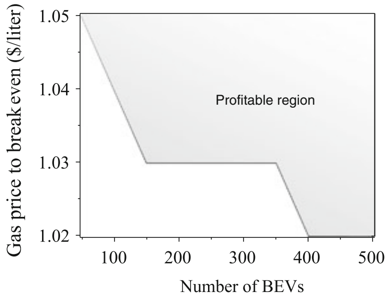

Fig. 7 Profitable region for a given BEV penetration, gas price. Electricity fixed at $.22/kWh

Francisco. We find $r_B \approx \$4100$, which is $\approx 0.68\ \%$ of the company's existing revenue r_E. Because current prices of gasoline and electricity may vary, we determine the gasoline price for which they become profitable. For a fixed electricity price of $.22 kWh, the price point where BEVs are exactly at parity with petroleum vehicles, without considering the cost of the switching stations ($c_{BSS}(\tau)$) is $1.02/l. With gasoline prices above this point, the company can pay back $c_{BSS}(\tau)$.

We now compute $\Delta_B(\tau) = r_B x - c_{BSS}(\tau)$. We are specifically interested in the point where $\Delta_B(\tau) = 0$; this represents the "break even" point where x BEVs can be operated for the exact cost that x SVs can. Under our assumption that switching infrastructure lasts 15 years, we amortize the $1.5M cost of three switching stations accordingly, yielding a cost of $100k/year. Figure 7 shows the when $\Delta_B(\tau) = 0$ for a fixed electricity price of $.22/kWh. For a given point on the line, if gas prices rise, the company accrues profits. We note that gas prices in San Francisco are currently above all points on this line ($1.08/l).

Revenue analysis for PHEVs

We now compute the costs of switching to PHEVs instead of BEVs and compare the two scenarios. We therefore derive C_P and charging rates next, and we find Δ_P in the following section.

Vehicle cost and specifications

For our PHEV analysis, we study the four cylinder Chevrolet Volt. The Chevrolet Volt has a retail cost of $41,000 (Reuters 2010). After the $7,500 tax credit, the price is $33,500. Using the $15,000 taxi replacement figure, $C_P = \$18{,}500$. Table 7 gives the values of the constants needed for our revenue analysis for the Chevrolet Volt. These figures were derived from the specifications given on their website (Chevrolet 2011).

Overall PHEV transition cost

Figure 8 shows the ROI in PHEVs, r_P, without any roadside charging. As with BEVs, at current prices PHEVs are less expensive to operate than SVs. We find r_P corresponding to current prices of $1.08/l and $.22 kWh is $\approx$ $3400, which is $\approx$0.57 % of r_E.

For PHEVs there is no fixed cost investment, so Δ_P can be derived by multiplying r_P by any given x to get the cost of transitioning x SVs to PHEVs. Therefore, the company can switch to PHEVs on a per vehicle basis. However, in the next section, we discuss why BEVs have more advantages than PHEVs.

PHEV versus BEV comparison

Even though PHEVs and BEVs are both currently profitable, BEVs are a likely better investment. Gas prices are volatile, and in the past three years we have seen the two highest prices ever for a liter of gas in the United States. Figure 9 shows gasoline prices per gallon (one gallon = 3.78 l) since 1971 adjusted for inflation (Zfacts 2010). In contrast, electricity prices have not been volatile; Fig. 10 shows the average electricity price *in California* (adjusted for inflation) since 1980 (Tom and Kurt 2007). Note that Fig. 10 does not show a price of $.22/kWh electricity (the figure used throughout this paper) because electricity prices in San Francisco are nearly double than in the rest of California and the US average, as shown in Table 8 (US Bureau of Labor Statistics 2011a). We could not find a long history of electricity prices in San Francisco alone. Because most PHEVs are still about 60 % petroleum based (for example, the Volt uses electricity for 40 % of its useable range

Table 7 Chevrolet Volt specifications (Chevrolet 2011)

Constant	Value for Chevrolet Volt
D in Eq. 2 (kWh/km)	.25
B_{g1} in Eq. 3 (kWh/s)	.0004
B_{g2} in Eq. 3 (kWh/s)	.0011
R_{p1} in Eq. 19 (days/PHEV)	.45
R_{p2} in Eq. 19 (days/PHEV)	.16
C_B in Eq. 19 (kWh)	16
C_P in Table 2	$18,500

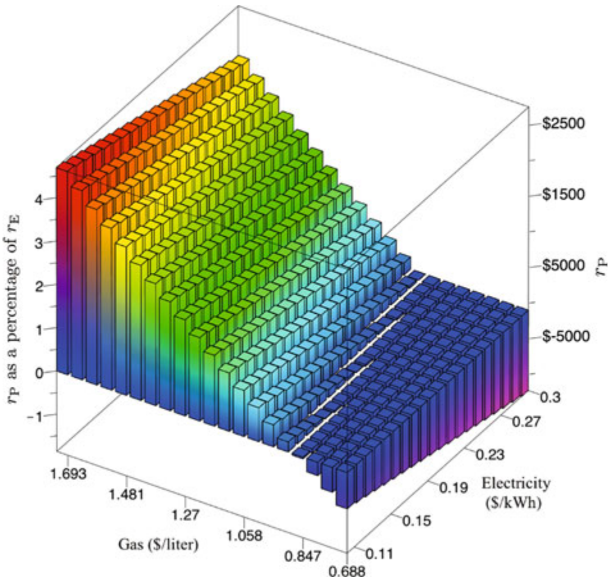

Fig. 8 Gas and electricity prices versus r_P as a function of r_E, r_P. No roadside charging

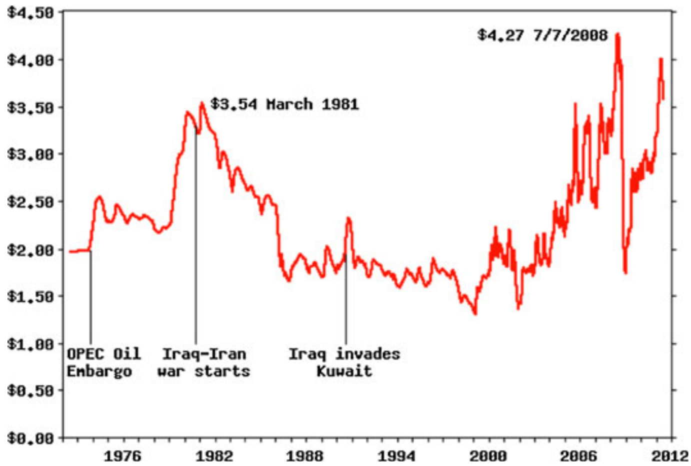

Fig. 9 Average US price per gallon of gas since 1971 (2011 dollars) (Zfacts 2010)

(Chevrolet 2011), if both of these trends continue as they have for the past 30+ years, BEVs are a better investment. Furthermore, battery prices are decreasing and are expected to continue decreasing (Boulanger et al. 2011; Hensley et al. 2010, 2009; Kamat 2009; Kanellos 2010). Because PHEVs are more expensive due to their ICEs, and the majority of the price of a BEV is its battery, BEVs are expected to decrease in price faster than PHEVs. Finally, if our ultimate goal is *complete petroleum independence*, PHEVs can only be used as an interim solution and would need to be replaced.

Sensitivity analysis

We now analyze whether our case study results would generalize to taxi companies in different cities. Although we cannot draw definite conclusions without re-running our study for a taxi company in a different city, we attempt to answer this question here.

1. *Average Trip Length and City Density*. Switching infrastructure is expensive, so it will initially be sparsely deployed, in contrast with current petroleum infrastructure. Consequently, the geography of a city affects the feasibility of BEVs. Large cities with widespread points of interest are less suitable than dense cities with concentrated points of interest. One way we can measure this for a given city is to determine the distribution over fare trip lengths. We can use the distribution of how far people commonly travel as a heuristic to estimate how many switching stations will be needed. Figure 11 shows the distribution of trip lengths for all fares the YCSF taxis completed during the study period. From the cumulative density function of this distribution, we find 85 % of all fares are less than 10 km, which is why few stations are needed in San Francisco. This figure shows a two-peaked distribution. From the probability mass function, we find 8 % of the fares are between 20 and 30 km (roughly 7 % of all trips are to the San Francisco International Airport, which is 24 km by highway from Union Square in downtown San Francisco).
   The following factors are important to determine the ROI of EVs in a region.
   The average trip length can also help us determine whether PHEVs or BEVs are better suited for a region. For a PHEVxxm, it does not matter how many trips are completed before battery depletion, because the financial benefit comes from the transportation savings on the first xx km. BEVs are range limited, however, and completing a large number of short trips (before battery depletion) is more profitable than a short number of long trips, due to the initial charge to each passenger that requests a fare. Therefore, PHEVs are likely better suited for cities with many long trips, whereas BEVs will be more profitable in cities with a large number of short trips, like San Francisco.
2. *Distribution over locations*. Closely related to the distribution over trip lengths is the distribution over locations discussed in "Switching station location and distribution over locations". We found roughly 90 % of YCSF fares start or end in fewer than 20 % of the grid locations. If this distribution was less concentrated, the average trip length

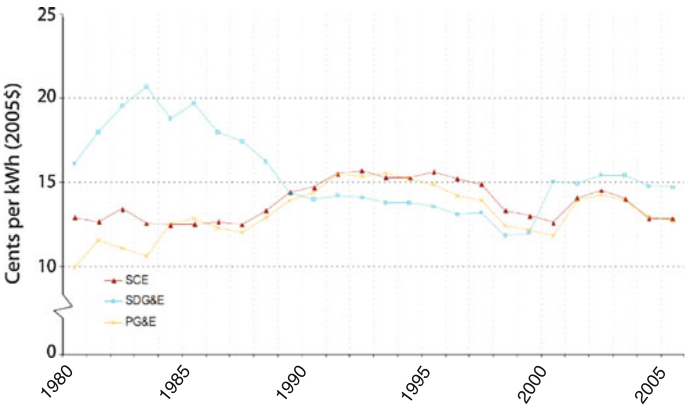

Fig. 10 Average electricity price given by three major utilities in California since 1980 (Tom and Kurt 2007)

Table 8 San Francisco versus US fuel prices (US Bureau of Labor Statistics 2011a)

May 2011: San Francisco electricity prices versus US average			
Item	US average	San Francisco	Percent difference
Electricity ($ per kWh)	0.129	0.226	75.2
Gasoline ($ per liter)	1.06	1.12	5.4

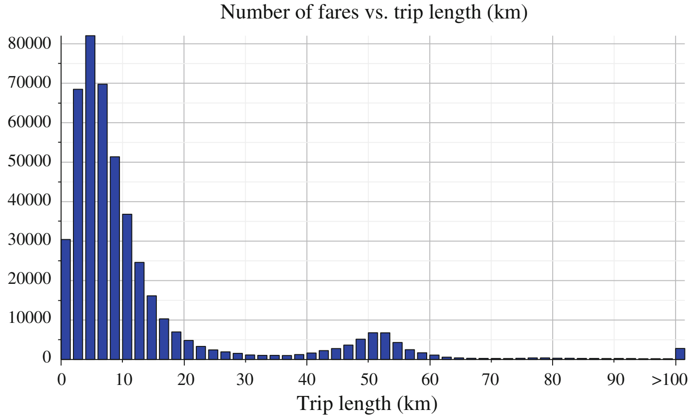

Fig. 11 Distribution of fare trip lengths (the *bar* for 50 represents all trips over 50 km)

shown in Fig. 11 may have increased. In our case study, trips in the downtown area within 3 km of Union Square accounted for more than half of all trips. This is highly conducive to centralized switching station placement. Taxis in larger cities may find they have to travel a greater distance out of their way to refuel.

3. *Gas and electricity prices*. Although current gas prices in San Francisco ($1.08/l) are higher than the rest of the United States, they are lower compared to the rest of the world. For example, the average price in London, England is $1.39/l (The Automobile Association 2011), and the average price in Toronto, Canada is $1.29/liter (GasBuddy 2011b). If the mobility patterns of taxis in these regions are similar to those in San Francisco, transitioning to EVs would be even more profitable. We also note that electricity prices in San Francisco are twice the United States national average, while gas prices are not (United States Department of Labor 2010).
4. *Temperature and weather*. Reference (Shidore and Bohn 2008) shows that at cold temperatures (<32 °F), over 10 % of the energy in a battery is lost compared to at 68 °F. It rarely snows or drops below freezing in San Francisco, even during the winter months, but taxi companies in cities with colder climates should expect worse performance. Furthermore, passengers in cities with extreme weather temperatures require more heating and cooling, which further drains the battery.

Limitations and future work

To obtain realistic results for our case study, we used only commercially available vehicles and their manufacturer specifications. However, predicting the outcome of a major transition prior to it occurring is an error-prone process. We now discuss some avenues for future work.

1. By introducing a stochastic model instead of a discrete Bayesian model, most of the limitations in "Model assumptions" could be resolved. We note however, that this is a large piece of future work and may not give much more insight into the problem.
2. Our process can only be used with taxi companies whose vehicles are brought back to a common location after each driver's shift. Future work could generalize the process to different types of taxi companies.
3. We have not included any analysis of vehicle maintenance costs. Maintenance costs for a fleet of EVs is thought to be lower than SVs (Boulanger et al. 2011), but we are not aware of any quantitative analysis comparing the two. A maintenance cost analysis for a large fleet of PHEVs/BEVs would greatly improve our cost model.
4. Our switching station optimization assumes that the locations can charge any number of batteries and can be placed anywhere in the city. In reality, distribution network limitations may place some restrictions on switching station placement and battery charging; areas with a fully utilized distribution network may not be able to accommodate the new load.
5. We have not considered real estate prices for switching stations, other than the cost of the stations themselves. We should account for the cost of acquiring space to build the switching station.
6. Obtaining a second data set from a different city would provide a better foundation for the sensitivity analysis section.
7. Batteries do not charge at a constant rate as we have assumed. A better assumption would be to use a two phase linear approximation; have a higher charge rate while the state of charge (SOC) is less than 80 %, and a lower rate when the SOC is above 80 % (Nissan 2011b).
8. Queueing delays at switching stations due to multiple taxis attempting to switch their batteries at once should be modeled.

We note that EVs are still manufactured using petroleum; the study of the overall petroleum use of a vehicle, including manufacturing is known as *life cycle* analysis (Elgowainy et al. 2009; Samaras and Meisterling 2008).

Conclusions

In this paper, we proposed a process to determine the ROI for a taxi corporation transitioning to electric vehicles. We first built a model of taxi fleet transportation, and then used the model to compute the economic costs of the transition. The model can be configured with a wide array of input parameters, including the type of vehicle to be tested, electricity and gasoline prices, and roadside charging/battery switching infrastructure assumptions. We then used our process to analyze a fleet of over 500 taxis in San Francisco. We found that PHEVs and BEVs are both currently profitable.

Electric vehicles are expected to play a large role in reducing petroleum consumption and global carbon emissions. The transition to EVs is a necessary but will likely be a difficult transition, but we can mitigate the negative effects of major transitions, such as a complete overhaul of the transportation industry, with careful planning. Careful planning requires analysis of several aspects of the transition, including financial feasibility and social factors. Our work presents a step towards providing information to taxi companies as to the extent, either positively or negatively, to which they might be affected.

Acknowledgments The authors would like to acknowledge Jason Baek, Augustin Chaintreau, Earl Oliver, Lisa Patel, and Dr. Catherine Rosenberg for their assistance with this research.

Appendix A: Brief CLGN background

In this appendix we provide a brief background on CLGNs. We assume knowledge of standard Bayesian networks; an excellent reference text is (Koller and Friedman 2009).

We first present some necessary definitions.

- The *graphical model* of a problem is a directed acyclic graph $G(V, E)$, where each vertex is a variable and each edge represents a causal effect. The variables may be known (we can directly observe or compute their values) or hidden (we estimate their value because we cannot observe their values directly).
- The set of parents $Pa(X)$ of a node X in a graphical model is defined as all nodes Y such that $(Y, X) \in E$ and $Y \neq X$.
- Variables can be either *discrete* or *continuous*; discrete variables can only take values from a countable set of values, such as the integers, whereas continuous variables can be any real number.
- A *Bayesian network* is a directed acyclic graph that defines the relationship $P(X|Pa(X))$ between every variable and its parents. The probability of any variable X is independent of all other variables in the network given its parents.

Hybrid models

Standard Bayesian networks contain only discrete variables. A *hybrid* model contains a mix of both continuous and discrete variables. Several different hybrid models exist; we chose to use conditional linear Gaussian networks (CLGNs). In linear Gaussian models, each variable X is modeled as a linear combination of its parents. CLGNs are extensions of linear Gaussian models that allow for both discrete and continuous variables.

In CLGNs, three types of relationships are defined:

- A discrete child with only discrete parents
- A continuous child with only continuous parents
- A continuous child with a mixture of continuous and discrete parents.

Note that CLGNs do not allow for discrete variables with continuous parents. Other models address this issue, but we do not need these extensions for our application.

Querying conditional linear gaussian networks

To query a variable is to return its Gaussian distribution. To query each of the three types of variables, we use the following formulas.

We consider the simplest case first; a discrete variable with only discrete parents. To express this conditional relationship, we use a discrete conditional probability table (CPT) as in standard Bayesian networks.

Next we consider a continuous variable with only continuous parents. A continuous variable X can take on any real number in the domain of X. Therefore, we cannot have a finite CPT because it would be infinitely large. Instead, we maintain a function of its parents' values that is used to generate a Gaussian over X. Let X have k parents with means $pa_1, pa_2, \ldots pa_k$. Under the CLGN model, we specify $k + 2$ parameters $\alpha_0, \alpha_1, \ldots, \alpha_k$, and a variance σ^2 and compute $P(X|pa_1, pa_2, \ldots pa_k)$ as

$$P(X|pa_1, pa_2, \ldots pa_k) = \mathcal{N}\left(\alpha_0 + \sum_{i=1}^{k} \alpha_i(t) \cdot pa_i(t), \sigma^2\right) \tag{22}$$

That is, the set of αs are linear combination constants; we are calculating a new Gaussian that is a linear combination of other Gaussians (its parents).

Before we examine the third case, we note how σ^2 is obtained. There are two widely used versions of CLGN's: those where the variance of each variable depends on the variances of its parents, and those where the variance is assumed to not depend on its parents (Koller and Friedman 2009). We are using the latter simpler model because we do not have data for the variables in Table 1. This model is not as accurate because it ignores covariance between variables and their parents, but is commonly used when the variances of variables in the network are not known, and still captures most of the meaningful relationships (Koller and Friedman 2009). Because we use this model, we do not present background on CLGNs where the variance of each variable X depends on $Pa(X)$; but note these models rely on the theory of *multivariate Gaussian distributions*.

Finally, we consider the most complex case, a continuous variable with both continuous and discrete parents. Let X be a continuous random variable with j discrete parents and k continuous parents. Let $\mathbf{D} = \{D_1, \ldots, D_j\}$ represent the discrete parents of X. Let $\mathbf{C} = \{C_1, \ldots, C_k\}$ represent the continuous parents of X; we denote the mean of the ith continuous parent c_i. Together, $\mathbf{D} \bigcup \mathbf{C} = Pa(X)$. For every combination $\mathbf{d}$ chosen from $\mathbf{D}$, we have a (possibly different) vector of $k + 2$ constants $\alpha_{d_0}, \alpha_{d_1}, \ldots, \alpha_{d_k}, \sigma_d^2$, and a variance $\sigma_d{}^2$ such that

$$P(X|\mathbf{D} = \mathbf{d}, \mathbf{C} = \mathbf{c}) = \mathcal{N}\left(\alpha_{d_0} + \sum_{i=1}^{k} \alpha_{d_i} \cdot c_i; \sigma_d^2\right) \tag{23}$$

Again, the set of αs are the linear combination constants.

The problem with this approach is that the set of all combinations of $\mathbf{d}$ may be massive; even if each discrete parent was binary, we would still have 2^d combinations and would need to store $(k + 2)2^d$ constants *for every variable*. A better idea is to store a function for each variable that calculates these $k + 2$ constants based on its parents at any time. This also allows us to set the α values based on $X's$ discrete *and* continuous parents, if needed. Therefore, we introduce a function $\phi_X(\mathbf{d}, \mathbf{c}) : \mathbb{R}^{j+k} \rightarrow \mathbb{R}^{k+1}$. This function ϕ_X takes in *all* of $Pa(X)$ and generates the α values used in the linear combination. Creating the ϕ functions requires knowledge of the problem; we need encode our knowledge of how variables are dependent upon each other into the network via the ϕ functions. If we were instead storing the constants, then we would need to derive the constants for each variable based on our knowledge of the problem.

Having introduced the ϕ_X function, we rewrite Eq. 23 as:

$$P(X|\mathbf{D} = \mathbf{d}, \mathbf{C} = \mathbf{c}) = \mathcal{N}\left(\alpha_{d_0} + \sum_{i=i}^{k} \alpha_i \cdot c_i; \sigma_{\mathbf{d}}^2\right) \tag{24}$$

$$\{\alpha_0, \ldots, \alpha_k\} = \phi_X(\mathbf{d}, \mathbf{c}) \tag{25}$$

Whenever we query a variable, we use Eqs. 24 and 25 to calculate the distribution.

Appendix B: General switching station optimization algorithm

As discussed in Related Literature, placing facilities to maximize the number of miles travelled or maximize the number of intercepted flows does not necessarily maximize revenue. We assume the taxi company's objective is to maximize their overall revenue, and therefore introduce a new optimization framework based on the discretized locations of the taxis and their charge levels. It is also a variation of the flow based facility location model.

We now provide the details of our approach to computing locations for switching stations. First, we show that the problem is NP-hard, which implies that it is unlikely to be able to be solved by an algorithm that runs in polynomial-time. Then, we formulate it as an integer program, and propose an algorithm for the problem that works on small instances.

We outline a proof that the switching station location problem is NP-hard by a reduction to the facility location problem. The facility location problem is stated as: given a set of clients has some demand from a facility and a cost to build each facility, find the optimal placement of facilities to minimize the cost of the facilities and the cost of serving the clients. The switching station location problem can be reduced to the facility location problem by treating the taxis as clients whose demand varies over time and the switching stations as the facilities that can meet that demand. Therefore, finding the set of optimal switching station locations is also NP-hard.

We now formally describe the switching station location problem. First, we introduce the necessary notation. Let L be the set of locations where a switching station can be placed. We denote a taxi by x and the set of all taxis by X. We assume knowledge of a cost function cost(l) for each location $l \in L$ that is the price of placing a station at l. We use $\mathrm{Loc}_t(x)$ to be the location of x at time t and $\mathrm{fare}_t(x)$ is True when x has a passenger and False when it does not. We use a binary variable $y(l)$ to indicate if a location has been selected for a switching station. The charge level of $x \in X$ at timestep t_k is denoted by $\mathrm{CL}(x, t_k)$. Let $\mathrm{o}_t(x, r)$ be the opportunity cost of an EV with charge level r at time t. For a taxi x, $\mathrm{o}_t(x, r)$ should be zero when x's battery is sufficiently charged; however, as its charge level drops, there is some opportunity cost because the driver will not be able to complete trips over some length, and thus may lose revenue because some passengers cannot be transported to their destination. In our analysis, we define $\mathrm{o}_t(x, r)$ to be the sum of taxi x's fares for the remainder of its shift, once it cannot complete a trip because its charge level r is too low. That is, if x cannot complete a trip at time t, then its opportunity cost is the fares for the trips it would have completed from time t until the end of its shift. Finally, we use tau to be the battery level at which a taxi will always swap its battery if is at the same location as a switching station.

The objective of our optimization problem is stated as *given the set of taxis and their temporal mobility patterns, find the optimal location(s) for switching stations such that the taxi company's profits are maximized.* Our mathematical formulation of the switching station location problem is as follows:

$$\min \sum_{l \in L} \mathrm{cost}(l) \cdot y(l) + \sum_{x \in T} \sum_{t} \mathrm{o}_t(x, \mathrm{c}_t(x)) \tag{26}$$

subject to:

$$y(l) = \{0,1\} \forall l \in L$$

$$CL(x,t_k) = \begin{cases} Full & \text{if } \mathrm{y}(\mathrm{Loc}_t(x)) = 1 \\ & \text{and } CL(x,t_{k-1}) < \tau \\ & \text{and } \mathrm{fare}_t(x) = \text{false} \\ CL(x,t_{k-1}) - u(t_{k-1},t_k), & \text{otherwise} \end{cases}$$

$$u(t_{k-1},t_k) = \text{energy used from } t_{k-1},t_k$$

when L does not contain too many locations (for example, as in our case study below), we can solve the switching station location problem optimally using brute force. That is, we find the value of Eq. 26 for all possible locations of $1,2,\ldots,k$ switching stations. The value of k is found by determining the number of switching stations sufficient so that no revenue is lost due to opportunity costs (i.e., we have $\sum_{x \in T}\sum_t \mathrm{o}_t(x, \mathrm{charge}_t(x)) = 0$). At this point, Eq. 26 is monotonically increasing when more switching stations are added, so we can safely conclude that Eq. 26 is minimized with k or fewer switching stations.

This brute force approach may not be feasible over larger areas with more locations. In this case, it is possible to use heuristic algorithms to find a solution, though these heuristics cannot guarantee the optimality of their solution. Algorithms such as simulated annealing, tabu search, and hill climbing are general optimization methods, and could be used to find approximate solutions to the switching station location problem (Glover and Laguna 1997; Kirkpatrick et al. 1983; Russell and Norvig 2003).

Our formulation of the switching station location problem relies on time series locations of the vehicles that will use the switching stations. Ideally, this location data is collected from multiple vehicles over multiple weeks; however, this data may not be obtainable in some situations. In this case, it is still possible to optimize the placement of switching stations using stochastic facility location algorithms [e.g., (Owen and Daskin 1998; Snyder 2004)]. Such algorithms are designed to optimize facility locations when there is a high amount of uncertainty in the input. These algorithms take a probability distribution of the amount of time vehicles spend at given locations as input. This distribution could be estimated from, e.g., road congestion statistics or logs of passenger pickups and drop-offs.

References

Aecom: Economic Viability of Electric Vehicles. Aecom Report (2009)

Bay City News: SFMTA Approves Controversial Taxi Medallion Plan. http://sanfrancisco.cbslocal.com/2011/09/02/111118/. Accessed 20 Oct 2012

Becker, T.A., Sidhu, I., Tenderich, B.: Electric Vehicles in the United States: A New Model with Forecasts to 2030. Center for Entrepreneurship & Technology (2009)

Berman, O., Larson, R., Fouska, N.: Optimal Location of Discretionary Service Facilities. Operations Research pp 1–36 (1990)

Berman, O., Larson, R., Fouska, N.: Optimal location of discretionary service facilities. Transp. Sci. **26**:1–11 (1992)

Better Place: Better Place to Bring Electric Taxi Program to the San Francisco Bay Area. http://www.betterplace.com/the-company-pressroom-pressreleases-detail/index/id/better-place-to-bring-electric-taxi-program-to-the-san-francisco-bay-area (2010). Accessed 27 April 2011

Bloomfield, N.G.: Better Place Unveils Europe's First Battery Switch Station in Denmark. http://www.greencarreports.com/news/1072111_better-place-delivers-electric-cars-battery-swaps-become-real (2012). Accessed 20 Oct 2012

Boulanger, A.G., Chu, A.C., Maxx, S., Waltz, D.L.: Vehicle electrification: status and issues. Proc. IEEE **99**(6):1–23 (2011)
Canizares, et al.: Towards an Ontario Action Plan For Plug-In Electric Vehicles (2010)
CBS San Francisco: San Francisco Taxi Fares Go Up. http://sanfrancisco.cbslocal.com/2011/09/02/111118/ (2011). Accessed 20 Oct 2012
Chevrolet: Chevrolet Volt Specifications. http://gm-volt.com/full-specifications/ (2011). Accessed 5 July 2011
Cornell University Law School.: US Code, Title 26,30. New Qualified Plug In Electric Drive Motor Vehicle Credit. http://www.nissanusa.com/ev/media/pdf/incentives/nissan-leaf-incentive-federal-2.pdf (2010). Accessed 5 July 2011.
Crawdad: GPS Mobility Data Set.http://crawdad.cs.dartmouth.edu/meta.php?name=epfl/mobility (2009). Accessed Oct 2010
Darovsky, et al.: Electric avenue: two case studies on the economic feasibility of the electrification of transportation. Master's thesis, Duke University (2010)
Delucchi, M.A., Lipman, T.E.: An analysis of the retail and lifecycle cost of battery-powered electric vehicles. Transportation Research Part D (2001)
Drezner, Z., Hamacher, H.: Facility Location. Applications and Theory. Springer, Berlin (2002)
Elgowainy, A., Burnham, A., Wang, M., Molburg, J., Rousseau, A.: Well-to-wheels energy use and greenhouse gas emissions analysis of plug-in hybrid electric vehicles. Center for Transportation Research, Energy Systems Division. ANL/ESD/09-2 (2009)
EPRI.: Comparing the Benefits and Impacts of Hybrid Electric Vehicle Options. EPRI Report (2001)
Farrington, R., Rugh, J.: Impact of Vehicle Air Conditioning on Fuel Economy, Tailpipe Emissions, and Electric Vehicle Range. Earth Technologies Forum, National Renewable Energy Laboratory (2000)
Galbraith, K.: Better Place Unveils Battery Swap Station. http://green.blogs.nytimes.com/2009/05/13/better-place-unveils-battery-swap-station/ (2009). Accessed 25 April 2011
Gao, H.O., Kitirattragarn, V.: Taxi owners' buying preferences of hybrid-electric vehicles and their implications for emissions in New York city. Transp. Res. Part A: Policy Pract. **42**(8):1064–1073 (2008)
García, I., Miguel, L.J.: Is the electric vehicle an attractive option for customers?. Energies **5**(1):71–91 (2012)
GasBuddy: San Francisco Gas Prices. http://www.sanfrangasprices.com/ (2011a). Accessed 12 May 2011
GasBuddy: Toronto Gas Prices. http://www.torontogasprices.com/ (2011b). Accessed 12 May 2011
Glover, F., Laguna, M.: Tabu Search. Kluwer Academic Publishers, Norwell (1997)
Hensley, R., Knupfer, S.,, Pinner, D.: McKinsey Quarterly: Electrifying Cars: How Three Industries Will Evolve. http://www.mckinseyquarterly.com/Electrifying_cars_How_three_industries_will_evolve_2370 (2009). Accessed 17 July 2011
Hensley, R., Knupfer, S.,, Pinner, D.: Green Tech Media: EV Batteries Plummet in Price: Down to $400 a kWh. http://www.greentechmedia.com/articles/read/ev-batteries-dropping-rapidly-in-price/ (2010). Accessed 17 July 2011.
Kamat, H.: California Air Resources Board: Lithium Ion Batteries for Electric Transportation: Costs and Markets. http://www.arb.ca.gov/msprog/zevprog/2009symposium/presentations/kamath.pdf (2009). Accessed 17 July 2011
Kempton, W., Tomic, J.: Vehicle-to-grid power implication: from stabilizing the grid to supporting large-scale renewable energy. J. Power Sour. **144**(1):280–294 (2005)
Kirkpatrick, S., Gelatt, C.D., Vecchi, M.P.: Optimization by simulated annealing. Science **220**(4598): 671–680 (1983)
Kleinrock, L.: Theory, Volume 1, Queueing Systems. Wiley, New York (1975)
Kliesch, J., Langer, T.: Plug-In Hybrids an Enviornmental and Economic Performance Outlook. American Council for an Energy-Efficient Economy Report (2006)
Koller, D., Friedman, N.: Probabilistic Graphical Models: Principles and Techniques. MIT Press, Cambridge (2009)
Korte, B., Vygen, J.: Combinatorial Optimization: Theory and Algorithms 3rd edn. Springer, Berlin (2006)
Kuby, M., Lim, S.: The flow-refueling location problem for alternative-fuel vehicles. Socio-Economic Plan. Sci. **39**(2):125–145 (2005)
Kuby, M., Lim, S.: Location of alternative-fuel stations using the flow-refueling location model and dispersion of candidate sites on arcs. Netw. Spatial Econ. **7**(2):129–152 (2006)
Kuby, M., Lines, L., Schultz, R., Xie, Z., Kim, J.G., Lim, S.: Optimization of hydrogen stations in Florida using the flow-refueling location model. Int. J. Hydrogen Energy **34**(15):6045–6064 (2009)
Kanellos, M.: Green Tech Media: EV Batteries Plummet in Price: Down to $400 a kwH. http://www.greentechmedia.com/articles/read/ev-batteries-dropping-rapidly-in-price/ (2010). Accessed 4 July 2011

National Research Council: Transitions to Alternative Transportation Technologies: Plug-in Hybrid Electric Vehicles. The National Academies Press, Washington (2010)

Nissan: Nissan Leaf Pricing Information for California. http://www.nissanusa.com/leaf-electric-car/incentives/show/California#/leaf-electric-car/feature/pricing_information (2011a). Accessed 5 July 2011

Nissan: The Nissan Leaf. http://www.nissanusa.com/leaf-electric-car/faq/list/charging#/leaf-electric-car/faq/list/charging (2011b). Accessed 28 June 2011

Owen, S.H., Daskin, M.S.: Strategic facility location: a review. Eur. J. Oper. Res. **111**(3):423–447 (1998)

Prud'homme, R.: Electric Vehicles: A Tentative Economic and Environmental Evaluation. International Transport Forum (2010)

Reuters: GM Sets $41,000 Price for Electric Chevy Volt. http://ca.reuters.com/article/domesticNews/idCATRE66Q4UO20100727 (2010). Accessed 4 Mar 2011

Russell, S.J., Norvig, P.: Artificial Intelligence: A Modern Approach. Prentice Hall, Englewood Cliffs (2003)

Samaras, C., Meisterling, K.: Life cycle assessment of greenhouse gas emissions from plug-in hybrid vehicles: Implications for policy. Environ. Sci. Technol. **42**(9):3170–3176 (2008)

Schaller Consulting: The New York City Taxi Fact Book. http://www.schallerconsult.com/taxi/taxifb.pdf (2006). Accessed 4 July 2011

Scientific American: Electric Cars: How Much Does it Cost per Charge?. http://www.scientificamerican.com/article.cfm?id=electric-cars-cost-per-charg (2009). Accessed 27 April 2011.

Shidore, N., Bohn, T.: Evaluation of Cold Temperature Performance of the JCS-VL41M PHEV Battery Using Battery HIL. Argonne National Laboratory, USA. http://www.autonomie.net/docs/6%20-%20Papers/CIL/evaluation_of_cold_temperature.pdf (2008)

Shukla, A., Pekny, J., Venkatasubramanian, V.: An optimization framework for cost effective design of refueling station infrastructure for alternative fuel vehicles. Comput. Chem. Eng. **35**(8):1–8 (2011)

Simpson, A.: Cost-benefit analysis of plug-in hybrid electric vehicle technology. 22nd International Battery, Hybrid and Fuel Cell Electric Vehicle Symposium (2006)

Snyder, L.V.: Facility location under uncertainty: a review. IIE Trans. **38**:547–564 (2004)

The Automobile Association: UK and Overseas Fuel Prices.http://www.theaa.com/motoring_advice/fuel/ (2011). Accessed 1 Aug 2011.

The Exploratorium: Cabspotting. http://cabspotting.org/ (2008). Accessed 12 Feb 2011.

Tom, G., Kurt, P.: California's Residential Electricity Consumption, Prices, and Bills 1980–2005, California Energy Commission Staff. http://www.energy.ca.gov/2007publications/CEC-200-2007-018/CEC-200-2007-018.PDF (2007). Accessed 5 July 2011

Tzeng, G.H., Lin, C.W., Opricovic, S.: Multi-criteria analysis of alternative-fuel buses for public transportation. Energy Policy **33**(11):1373–1383 (2005)

US Department of Labor: Average energy prices in the San Francisco Area. http://www.bls.gov/ro9/cpisanf_energy.pdf (2010). Accessed 1 July 2011

Upchurch, C., Kuby, M., Lim, S.: A model for location of capacitated alternative-fuel stations. Geogr Anal **41**:1–22 (2009)

US Bureau of Labor Statistics: Average Energy Prices In The San Francisco Area: May 2011. http://www.bls.gov/ro9/cpisanf_energy.htm (2011a). Accessed 14 July 2011

US Bureau of Labor Statistics: San Francisco Electricity Prices. http://www.bls.gov/ro9/cpisanf_energy.htm (2011b). Accessed 8 June 2011

Wirasingha, S., Schofield, N., Emadi, A.: Feasibility analysis of converting a chicago transit authority (cta) transit bus to a plug-in hybrid electric vehicle. In: Vehicle Power and Propulsion Conference, 2008. VPPC '08. IEEE, pp. 1 –7 (2008)

Yarow, J.: The Cost of a Better Place Battery Swapping Station: $500,000. http://www.businessinsider.com/the-cost-of-a-better-place-battery-swapping-station-500000-2009-4 (2009). Accessed 11 Jan 2011

Yellow Cab San Francisco: Yellow Cab San Francisco Rates. http://www.yellowcabsf.com/our-service/cab-fares/ (2011). Accessed 4 July 2011

Zfacts: Current Gas Prices and Price History. Figures from Department of Energy. http://zfacts.com/p/35.html (2010). Accessed 12 July 2011

Author Biographies

Tommy Carpenter is a Ph.D. student in Computer Science at the University of Waterloo. His research focuses on reducing barriers to electric vehicle adoption for both fleets and private consumers.

Andrew Curtis is a software engineer in Palo Alto, California. He finished his Ph.D. at the University of Waterloo in 2012. His dissertation research found ways to reduce the operational costs of data center networks using software-defined networking, graph theory, and optimization algorithms.

S. Keshav is a Professor, ACM Fellow, Canada Research Chair, and Cisco Systems Research Chair in Smart Grid in the Cheriton School of Computer Science at the University of Waterloo.