Planning under Uncertainty for Aggregated Electric Vehicle Charging using Markov Decision Processes

Erwin Walraven
Delft University of Technology
Mekelweg 4, 2628 CD
Delft, The Netherlands

Matthijs T. J. Spaan
Delft University of Technology
Mekelweg 4, 2628 CD
Delft, The Netherlands

Abstract

The increasing penetration of renewable energy sources and electric vehicles raises important challenges related to the operation of electricity grids. For instance, the amount of power generated by wind turbines is time-varying and dependent on the weather, which makes it hard to match flexible electric vehicle demand and uncertain wind power supply. In this paper we propose a vehicle aggregation framework which uses Markov Decision Processes to control charging of multiple electric vehicles and deals with uncertainty in renewable supply. We present a grouping technique to address the scalability aspects of our framework. In experiments we show that the aggregation framework maximizes the profit of the aggregator while reducing usage of conventionally-generated power and cost of customers.

Introduction

Smart grids offer several opportunities and challenges for the field of Artificial Intelligence, such as planning and scheduling of electric vehicle charging (Rigas, Ramchurn, and Bassiliades 2015). For example, the increased penetration of renewable energy sources and electric vehicles (EVs) in distribution networks gives rise to the development of intelligent planning methods for so-called *aggregators*. In smart grids these aggregators represent flexible charging demand of a large number of EVs, which can be shifted to periods with sufficient renewable supply such that peak loads are reduced and renewable supply is fully exploited.

In this paper we consider uncertain wind power production combined with the need to coordinate charging of a large number of EVs to take advantage of zero-cost renewable energy. To make sure that multiple vehicles charge their batteries when renewable supply is available, we present a framework based on the Multiagent Markov Decision Process (MMDP) formalism (Boutilier 1996). The development of such a framework poses challenges related to the number of agents involved and the uncertainty associated with renewable energy sources. We focus on the first challenge in this paper, and for the second challenge we build upon recent work related to planning under uncertainty in domains with renewable energy (Walraven and Spaan 2015).

The main contributions of this paper can be summarized as follows. First, we present an electric vehicle aggregation framework which coordinates charging of a collection of EVs using MMDPs. Second, we show how the computation of state value functions can be combined with tree-based representations of uncertainty in wind power. Third, we develop an abstraction of the original MMDP which groups vehicles based on deadlines to keep the number of joint states and actions manageable when increasing the number of vehicles. We also show how the enumeration of MMDP states and actions can be bounded to reduce the number of enumerated states and actions during the computation of value functions.

Our experimental evaluation shows that the aggregation framework maximizes the profit of an aggregator while reducing the cost of individual customers, and the framework reduces usage of conventionally-generated power. Moreover, the experiments show that our MMDP formulation based on groups of vehicles improves the scalability of our framework.

Background

In this section we provide background information related to aggregation in smart grids, wind forecasting and Markov Decision Processes.

Aggregators in Smart Grids

Aggregators in electricity grids are new entities that are acting between individual customers and the utility operator (Gkatzikis, Koutsopoulos, and Salonidis 2013). From the perspective of the utility operator, an aggregator represents a large number of vehicles that require power to charge their batteries. EVs provide a certain amount of flexibility since typically they do not need to be charged immediately. The aggregator is responsible for the communication technology between it and the charging points, allowing for direct control and coordination of connected vehicles.

Individual customers can be incentivized to participate in aggregated charging of vehicles by providing a financial compensation. For instance, customers can sell their flexibility to the aggregator, and get a lower charging tariff in return. From an aggregator point-of-view it is important that the cost associated with the technologies and financial compensations paid to customers are less than the profits that can be made by efficiently controlling vehicle charging.

Wind Speed Forecasting using Scenarios

Wind forecasting methods can be categorized as either physical or statistical, where the latter are suitable for short-term prediction (Giebel et al. 2011). We use a short-term forecasting method that seeks to find analogs (Van den Dool 1989) between observed wind speed and historical wind data (Walraven and Spaan 2015).

The average wind speed during hour t is denoted by w_t, and becomes known at the start of hour $t + 1$.[1] At the start of hour t, wind speed forecasts $\hat{w}_t, \hat{w}_{t+1}, \ldots$ can be computed as follows. Given a sequence of past observations $w_{t-b}, \ldots, w_{t-2}, w_{t-1}$ of length b, we identify similar sequences in a historical dataset containing wind speed measurements based on the Euclidean distance (Walraven and Spaan 2015). For each identified sequence $\hat{w}_{t-b}, \ldots, \hat{w}_{t-2}, \hat{w}_{t-1}$, the subsequent historical wind speed measurements $\hat{w}_t, \hat{w}_{t+1}, \ldots, \hat{w}_{t+y}$ provide a scenario of length y which encodes future wind speed.

Probabilistic wind speed forecasts can be encoded using scenario trees (Conejo, Carrion, and Morales 2010), which is a commonly used technique in the energy domain. Scenario trees can also be combined with wind forecasting methods such as ARMA models (Torres et al. 2005), and therefore the planning methods that we present in this paper are not limited to analog-based wind forecasting. Furthermore, the size of the tree can be managed using scenario reduction techniques (Dupačová, Gröwe-Kuska, and Römisch 2003).

Markov Decision Processes

In this paper we use techniques based on the Markov Decision Process (MDP) formalism (Puterman 1994) and its extension to multiple agents (Boutilier 1996). An MDP is a tuple (S, A, P, R, T), where S is a finite set of states and A is a finite set of actions. The function $P : S \times A \times S \to \mathbb{R}$ defines the state transition probabilities, where $P(s, a, s')$ is the probability to transition from state s to state s' after executing action a. The function $R : S \times A \times S \to \mathbb{R}$ defines the reward function, where $R(s, a, s')$ is the immediate reward received when transitioning from state s to s' after executing action a. The feasible set of actions that can be executed in state s is denoted by $A(s)$, and the MDP has a finite time horizon T. An optimal solution to the MDP can be defined using an optimal value function $V_t^* : S \to \mathbb{R}$ for each timestep $t = 0, \ldots, T - 1$ satisfying the Bellman optimality equation:

$$V_t^*(s) = \max_{a \in A(s)} \sum_{s' \in S} P(s, a, s')(R(s, a, s') + V_{t+1}^*(s')).$$

The optimal solution is a policy $\pi_t^* : S \to A$ which can be used by the decision maker to select an optimal action in each timestep $t = 0, \ldots, T - 1$, and can be defined as follows:

$$\pi_t^*(s) = \arg\max_{a \in A(s)} \sum_{s' \in S} P(s, a, s')(R(s, a, s') + V_{t+1}^*(s')).$$

---

[1]Note that throughout the paper we assume hourly intervals, but our method can be trivially generalized to other intervals.
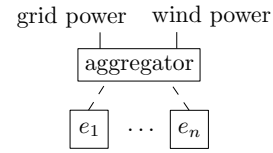
grid power   wind power

Figure 1: Vehicle aggregation with conventionally-generated grid power, wind power and n electric vehicles

The MMDP formalism (Boutilier 1996) generalizes MDPs to the multiagent case, in which the state space is defined by taking the Cartesian product of the state spaces of the individual agents, and actions represent the joint actions that can be executed by the agents. An MMDP can still be considered as a regular MDP, and can be solved using the same algorithms (e.g., value iteration).

In our framework the individual agents are transition-independent (i.e., P can be computed as the product of individual transition functions defined over the individual states and actions of each agent), as the decision whether or not to charge a particular vehicle only affects that vehicle's state of charge. However, since they are coupled through the joint reward function (only a certain number of vehicles can be charged for free using renewable energy), the value function is not factored. While specific solution algorithms have been developed for transition-independent Decentralized MDPs (Becker et al. 2003; Dibangoye et al. 2013), these do not apply to the centralized MMDP model.

Aggregated EV Charging

In this section we propose an aggregation framework for electric vehicle charging, and we formalize the optimization problem that needs to be solved by the aggregator.

We propose a vehicle aggregation framework as shown in Fig. 1. The aggregator is responsible for charging n electric vehicles and is able to use wind power generated by small-scale wind turbines in the residential area, such as wind turbines mounted on tall apartment buildings. Wind power has zero marginal cost, and we assume that excess of wind power can be sold to the utility operator. If the amount of wind power is not sufficient to charge the vehicles in time, additional conventionally-generated power can be bought from the utility operator.

Now we formally introduce the optimization problem that needs to be solved by the aggregator. We consider an ordered set $E = (e_1, \ldots, e_n)$ containing n electric vehicles. A vehicle e_i is connected to its charging point at the start of hour c_i, and needs to charge h_i hours before hour d_i starts. Thus, we can define each vehicle e_i as a tuple $e_i = (c_i, d_i, h_i)$. We assume that the charging rate of each charging point is equal to z kW and that each charging point can only accommodate a single vehicle.

The aggregator is able to buy power from the utility company and pays p_t^b per kWh during hour t. If the wind turbine produces more power than needed, excess wind power can be sold to the utility company for p_t^s per kWh during hour t. The aggregator receives a fixed payment m_i from each vehicle $e_i \in E$ once charging has finished, which is dependent

on the amount of energy used to charge the vehicle.

The power generated by the wind turbine during hour t is $g(w_t)$ kW, where w_t is the wind speed during hour t. The mapping from wind speed to wind power can be modeled using a sigmoid power curve (Ströhle et al. 2014), as shown below:

$$g(w_t) = C \cdot (1 + e^{6 - \frac{2}{3} w_t})^{-1},$$

where C is the rated capacity of the wind turbine.

In order to define the objective function of the aggregator, we introduce decision variables corresponding to the charging decisions of the vehicles. Note that as the aggregator is contractually obligated to charge all vehicles by their deadline (if feasible given deadline and required charge), its payments m_i are not present in the objective function. Variable $x_{i,t}$ equals 1 if vehicle e_i charges during hour t, and is 0 otherwise. The total number of charging vehicles during hour t can be defined as $x_t = \sum_{i=1}^{n} x_{i,t}$. The optimization problem of the aggregator can be formulated as follows:

$$\max \quad \sum_{t=0}^{T-1} f(x_t, w_t)$$

$$\text{s.t.} \quad \sum_{t=c_i}^{d_i - 1} x_{i,t} = h_i \qquad i = 1, \dots, n$$

where the function f computes the benefit to be had by the aggregator when charging x_t vehicles if the wind speed is w_t during hour t. The function can be defined as follows:

$$f(x_t, w_t) = \begin{cases} p_t^s \cdot (g(w_t) - x_t \cdot z) & g(w_t) > x_t \cdot z \\ p_t^b \cdot (g(w_t) - x_t \cdot z) & \text{otherwise} \end{cases}. \tag{1}$$

Note that this function returns negative values if the amount of wind power $g(w_t)$ is not sufficient to charge x_t vehicles, because in such cases additional power needs to be bought from the utility operator. The total profit of the aggregator can be defined as $\sum_{i=1}^{n} m_i + \sum_{t=0}^{T-1} f(x_t, w_t)$.

If the wind speed over time and the parameters of the vehicles are known, then the optimization problem can be solved using mixed-integer programming. However, the aggregator does not know precisely how much wind power will be generated in the future, and needs to make decisions under uncertainty. In the next two sections we will discuss how wind uncertainty can be modeled, and we present a planning algorithm to make charging decisions under uncertainty.

Planning for Aggregated EV Charging

In this section we show how the planning problem for aggregated EV charging can be formulated as a Multiagent Markov Decision Process (MMDP). After introducing the representation of individual vehicles, we show how MMDP value functions can be combined with wind uncertainty represented by a scenario tree. Next, we introduce an MMDP abstraction based on vehicle groups to keep the number of joint states and actions manageable when increasing the number of electric vehicles.
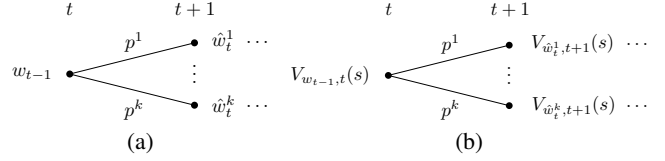

Figure 2: (a) Scenario tree representing w_{t-1}, and k branches corresponding to forecasts of w_t and their probabilities. (b) Value tree containing a value function for hour t, and k value functions for hour $t+1$.

States and Actions of Individual Vehicles

First we describe how the aggregated EV charging problem can be formulated as MMDP, in which each agent represents an electric vehicle. At the start of hour t, we define the state h_i^t of a vehicle as the remaining number of hours during which it needs to charge (assuming a vehicle should be fully charged by the deadline). Since charging must finish before the deadline has passed, it should hold that $h_i^{d_i} = 0$.

Each agent has two actions which it can execute: *charge* and *idle*. Charging is feasible if $h_i^t > 0$ and reduces its demand by one hour: $h_i^{t+1} = h_i^t - 1$. Being idle is feasible if $h_i^t < (d_i - t)$ and does not affect its state of charge (i.e., $h_i^{t+1} = h_i^t$). Based on the feasibility of the actions, we use a state-dependent action space to ensure that deadlines are not being violated.

The joint states and actions of the MMDP can be created by taking the Cartesian product of the states and actions of individual agents. The joint reward function of the agents can be computed using the function $f(x_t, w_t)$ (1), where x_t is the number of charging vehicles and w_t is the wind speed during hour t.

Computing Value Functions in Scenario Trees

Instead of using a set of scenarios as a representation of the wind forecast (Walraven and Spaan 2015), we use a scenario tree representation which encodes the scenarios as a tree, as shown in Fig. 2a. The tree is constructed at the start of hour t, when w_{t-1} becomes known, and forecasted wind speed values are represented by branches in the tree with a corresponding probability.

Rather than encoding wind uncertainty in the state transitions of the MMDP formulation, we introduce separate value functions associated with the nodes of the scenario tree, which allows us to naturally compute value functions for each path of the scenario tree (Leterme et al. 2014). A tree-based representation is beneficial since it does not require separate state variables to encode time-dependent wind forecasts in the state description of the MMDP.

Fig. 2b shows a value function $V_{w_{t-1},t}(s)$ that can be used to select an action at the start of hour t, and the corresponding tree has the same structure as the scenario tree in Fig. 2a. There are k possible realizations for the wind speed during hour t, represented by $\hat{w}_t^1, \dots, \hat{w}_t^k$, and there is a probability p^j and value function $V_{\hat{w}_t^j, t+1}(s)$ corresponding to each realization. The value function $V_{w_{t-1},t}(s)$ can be computed
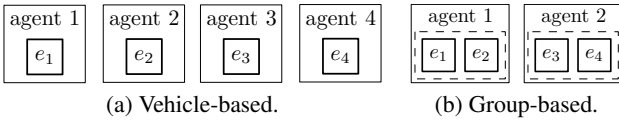
(a) Vehicle-based.   (b) Group-based.

Figure 3: Vehicle-based and group-based MMDPs

as shown below:

$$V_{w_{t-1},t}(s) = \max_{a \in A(s)} \sum_{j=1}^{k} \sum_{s' \in S} (p^j \cdot P(s,a,s') \cdot$$

$$(R(s,a,s',\hat{w}_t^j) + V_{\hat{w}_t^j, t+1}(s')))$$

where the reward function $R(s,a,s',\hat{w}_t^i)$ is an augmented reward function such that the reward can be dependent on the wind speed $\hat{w}_t^i$ during hour t. The value functions for the entire scenario tree can be computed using dynamic programming, in which the value function of each node is computed using the value functions of its child nodes, similar to the example above. In Fig. 2b we show the tree for just one step ahead. However, the value functions $V_{\hat{w}_t^j, t+1}$ also need to be computed recursively based on the value functions in multiple subsequent branches. Eventually, an optimal action can be chosen using the value function associated with the root of the tree. Finally, such a tree structure offers many possibilities besides dynamic programming, for instance heuristic search or branch-and-bound methods.

Group-Based MMDP Abstraction

In order to reduce the number of joint states and actions when increasing the number of electric vehicles, we present a group-based MMDP abstraction in which each agent represents a group of vehicles. The difference between vehicle-based and group-based MMDP formulations is illustrated in Fig. 3. The grouping technique is based on deadlines of vehicles, which is formalized below.

Definition 1 (Vehicle group). *A vehicle group $G_d \subseteq E$ is defined as a subset of vehicles whose deadline is equal to d. In other words, for each $e_i \in G_d$ it holds that $d_i = d$.*

The state of group G_d at the start of hour t is defined as $s_d^t = \sum_{e_i \in G_d} h_i^t$, which is simply the aggregated demand of the vehicles belonging to the group. It should hold that $s_d^d = 0$, since the deadline of the vehicles belonging to the group is identical. The action space A_d contains charging actions corresponding to group G_d. Each action $a \in A_d$ corresponds to the number of vehicles that is charging within the group. After executing action a, the demand of the entire group is reduced accordingly: $s_d^{t+1} = s_d^t - a$. Similar to the vehicle-based formulation, the joint reward can be computed using the function $f(x_t, w_t)$.

Even with grouping of vehicles, obstacles to scalability might remain. In particular, it might be the case (and even likely in a typical overnight charging scenario) that many vehicles share the same deadline and hence certain G_d sets will be large, resulting in large A_d sets. We offer two solutions to this problem, both of which result in suboptimal

policies. First, the size of each A_d can be limited by considering charging only multiples of l vehicles, i.e.,

$$A_d = \{0, l, 2l, 3l, \ldots, |G_d|\}. \tag{2}$$

The loss of fine-grained control will typically be compensated by the ability to solve for much larger sets of vehicles. Second, our group-based planner only requires that all vehicles in a group share the same deadline, hence an aggregator could create many G_d sets. If the available renewable energy is split among them equally (for instance), each such set can be planned for separately.

Planning with Group-Based MMDPs

A group-based MMDP can directly be solved by computing value functions in the scenario tree. In this section we present bounds on the feasible states and actions of vehicle-based and group-based MMDPs, which are important to avoid enumeration of unreachable parts of the state space, and they ensure that charging is finished by the deadline.

In a vehicle-based MMDP the state enumeration can be reduced by observing that some parts of the state space are not reachable. For timesteps $t' \geq t$ the enumerated states $h_i^{t'}$ can be bounded as follows:

$$\max(0, h_i^t - (t' - t)) \leq h_i^{t'} \leq \min(h_i^t, d - t').$$

The lower bound is achieved when charging as fast as possible during hours $t, \ldots, t' - 1$, and the upperbound is achieved when being idle as much as possible during this period. The state-dependent action space can be restricted using the conditions we discussed earlier.

Now we consider a group G_d, for which we can assume that s_d^t is known at the start of hour t, as well as h_i^t for each $e_i \in G_d$. This assumption can be made since the aggregator is able to observe the states of the individual vehicles before making a decision for hour t. Based on the bounds on the demand of the individual vehicles, the feasible set of states at time $t' \geq t$ for group G_d is given by:

$$\sum_{e_i \in G_d} \max(0, h_i^t - (t' - t)) \leq s_d^{t'} \leq \sum_{e_i \in G_d} \min(h_i^t, d - t').$$

These bounds have been constructed by taking the sum of the lower and upper bounds corresponding to the individual vehicles belonging to the group. To reduce the number of enumerated actions $a \in A_d$ for a state $s_d^{t'}$ ($t \leq t' < d$) we use the following bounds:

$$a \geq \max\left(0, s_d^{t'} - \lceil s_d^{t'+1} \rceil\right),$$

$$a \leq \min\left(|G_d|, s_d^{t'} - \lfloor s_d^{t'+1} \rfloor\right),$$

where $\lceil s_d^{t'+1} \rceil$ and $\lfloor s_d^{t'+1} \rfloor$ denote the upper and lower bound on $s_d^{t'+1}$, respectively. These bounds ensure that the planner does not violate deadlines of groups.

Experiments

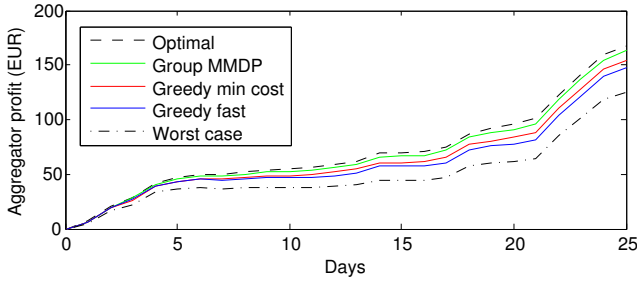This section describes the results of our experiments. We use historical wind data from the Sotavento wind farm in

Figure 4: Cumulative profit made by the aggregator

Figure 5: Running time comparison between vehicle-based and group-based MMDP formulations (log scale)

(a) Running time   (b) Profit

Figure 6: Effect of action space compression (2)

Spain.[2] We simulate the hourly average wind speed for the period from September 2, 2012 until September 26, 2012. The forecasts are based on data from the period September 1, 2009 until December 31, 2009. Unless stated otherwise, the capacity of the wind turbine involved is 50 kW. We assume that the charging rate of the vehicles is equal to 3 kW, which corresponds to a compact hatchback. The electricity price during the simulation is time-dependent, for which we use data from a European power market, which gives us an hourly electricity price (unit EUR/kWh). The feed-in tariff is assumed to be 50 percent of the tariff for buying power.

Aggregator Profit

First we investigate whether the aggregator can make a profit by coordinating vehicles. We simulate 25 days, and during each day we charge 30 vehicles. For each vehicle $e_i \in E$, the payment m_i is 10 percent lower than the minimum cost the customer would pay to the utility operator without participation, which provides a clear incentive for the customers to subscribe to the aggregator. In order to compensate for the discount given to customers, the aggregator needs to efficiently use zero-cost wind power.

Fig. 4 shows the cumulative daily profit of the aggregator for several different planners, which needs to be maximized. In addition to our MMDP planner with groups, we use a greedy planner which charges each vehicle during its individual cheapest hours (i.e., minimal cost), and another greedy planner which charges the vehicles as fast as possible. Lower- and upper bounds on the profit have been computed using a mixed-integer programming formulation, which computes omniscient optimal and worst case charging schedules based on the actual wind speed during the day. In practice it would not be possible to find such schedules, since wind speed in the future is uncertain.

From the experiment we derive two conclusions. First, the aggregator is able to make profit by coordinating vehicles, even if it provides financial compensation to customers. Second, the group-based MMDP planner outperforms two greedy planners in terms of profit, and its profit is close to the profit of the omniscient optimal planner.

Vehicle-Based and Group-Based MMDPs

Next we study the influence of grouping on the scalability of MMDP formulations for electric vehicle charging. To
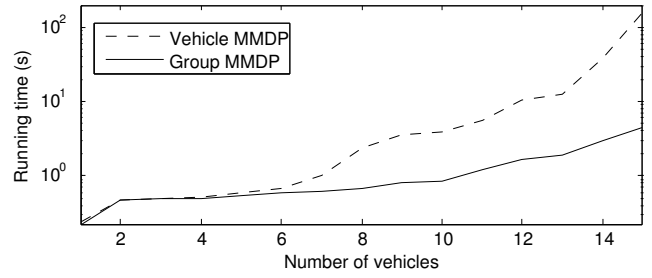
[2]Data is available at www.sotaventogalicia.com.
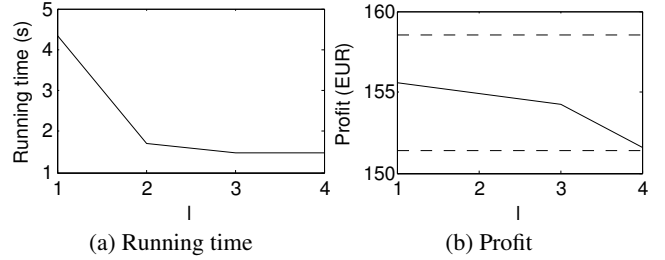
study the difference between vehicle-based and group-based MMDPs, we constructed a set of EVs $E' = (e_1, \ldots, e_{15})$, in which the first three vehicles do not have common deadlines. When we run vehicle-based and group-based planners on the first $1 \le \delta \le 15$ vehicles of E', we expect that grouping only provides improved scalability if $\delta > 3$. In Fig. 5 we show the running times of vehicle-based and group-based MMDPs for an increasing δ (i.e., number of vehicles), which confirms our expectation that group-based formulations require less computation time if groups of vehicles can be created. Note that a log scale is used for the y-axis representing the running time.

Action Space Compression

When after grouping large sets of vehicles remain, it may be desirable to perform action space compression, as defined in Eq. 2. For a case of 15 vehicles, Fig. 6a shows the effect on runtime of increasing l (the level of discretization of the action space) and Fig. 6b the corresponding profit. We can see that as expected a small loss is incurred, but that runtime decreases significantly. The dashed lines represent the profit of the optimal and greedy minimal cost planners, which shows that the MMDP planner still performs better than the greedy planners.

Grid Power Consumption

Although the main objective of the aggregator is optimizing its profit, it may be able to reduce power consumption from the grid, since it is able to charge vehicles during periods in which wind speed is high. Fig. 7 shows the cumulative grid power consumption corresponding to the simulation of the previous experiment. We observe that the grid power con-
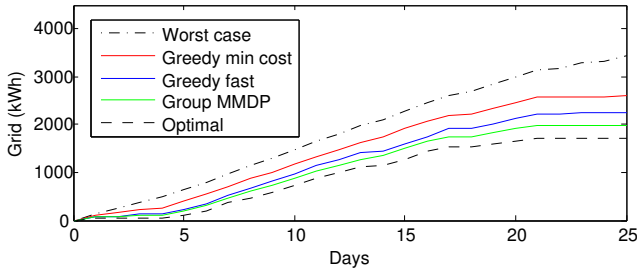
Figure 7: Cumulative power consumption of the vehicles

Figure 8: (a) Aggregator profit as a function of turbine capacity (b) Profit as a function of vehicle discount

sumption of the MMDP planner is lower than the power consumption of the greedy planners involved in the experiment. Therefore, we conclude that an aggregator that aims to maximize its profit also reduces grid power consumption, which can be considered as one of its side effects.

Influence of Wind Turbine Capacity

Until now we assumed a fixed capacity of the wind turbine, but it can be expected that the capacity of the wind turbine influences the profit of the aggregator. In order to test the influence of the turbine capacity on the profit, we run simulations in which we charge 20 vehicles during each day, and we assume that wind power cannot be sold to the utility operator. It should be noted that wind power must be used in practice, but it eliminates the influence of selling wind power in our experiment. Small-scale wind power involves turbines with a capacity of at most 50 kW, and therefore we repeat the simulation for an increasing turbine capacity up to 50kW, as shown in Fig. 8a.

We can derive three conclusions. First, if the turbine capacity is too low then the aggregator is not able to make profit. This is caused by the fact that the charging cost will exceed the customer payments if there is almost no wind power available. Second, a relatively small wind turbine may already be sufficient to make profit. Third, the profit is also positive if the turbine capacity is relatively low (e.g., 10 kW), which shows that it is likely that our framework can be used in the residential area where wind turbines typically have a limited capacity (Ayhan and Sağlam 2012).

Influence of Customer Payments

In the previous experiment we observed that the financial compensation paid to the customers influences the profit of the aggregator, and we expect that profit becomes negative if compensation is too high compared to the usage of zero-cost wind power. In the current experiment we assume that the payments m_i are α percent lower than the minimum cost the customer would pay to the utility operator without participation ($0 < \alpha \leq 100$), and we run simulations for an increasing value of α. The parameter α is called the vehicle discount. In Fig. 8b we show the profit of the aggregator as a function of the vehicle discount, which confirms our expectation that it is impossible to make profit if the discount is too high. In order to provide an incentive to customers of EVs to participate, it is sufficient to have a small nonzero α, and
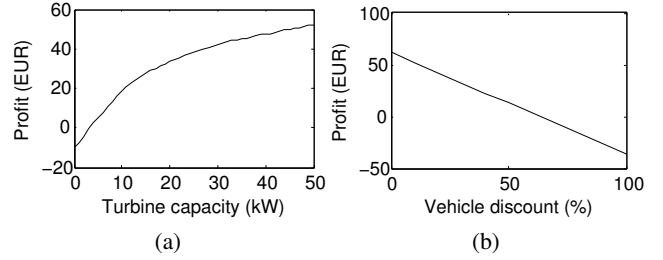
therefore we conclude that the payments m_i of our framework provide an incentive to customers to participate.

Related Work

Markov Decision Processes have been used in recent work to control EV charging. Leterme et al. (2014) discuss an MDP-based approach to control EV charging for wind balancing purposes, in which wind uncertainty is encoded as a tree, but in contrast to our work the paper does not focus on control of individual EVs. An optimization problem similar to our aggregated charging problem is studied by Huang et al. (2015), in which EVs are also clustered based on remaining parking time. Rather than applying an exact solving algorithm, the authors use Monte Carlo simulations to address scalability. Aggregators can learn a consumption pattern of their fleet before buying energy in the day-ahead market using reinforcement learning, as shown by Vandael et al. (2015). Other objective functions for MDPs and EVs have also been studied in existing work, such as minimization of waiting time at shared charging stations with multiple charging points (Zhang et al. 2014).

Conclusions

In this paper we consider the problem of charging electric vehicles in the residential area using renewable energy. We present an aggregated charging technique based on Multiagent Markov Decision Processes which accounts for the uncertainty in renewable supply and coordinates the charging process of several EVs. We use groups of vehicles to create an abstraction of the MMDP, which reduces the number of joint states and actions. Our experiments show that our framework is able to charge a collection of EVs, reduces cost of the individual customers and reduces consumption of conventionally-generated power. In future work, additional grouping of vehicles based on charging rate and spatial location will allow us to take network constraints and physical power flows into account.

Acknowledgements

References

Ayhan, D., and Sağlam, S. 2012. A technical review of building-mounted wind power systems and a sample simulation model. *Renewable and Sustainable Energy Reviews* 16(1):1040–1049.

Becker, R.; Zilberstein, S.; Lesser, V.; and Goldman, C. V. 2003. Transition-Independent Decentralized Markov Decision Processes. In *Proceedings of the 2nd International Conference on Autonomous Agents and Multiagent Systems*, 41–48.

Boutilier, C. 1996. Planning, Learning and Coordination in Multiagent Decision Processes. In *Proceedings of the 6th Conference on Theoretical Aspects of Rationality and Knowledge*, 195–210.

Conejo, A. J.; Carrion, M.; and Morales, J. M. 2010. *Decision Making Under Uncertainty in Electricity Markets*. Springer.

Dibangoye, J. S.; Amato, C.; Doniec, A.; and Charpillet, F. 2013. Producing efficient error-bounded solutions for transition independent decentralized MDPs. In *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, 539–546.

Dupačová, J.; Gröwe-Kuska, N.; and Römisch, W. 2003. Scenario reduction in stochastic programming: An approach using probability metrics. *Mathematical Prog. Series A*.

Giebel, G.; Brownsword, R.; Kariniotakis, G.; Denhard, M.; and Draxl, C. 2011. The State-Of-The-Art in Short-Term Prediction of Wind Power. Technical report, Deliverable of the ANEMOS.plus project.

Gkatzikis, L.; Koutsopoulos, I.; and Salonidis, T. 2013. The Role of Aggregators in Smart Grid Demand Response Markets. *IEEE Journal on Selected Areas in Communications* 31(7):1247–1257.

Huang, Q.; Jia, Q.-S.; Qiu, Z.; Guan, X.; and Deconinck, G. 2015. Matching EV Charging Load With Uncertain Wind : A Simulation-Based Policy Improvement Approach. *IEEE Transactions on Smart Grid* 6(3):1425–1433.

Leterme, W.; Ruelens, F.; Claessens, B.; and Belmans, R. 2014. A Flexible Stochastic Optimization Method for Wind Power Balancing With PHEVs. *IEEE Transactions on Smart Grid* 5(3):1238–1245.

Puterman, M. L. 1994. *Markov Decision Processes: Discrete Stochastic Dynamic Programming*. John Wiley & Sons, Inc.

Rigas, E.; Ramchurn, S.; and Bassiliades, N. 2015. Managing electric vehicles in the smart grid using artificial intelligence: A survey. *IEEE Transactions on Intelligent Transportation Systems* 16(4):1619–1635.

Ströhle, P.; Gerding, E. H.; De Weerdt, M. M.; Stein, S.; and Robu, V. 2014. Online Mechanism Design for Scheduling Non-Preemptive Jobs under Uncertain Supply and Demand. In *Proceedings of the 13th International Conference on Autonomous Agents and Multiagent Systems*, 437–444.

Torres, J. L.; García, A.; De Blas, M.; and De Francisco, A. 2005. Forecast of hourly average wind speed with ARMA models in Navarre (Spain). *Solar Energy* 79(1):65–77.

Van den Dool, H. M. 1989. A New Look at Weather Forecasting through Analogues. *Monthly Weather Review* 117(10):2230–2247.

Vandael, S.; Claessens, B.; Ernst, D.; Holvoet, T.; and Deconinck, G. 2015. Reinforcement Learning of Heuristic EV Fleet Charging in a Day-Ahead Electricity Market. *IEEE Transactions on Smart Grid* 6(4):1795–1805.

Walraven, E., and Spaan, M. T. J. 2015. Planning under Uncertainty with Weighted State Scenarios. In *Proceedings of the 31st Conference on Uncertainty in Artificial Intelligence*, 912–921.

Zhang, T.; Chen, W.; Han, Z.; and Cao, Z. 2014. Charging Scheduling of Electric Vehicles With Local Renewable Energy Under Uncertain Electric Vehicle Arrival and Grid Power Price. *IEEE Transactions on Vehicular Technology* 63(6):2600–2612.