Planning Under Uncertainty for Aggregated Electric Vehicle Charging with Renewable Energy Supply

Walraven, Erwin; Spaan, Matthijs

DOI
10.3233/978-1-61499-672-9-904

Publication date
2016

Document Version
Final published version

Published in
ECAI 2016 - 22nd European Conference on Artificial Intelligence

Citation (APA)
Walraven, E., & Spaan, M. T. J. (2016). Planning Under Uncertainty for Aggregated Electric Vehicle Charging with Renewable Energy Supply. In G. A. Kaminka, M. Fox, P. Bouquet, E. Hüllermeier, V. Dignum, F. Dignum, & F. van Harmelen (Eds.), ECAI 2016 - 22nd European Conference on Artificial Intelligence (pp. 904-912). (Frontiers in Artificial Intelligence and Applications; Vol. 285). IOS Press. https://doi.org/10.3233/978-1-61499-672-9-904

ECAI 2016
G.A. Kaminka et al. (Eds.)

doi:10.3233/978-1-61499-672-9-904

Planning Under Uncertainty for Aggregated Electric Vehicle Charging with Renewable Energy Supply

Erwin Walraven[1] and **Matthijs T. J. Spaan**[1]

Abstract. Renewable energy sources introduce uncertainty regarding generated power in smart grids. For instance, power that is generated by wind turbines is time-varying and dependent on the weather. Electric vehicles will become increasingly important in the development of smart grids with a high penetration of renewables, because their flexibility makes it possible to charge their batteries when renewable supply is available. Charging of electric vehicles can be challenging, however, because of uncertainty in renewable supply and the potentially large number of vehicles involved. In this paper we propose a vehicle aggregation framework which uses Markov Decision Processes to control 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, reduces cost of customers and reduces consumption of conventionally-generated power.

1 INTRODUCTION

The emergence of renewable energy sources in electricity grids is accompanied by several challenges [29]. For instance, power produced by solar panels and wind turbines is dependent on the weather and may cause power production peaks outside the secure range of the grid. Moreover, when many consumers use cheap electricity when renewables have a high output, the grid may become significantly congested. Traditionally such problems were addressed by expensive reinforcements of the grid, but this can be very costly [34]. A recent development is intelligently controlling generation and consumption of local consumers, and thereby creating a smart distribution grid.

Smart distribution grids offer several opportunities and challenges for the field of Artificial Intelligence, such as planning and scheduling of charging of electric vehicles [24]. In order to reduce peak loads and exploit locally produced renewable energy, such as small-scale wind power, shifting flexible electric vehicle charging demand to periods with sufficient renewable supply requires planning algorithms for so-called *aggregators*. These aggregators are entities in smart distribution grids responsible for coordinating a large number of vehicles, and need to be able to deal with uncertain information regarding the availability of renewable supply.

In this paper we consider uncertain wind power production combined with the need to coordinate charging of a large number of electric vehicles (EVs), to take advantage of renewable energy and to reduce consumption of conventionally-generated power. To make sure that vehicles charge their batteries when renewable supply is available, we present an aggregation framework based on the Multiagent Markov Decision Process (MMDP) formalism [5]. The development of such a framework poses challenges related to the number of agents involved and the uncertainty associated with renewable energy sources. The first challenge is the main topic of this paper, and for the second challenge we build upon recent work related to modeling uncertainty of renewables [35].

Our main contributions can be summarized as follows. First, we present an electric vehicle aggregation framework which coordinates charging of collection of EVs using MMDPs. Second, we describe how the computation of value functions can be combined with tree-based representations of uncertainty in renewable wind power, such that the aggregation framework naturally accounts for uncertainty in renewable supply. 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 show how the enumeration of MMDP states and actions can be limited to reduce the number of enumerated states and actions during the computation of value functions.

In experiments based on realistic data we show that our aggregation framework is able to optimize the profit of an aggregator while reducing cost of individual consumers. Moreover, we show that electric vehicles are charging when renewable supply is available, such that consumption of conventionally-generated grid power is reduced. The experiments also show that the group-based abstraction makes our framework sufficiently scalable to control vehicles in a realistically-sized street or a small neighborhood.

The structure of the paper is as follows. In Section 2 we introduce background information about aggregation in smart grids, wind forecasting and Markov Decision Processes. Section 3 formalizes the aggregated electric vehicle charging problem. We present the corresponding MMDP formulation in Section 4, and in Section 5 we discuss an abstraction of the MMDP to improve scalability. Section 6 describes our experimental results, and the remaining sections discuss related work and our conclusions.

2 BACKGROUND

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

2.1 Aggregators in Smart Grids

Aggregators in electricity grids are new entities that are acting between individual customers and the utility company [13]. From the perspective of the utility company, 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.

[1] Delft University of Technology, Delft, The Netherlands. Email: e.m.p.walraven@tudelft.nl, m.t.j.spaan@tudelft.nl.

The flexibility of EVs can be used to address grid congestion problems. For example, during the early morning and the evening the total power demand is high since many people are at home. Current distribution grids have sufficient capacity to deal with the demand of conventional devices during such periods. However, a large number of EVs require a significant amount of power for charging, for which the capacity may not be sufficient [21, 30]. Flexibility of EVs can be used to address this problem, since EV demand can be shifted to periods in which either renewable power supply or sufficient grid capacity is available [2]. Since demand shifting for a large number of EVs requires coordination, aggregators have been proposed to control flexible demand of a large number of EVs. An aggregator is responsible for the communication technology between it and the charging points, allowing for direct control and coordination of vehicles connected to the network.

Individual customers can be incentivized to participate in aggregated charging of vehicles by providing a financial compensation. For instance, customers can sell their flexibility 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 vehicles of customers.

2.2 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 [11]. We use a short-term forecasting method that finds analogs [31] between observed wind speed and historical wind data [35].

The average wind speed during hour t is denoted w_t, and becomes known at the start of hour $t+1$.[2] 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 [35]. 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 describing future wind speed.

Probabilistic wind speed forecasts can be encoded using scenario trees [7]. Scenario trees can also be combined with wind forecasting methods such as ARMA models [28], 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 [10].

2.3 Markov Decision Processes

We use techniques based on the Markov Decision Process (MDP) formalism [23] and its extension to multiple agents [5]. 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 $P : S \times A \times S \rightarrow \mathbb{R}$ 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. Similarly, the function $R : S \times A \times S \rightarrow \mathbb{R}$ 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 $A(s)$, and the MDP has a finite time horizon T. A policy is a function $\pi : S \rightarrow A$ which maps states to actions and this function can be used by a decision maker to select an action for a given state. Optimal policies can be defined in terms of a value function $V^\pi : S \rightarrow \mathbb{R}$. The value of a state s under policy π, denoted by $V^\pi(s)$, is defined as the expected reward when starting from state s and following policy π thereafter. For an optimal policy π^* it holds that $V^{\pi^*}(s) \geq V^\pi(s)$ for each state $s \in S$ and for each policy π. The optimal value function of a finite-horizon MDP can be computed as follows:

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

for $t = 0, \ldots, T-1$. The corresponding time-dependent optimal policy $\pi_t^* : S \rightarrow A$ can be defined as follows:

$$\pi_t^*(s) = \arg\max_{a \in A(s)} \sum_{s' \in S} P(s, a, s')(R(s, a, s') + V_{t+1}^*(s')), \quad (2)$$

for $t = 0, \ldots, T-1$. Note that the value $V_T^*(s)$, corresponding to the final recursive step, can be defined as zero. Alternatively, it can represent a final reward corresponding to state s.

The MMDP formalism [5] generalizes MDPs to the multiagent case, in which a state $s \in S$ characterizes the joint state of the agents and actions $a \in A$ 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).

[2] Note that throughout the paper we assume hourly intervals, but our method is trivially generalized to other intervals.

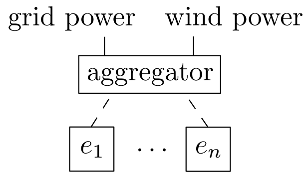


Figure 1: Vehicle aggregation with conventionally-generated grid power, wind power and n electric vehicles.

3 AGGREGATED EV CHARGING

We propose a vehicle aggregation framework as shown in Figure 1. The aggregator is responsible for charging n EVs 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 negligible marginal cost, and excess of wind power can be sold to the utility company. 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 company.

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 the start of hour d_i. 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 EV $e_i \in E$ 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 as follows [26]:

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

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 charge required), 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:

$$\begin{aligned} \max \quad & \sum_{t=0}^{T-1} f(x_t, w_t) \\ \text{s.t.} \quad & \sum_{t=c_i}^{d_i-1} x_{i,t} = h_i \qquad i = 1, \dots, n, \end{aligned}$$

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:

$$f(x_t, w_t) = \begin{cases} p_t^s \cdot (g(w_t) - x_t \cdot z) & g(w_t) > x_t \cdot z \\ p_t^b \cdot (g(w_t) - x_t \cdot z) & \text{otherwise} \end{cases}. \tag{4}$$

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 company. The total profit of the aggregator can be defined as:

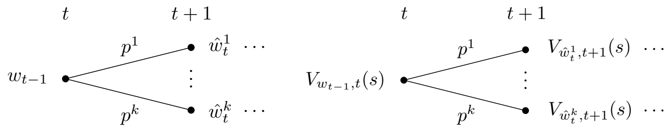
$$\sum_{i=1}^{n} m_i + \sum_{t=0}^{T-1} f(x_t, w_t). \tag{5}$$

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 this paper we address this problem using the MDP formalism because of two reasons. First, it allows us to conveniently separate the reasoning about exogenous wind uncertainty and the reasoning about electric vehicles, as we will show in the next section. Second, MDPs are particularly powerful in situations where the decision maker is able to control the degree of uncertainty that will be encountered in the future. For example, charging overnight before driving to work influences the uncertain demand of the vehicle at the end of the day, since the battery level upon arrival depends on the initial battery level and the distance. This paper only focuses on supply uncertainty and the problem representation, but we selected the MDP formalism based on its potential for extension to uncertainty in charging demand.

4 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). First we discuss how MDP value functions can be computed in scenario trees which encode wind forecasts. Thereafter we introduce an MMDP model in which each agent represents an electric vehicle that needs to be charged.



(a) (b)

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 $t+1$.

4.1 Computing Value Functions in Scenario Trees

We use a scenario tree representation which encodes the scenarios as a tree, as illustrated in Figure 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 j in the tree with a corresponding probability p^j. We introduce separate value functions associated with the nodes of the tree, which allows us to separate the exogenous wind uncertainty and the state transitions of the MMDP model [19]. The tree representation allows us to encode time-dependent wind forecasts, and by doing so we can avoid separate time-dependent MMDP state variables to encode wind uncertainty as part of the state transitions.

Figure 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 Figure 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 as shown below:

$$\begin{aligned} V_{w_{t-1},t}(s) = \max_{a \in A(s)} \sum_{j=1}^{k} \sum_{s' \in S} (p^j \cdot P(s,a,s') \cdot \\ \left(R(s,a,s',t,\hat{w}_t^j) + V_{\hat{w}_t^j,t+1}(s') \right) \right), \end{aligned} \tag{6}$$

where the function $R(s,a,s',t,\hat{w}_t^i)$ is an augmented reward function that is also dependent on the wind speed $\hat{w}_t^i$ during hour t. The state transitions of the MMDP model do not depend on the wind speeds, whereas the augmented reward function allows us to define a reward function that is dependent on both the state and wind speed.

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. In Figure 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. The wind forecast encoded by the scenario tree consists of a finite number of future timesteps, and therefore we have a finite planning horizon. Eventually, an optimal action can be chosen using the value function associated with the root of the tree. The tree representation of the value function corresponds to the recursive formulation in Equation 1, which we formalize below.

Proposition 1. *The value function in Equation 6 defines an optimal value function for an MDP with wind-dependent rewards, whose state transitions are independent of the wind transitions encoded by the scenario tree.*

Proof. We show that Equation 6 can be derived from Equation 1. For the purpose of the proof we make a distinction between an MMDP state s and a global state $\langle s, t, w_{t-1}\rangle$. The MMDP state encodes the EV charging state. The global state encodes both the MMDP state as well as the wind speed during the previous time period and a time step index. Since an MMDP is an MDP, it suffices to use the equation of an optimal MDP value function in the derivation. The value function at the start of hour t maps global states to values and can be defined as follows using Equation 1:

$$V(\langle s, t, w_{t-1}\rangle) = \max_{a \in A(s)} \sum_{\langle s', t+1, \hat{w}_t^j\rangle \in Q_t} P(\langle s, t, w_{t-1}\rangle, a, \langle s', t+1, \hat{w}_t^j\rangle) \cdot \Big(R(\langle s, t, w_{t-1}\rangle, a, \langle s', t+1, \hat{w}_t^j\rangle) + V(\langle s', t+1, \hat{w}_t^j\rangle)\Big), \quad (7)$$

where $Q_t = \{\langle s', t+1, \hat{w}_t^j\rangle \mid s' \in S, \hat{w}_t^j \in \{\hat{w}_t^1, \ldots, \hat{w}_t^k\}\}$ contains all possible global states at the start of hour $t+1$. The MMDP state transitions are independent of the wind transitions and the wind speed transitions are independent of the actions. Hence, it holds that $P(\langle s, t, w_{t-1}\rangle, a, \langle s', t+1, \hat{w}_t^j\rangle) = p^j \cdot P(s, a, s')$. The reward function of the MMDP model depends on the wind speed. Therefore, we define $R(\langle s, t, w_{t-1}\rangle, a, \langle s', t+1, \hat{w}_t^j\rangle) = R(s, a, s', t, \hat{w}_t^j)$ to simplify notation. The variable w_{t-1} can be left out because the reward received after hour t does not depend on the wind speed during hour $t-1$. Now the aforementioned value function can be simplified as follows:

$$V(\langle s, t, w_{t-1}\rangle) = \max_{a \in A(s)} \sum_{j=1}^{k} \sum_{s' \in S} p^j \cdot P(s, a, s') \cdot \Big(R(s, a, s', t, \hat{w}_t^j) + V(\langle s', t+1, \hat{w}_t^j\rangle)\Big). \quad (8)$$

The sum operators still define a sum over all elements in Q_t. Since the transitions of the time step counter t are assumed deterministic, the summation over all possibilities for $t+1$ can be left out. The resulting value function can be transformed to Equation 6 by defining $V(\langle s, t, w_{t-1}\rangle) = V_{w_{t-1},t}(s)$ and $V(\langle s', t+1, \hat{w}_t^j\rangle) = V_{\hat{w}_t^j, t+1}(s')$, which is a simplification of the notation. This step completes the derivation of Equation 6 from Equation 1. An identical derivation can be used to recursively transform the value function equations in the other nodes of the value function tree. Since we consider finite-horizon forecasts and thus a value function tree with a finite number of leafs, this concludes the proof. □

4.2 Vehicle-Based MMDP formulation

Now 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, it should hold that $h_i^{d_i} = 0$.

Each agent has two actions which it can execute: *charge* and *idle*. The *charge* action reduces the demand by one hour: $h_i^{t+1} = h_i^t - 1$, and the *idle* action does not affect its state of charge (i.e., $h_i^{t+1} = h_i^t$). We use a state-dependent action space to ensure that vehicles are guaranteed to meet their deadline. In state h_i^t the *idle* action can only be executed if $h_i^t < d_i - t$, which ensures that there is always enough time left to complete charging before the deadline. The action *charge* can be executed if $h_i^t > 0$, and must be executed if $h_i^t = d_i - t$. By using the state-dependent action space that we just described, it is guaranteed that $h_i^{d_i} = 0$. This is formalized in the following proposition.

Proposition 2. *The state-dependent action space ensures that a vehicle e_i always completes charging before its deadline d_i.*

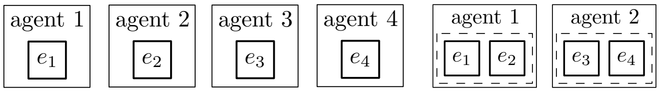
Proof. In order to show that a vehicle always finishes charging before its deadline, we need to show that the action *idle* is never executed in situations where it would lead to a violation of the deadline. For this purpose we assume the contrary, namely that the *idle* action is executed in state h_i^t, leading to a state h_i^{t+1} in which the demand is one higher than the time left for charging: $h_i^{t+1} = (d_i - (t+1)) + 1$. Since the *idle* action was executed, it holds that $h_i^t = h_i^{t+1}$. Now we derive $h_i^t = h_i^{t+1} = (d_i - (t+1)) + 1 = d_i - t$. In state h_i^t, however, action *charge* must have been executed according to our state-dependent action space. This contradicts the assumption that *idle* was executed in state h_i^t. We can conclude that the action *idle* is never executed if it leads to a situation in which it violates a deadline, and we can conclude that our state-dependent action space ensures that vehicles meet their deadline (i.e., $h_i^{d_i} = 0$). □

Until now we defined the states and state-dependent action space for an individual vehicle. For multiple vehicles the joint states and actions of the MMDP can be created by taking the Cartesian product of the states and actions of individual vehicles. For example, if there are two vehicles with states h_1^t and h_2^t at the start of hour t, then their joint state is (h_1^t, h_2^t) and an example of a joint action is (*charge*, *idle*). The joint reward function of the agents can be computed using the function $f(x_t, w_t)$ defined in Equation 4, where x_t is the number of charging vehicles and w_t is the wind speed during hour t. For instance, if a joint action dictates that x_t vehicles need to charge during step t when the wind speed is w_t, then the MMDP reward is equal to $f(x_t, w_t)$. The state transitions of the electric vehicles are assumed deterministic and therefore we do not define a probabilistic transition function. The probabilistic transitions of wind speed are encoded separately using the scenario tree, as discussed in the previous section.

In our MMDP formulation the individual vehicles 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 vehicle), 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. Specific solution algorithms have been designed for transition-independent Decentralized MDPs [4, 9], in which vehicles would take decisions in a decentralized manner. However, these solution techniques do not apply to our MMDP model in which an aggregator controls vehicles in a centralized manner. Other solution algorithms for transition-independent MMDPs exploit sparse reward structures [25], in which only a small subset of the joint actions has a non-zero reward. The latter is not the case in our model.

4.3 Reducing Enumerated States

In this section we present an optimization which reduces the number of states that need to be enumerated in each node of the value function tree when recursively computing the value functions. The number of enumerated states can be reduced by observing that some parts of the state space cannot be reached. For instance, states representing a situation in which a deadline is going to be violated will never be encountered, as stated in Proposition 2, and therefore such states do not need to be considered. When recursively computing a value function $V_{w_{t'-1},t'}(s)$ corresponding to time $t' \geq t$, it is necessary to



(a) Vehicle-based. (b) Group-based.

Figure 3: Vehicle-based and group-based MMDPs.

determine which states s need to be enumerated. For instance, suppose that state $s = (h_1^{t'}, h_2^{t'})$ encodes the joint state of two vehicles at time t', then all possible combinations of $h_1^{t'}$ and $h_2^{t'}$ can be enumerated in order to enumerate all possible states s. The states $h_i^{t'}$ which need to be enumerated for vehicle $e_i \in E$ can be defined as follows:

$$\max(0, h_i^t - (t' - t)) \leq h_i^{t'} \leq \min(h_i^t, d - t'). \quad (9)$$

The lowerbound 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 actions $a \in A(s)$ that need to be enumerated during the computation of $V_{w_{t'-1},t'}(s)$ can be defined using the state-dependent action space.

5 GROUP-BASED MMDPs

In order to reduce the number of joint states and actions when increasing the number of electric vehicles, we present a group-based MMDP formulation in which each agent represents a group of vehicles. The difference between vehicle-based and group-based MMDP formulations is illustrated in Figure 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. 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. 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, for multiple groups the joint states and joint actions can be defined by taking the Cartesian product of the states and actions of the groups. For example, if there are two groups with states s_1^t and s_2^t, then the joint state of the groups is (s_1^t, s_2^t). If there is one vehicle that is charging within both groups, then $(1, 1)$ would be a joint action. In the next section we will elaborate on the state-dependent action space which ensures that the planner does not violate the deadline of a group, similar to the state-dependent action space of the vehicle-based formulation. The joint reward can be computed using the function $f(x_t, w_t)$, similar to the vehicle-based formulation, where x_t is the number of charging vehicles and w_t is the wind speed during hour 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. A potential solution to this problem is restricting the A_d sets, by considering charging only multiples of l vehicles, i.e.,

$$A_d = \{0, l, 2l, 3l, \ldots, |G_d|\}. \quad (10)$$

The loss of fine-grained control will typically be compensated by the ability to solve for larger sets of vehicles. This aspect will also be studied in our experiments.

Example 1 (Vehicle grouping)**.** *In our example formulation we consider six electric vehicles connected to an aggregator at time $t = 0$. The relevant properties of the individual vehicles are shown in Table 1. First we compare the number of states and actions of vehicle-based and group-based MMDP models. When formulating a vehicle-based MMDP, the total number of states is equal to $\prod_{i=1}^{6}(h_i^0 + 1) = 2160$ and the number of actions is equal to $2^6 = 64$. A group-based MMDP formulation can be created by defining a group G_4 with demand* 3*, a group G_5 with demand* 3 *and a group G_6 with demand* 11*. The number of states in such a formulation is equal to* $(3 + 1) \cdot (3 + 1) \cdot (11 + 1) = 192$ *and the number of actions equals* $3 \cdot 2 \cdot 4 = 24$*. Clearly, the total number of states and actions decreased compared to the vehicle-based MMDP formulation. A Dynamic Bayesian network representation of the group-based MMDP is shown in Figure 4. It should be noted that the wind speed transitions in the actual implementation are encoded in a tree-based fashion, as discussed in Section 4.1.*

Table 1: Deadlines and demand of example vehicles.

i	1	2	3	4	5	6
d_i	4	4	5	6	6	6
h_i^0	2	1	3	4	5	2

5.1 Planning with Group-Based MMDPs

A group-based MMDP can directly be solved by computing value functions in the scenario tree. However, due to the aggregation of multiple vehicles into groups it becomes less straightforward which states and actions need to be enumerated in each node of the tree. In this section we first define which states need to be enumerated, and thereafter we discuss the state-dependent action space which ensures that the planner does not violate deadlines of vehicles.

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. When recursively computing the value function $V_{w_{t-1},t}(s)$, it is necessary to know which states $s = s_d^{t'}$ need to be enumerated for timesteps $t' \geq t$. For this purpose we generalize the bounds shown in Equation 9 to bounds on the demand of a group as shown below.

$$\sum_{e_i \in G_d} \max(0, h_i^t - (t' - t)) \leq s_d^{t'} \leq \sum_{e_i \in G_d} \min(h_i^t, d - t') \quad (11)$$

The lower bound has been defined by taking the sum of the lower bounds on the demand $h_i^{t'}$ for each vehicle $e_i \in G_d$. Similarly, the upper bound has been defined by taking the sum of the upper bounds on the demand. The resulting bounds can be used to ensure that we do not enumerate unreachable states in case we use a group-based formulation.

Similar to the vehicle-based MMDP formulation, the executed actions need to ensure that the demand of an entire group is decreased

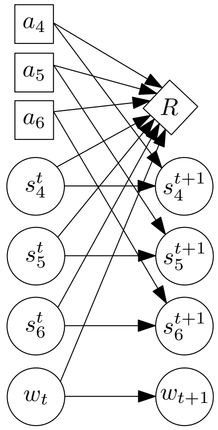


Figure 4: Dynamic Bayesian network corresponding to the group-based MMDP of the example instance.

to zero before its deadline. Therefore, we define a state-dependent action space for a group-based state $s_d^{t'}$. For convenience we let $\lfloor s_d^{t'} \rfloor$ denote the lower bound on $s_d^{t'}$ and $\lceil s_d^{t'} \rceil$ denotes the upper bound on $s_d^{t'}$. In other words, we obtain the following equations:

$$\lfloor s_d^{t'} \rfloor = \sum_{e_i \in G_d} \max(0, h_i^t - (t' - t)), \tag{12}$$

$$\lceil s_d^{t'} \rceil = \sum_{e_i \in G_d} \min(h_i^t, d - t'). \tag{13}$$

Now we can restrict the actions $a \in A(s_d^{t'})$ for a state $s_d^{t'}$ ($t \leq t' < d$) as follows:

$$\max\left(0, s_d^{t'} - \lceil s_d^{t'+1} \rceil\right) \leq a \leq \min\left(|G_d|, s_d^{t'} - \lfloor s_d^{t'+1} \rfloor\right). \tag{14}$$

In the computation of the state-dependent action space $A(s_d^{t'})$ we also use the lower- and upper bound on $s_d^{t'+1}$. These bounds have been properly defined in Equations 12 and 13. The state-dependent action space ensures that in state $s_d^{t'}$ an action is selected in such a way that $\lfloor s_d^{t'+1} \rfloor \leq s_d^{t'+1} \leq \lceil s_d^{t'+1} \rceil$. It holds that $0 = \lfloor s_d^d \rfloor \leq s_d^d \leq \lceil s_d^d \rceil = 0$, which implies that the total group demand is reduced to zero before the deadline.

Proposition 3. *The state-dependent action space for a group G_d ensures that all vehicles $e_i \in G_d$ always complete charging before their deadline d.*

Proof. If the group demand $s_d^{t'+1}$ at time $t'+1$ is higher than $\lceil s_d^{t'+1} \rceil$, then it is impossible to reduce the demand to zero before the deadline. We will show that this never occurs. If $s_d^{t'} > \lceil s_d^{t'+1} \rceil$, then the state-dependent action space defines that at least $s_d^{t'} - \lceil s_d^{t'+1} \rceil$ vehicles will be charged, such that $s_d^{t'+1} \leq s_d^{t'} - (s_d^{t'} - \lceil s_d^{t'+1} \rceil) = \lceil s_d^{t'+1} \rceil$. In other words, any action executed in state $s_d^{t'}$ guarantees that $s_d^{t'+1}$ does not exceed $\lceil s_d^{t'+1} \rceil$. Therefore, we can conclude that the state-dependent action space ensures that the planner does not violate the deadline of a group. □

We have shown that our group-based formulation defines states for groups of vehicles, while still being able to meet the deadlines of all individual vehicles in the EV fleet. It should be noted, however, that the group-based MMDP formulation does not define a Markovian state representation for the original EV charging problem. In other words, the state representation of the group-based formulation does not preserve sufficient information to derive the individual states of all the vehicles within the groups. Due to the aggregation of multiple vehicles into one group the upper bound on the number of vehicles that still needs to charge (i.e., the upper bound on a) may overestimate the number of vehicles that is actually available for charging. In the example below we discuss this potential overestimate in an example. An overestimate might only occur during planning when selecting the actions to compute value functions. When the resulting value function is used to select actions to control the vehicles, then such an overestimate never occurs, because the feasible actions can be determined using the actual state of the individual vehicles.

Example 2 (Overestimate of demand)**.** *Using the previous example instance we illustrate why infeasible actions may be enumerated during the computation of value functions. We consider group G_4 containing two vehicles with demand $h_1^t = 2$ and $h_2^t = 1$ at time $t = 0$. By definition it holds that $s_4^t = 3$. We consider the group-based state $s_4^{t'}$ at time $t' = 1$, for which it holds that $1 \leq s_4^{t'} \leq 3$. In state $s_4^{t'} = 2$, the upper bound on the number of vehicles with non-zero demand is $\min(|G_4|, s_4^{t'} - \sum_{e_i \in G_4} \max(0, h_i^t - ((t'+1) - t))) = \min(2, 2 - \max(0, 2-2) - \max(0, 1-2)) = 2$, which represents that we can charge at most two vehicles simultaneously in this state. However, it may be possible that $h_1^{t'} = 2$ and $h_2^{t'} = 0$, and then only one vehicle can be charged. In this case the number of vehicles with non-zero demand is overestimated by* 1.

6 EXPERIMENTS

This section describes the results of our experiments. We use historical wind data from the Sotavento wind farm in Spain.[3] 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 car. The electricity price during the simulation is time-dependent, for which we use data from a European power market, which gives us an hourly price (unit EUR/kWh). Unless stated otherwise, the feed-in tariff is 50 percent of the tariff for buying power. To define EVs we use realistic vehicle arrival and departure times from a Dutch mobility study, conducted by Statistics Netherlands [6].

6.1 Aggregator Profit and Power Consumption

First we investigate whether the aggregator is able to make profit by coordinating vehicles. We simulate 25 days, and during each day we charge 20 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 company without participation, which provides an 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. It is estimated that there is a 25 percent market share of EVs starting in 2020 [14], hence 20 vehicles can represent a realistically-sized street or a small neighborhood.

Figure 5 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 individually cheapest hours (i.e., min cost), and another greedy planner which charges the vehicles as fast as

[3] Data is available on www.sotaventogalicia.com.

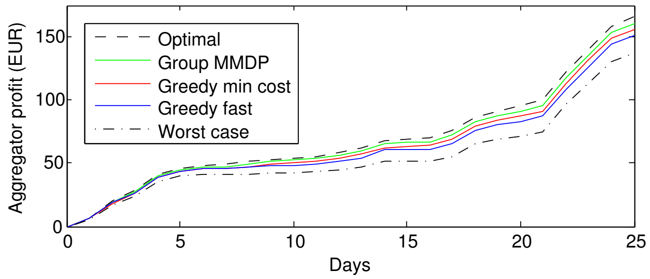


Figure 5: Cumulative profit made by the aggregator.

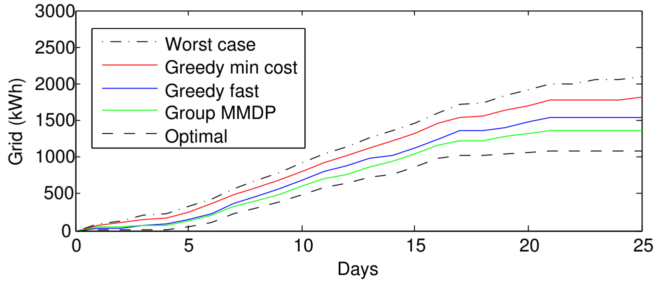


Figure 6: Cumulative power consumption of the vehicles.

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 wind speed during the day. In practice it would not be possible to find such schedules, since wind speed in the future is uncertain.

We conclude that the aggregator is able to make profit by coordinating vehicles, even if it provides a financial compensation to customers of the vehicles. Moreover, 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.

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. Figure 6 shows the cumulative grid power consumption corresponding to the simulation of the previous experiment. We observe that the grid power consumption 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.

6.2 Vehicle-Based and Group-Based MMDPs

Next we study the influence of grouping on the scalability of MMDP formulations for electric vehicle charging. To 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 $1 \leq \delta \leq 15$ vehicles of E', we expect that grouping only provides improved scalability if $\delta > 3$. In Figure 7 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.

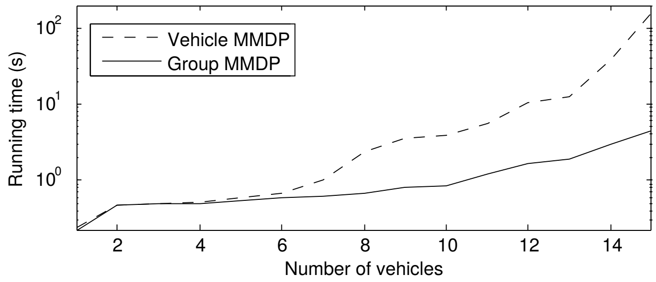


Figure 7: Running time comparison between vehicle-based and group-based MMDP formulations (log scale).

6.3 Action Space Compression

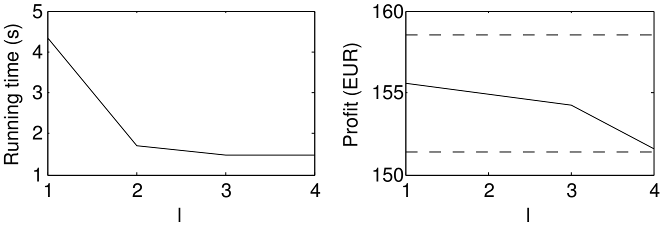
When after grouping large sets of vehicles remain, it may be desirable to perform action space compression to reduce the number of enumerated actions, as defined in Equation 10. This means that the planner only considers charging multiples of l vehicles. For a case of 15 vehicles, Figure 8a shows the effect on runtime of increasing l (i.e., the level of discretization of the action space) and Figure 8b shows the corresponding profit. We can see that as expected a small loss is incurred, but the running time required for the computation of the value functions decreases significantly. The dashed lines represent the profit of the omniscient optimal and greedy min cost planners in the simulation. Our MMDP planner still makes more profit compared to the greedy min cost planner in the simulation.

6.4 Influence of Wind Turbine Capacity

Until now we assumed a fixed turbine capacity, but it can be expected that the turbine capacity influences the profit of the aggregator. In order to study this influence, we run simulations in which we charge 15 vehicles during each day, and we assume that wind power cannot be sold to the utility company. The latter is assumed because this 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 Figure 9a. 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 experiment shows that it is likely that our framework can be used in the residential area where wind turbines typically have a capacity up to a few kilowatts [3].

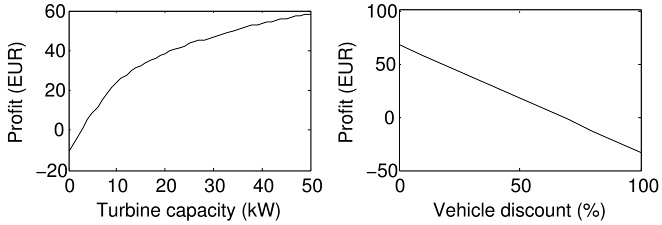
6.5 Influence of Customer Payments

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 the compensations are 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 company 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 Figure 9b 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



(a) Running time (b) Profit

Figure 8: Effect of action space compression (10).



(a) (b)

Figure 9: Profit for increasing turbine capacity (a) and discount (b).

non-zero α, and therefore we conclude that the payments m_i of our framework provide an incentive to customers to participate.

7 RELATED WORK

Leterme et al. discuss an MDP-based approach to control EVs for wind balancing, in which wind uncertainty is encoded as a tree [19], but in contrast to our work their solution does not control individual EVs. Huang et al. [16] cluster EVs based on remaining parking time and use Monte Carlo simulations to estimate a value function. Our scenario-tree encoding of the wind uncertainty provides a more advanced representation of wind uncertainty and cannot directly be combined with their approach. Other objective functions, such as waiting time at charging stations, have also been studied in existing work [37]. Aggregators can use reinforcement learning to learn a consumption pattern of their fleet before buying energy in the day-ahead market [32]. Currently our work only focuses on uncertainty in renewable supply, and it does not model bids in a day-ahead or intraday energy market.

In the power systems community research has focused on matching demand and supply in the unit commitment problem using multi-stage stochastic programming and mixed-integer programming, where exogenous uncertainty in the supply is also characterized using scenarios [22]. Multi-stage stochastic programming methods are typically used for problems with exogenous uncertainty that cannot be controlled by the decision maker [8], whereas Markov Decision Processes are well-suited if control actions influence the uncertainty encountered in the future. For example, stochastic state transitions in our MDP models can also be used to model uncertainty in arrival time and departure time of electric vehicles, which is hard to model in a multi-stage stochastic programming formulation. Research has also focused on inclusion of network characteristics in aggregate models of multiple EVs [17]. Compared to our work, existing work in this area focuses more on modeling the electrical aspects and the impact on the power system. Congestion management schemes have been developed for electric vehicles, which typically assume a deterministic setting in which there is no uncertainty during optimization and execution [33]. Our work can be used for congestion management if renewable supply is uncertain.

Reducing computational requirements by aggregating states of MDPs has been studied in the context of stochastic bisimulation [12], which is an exact method to compute an equivalent smaller-sized MDP, and symmetry reduction [18]. Both methods can theoretically be combined with our work, but require a given MDP which needs to be minimized [20] and often require full state-space enumeration. The latter leads to problems in the multiagent setting because of the exponential growth of the number of states. Our group-based model can be created without needing an initial model, but the abstraction method is not exact. Other abstraction methods include temporal abstractions, such as macro-actions [15] and Semi-MDPs [27], which would allow an aggregator to solve an abstract planning problem to select sub-policies rather than actions. However, these abstractions do not address scalability problems that follow from the large number of EVs, and it is hard to combine such abstraction techniques with exogenous wind uncertainty.

Constrained MDPs [1] include constraints in the dual formulation of a linear program. This framework can also be used to impose constraints to make sure that deadlines are satisfied, but it would be difficult to separate the reasoning about exogenous wind uncertainty in the corresponding linear programming formulations. Moreover, linear programs for Constrained MDPs are typically based on the assumption that the planning horizon is infinite.

8 CONCLUSIONS

In this paper 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 and it reduces the running time required to compute MMDP solutions. 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. Moreover, our work demonstrates that AI methods have the potential to support the development of smart grids. For example, an interesting application of our work can be found in parking garages with local grid capacity constraints, where charging of a large number of EVs needs to be coordinated and peak loads must be prevented.

In future work we aim to include information about uncertain demand in our MMDP formulations, which can be naturally included in stochastic state transitions. Our work can also be extended to asynchronous events and actions using Generalized Semi-MDPs [36], and it can be combined with wind scenario trees generated by ARMA models [28]. Another interesting direction is creating groups of vehicles based on additional characteristics besides their deadline, such as the charging rate and spatial location in the network. Our method can also be combined with power flow computations to derive the power flows through the network. This is useful if capacity violations must be prevented in a congested network.

ACKNOWLEDGMENTS

The work presented in this paper is funded by the Netherlands Organisation for Scientific Research (NWO), as part of the Uncertainty Reduction in Smart Energy Systems program.

REFERENCES

[1] E. Altman, *Constrained Markov Decision Processes*, CRC Press, 1999.

[2] P. Bach Andersen, J. Hu, and K. Heussen, 'Coordination strategies for distribution grid congestion management in a multi-actor, multi-objective setting', in *3rd IEEE PES Innovative Smart Grid Technologies Europe*, (2012).

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

[4] R. Becker, S. Zilberstein, V. Lesser, and C. V. Goldman, 'Transition-Independent Decentralized Markov Decision Processes', in *Proceedings of the 2nd International Conference on Autonomous Agents and Multiagent Systems*, (2003).

[5] C. Boutilier, 'Planning, Learning and Coordination in Multiagent Decision Processes', in *Proceedings of the 6th Conference on Theoretical Aspects of Rationality and Knowledge*, pp. 195–210, (1996).

[6] Centraal Bureau voor de Statistiek (CBS) / Rijkswaterstaat (RWS). Onderzoek Verplaatsingen in Nederland 2013 - OViN 2013, 2014. DANS. http://dx.doi.org/10.17026/dans-x9h-dsdg.

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

[8] B. Defourny, D. Ernst, and L. Wehenkel, 'Multistage Stochastic Programming: A Scenario Tree Based Approach to Planning under Uncertainty', in *Decision Theory Models for Applications in Artificial Intelligence: Concepts and Solutions*, eds., L E Sucar, E F Morales, and J Hoey, 97–143, Information Science Publishing, (2012).

[9] J. S. Dibangoye, C. Amato, A. Doniec, and F. Charpillet, 'Producing efficient error-bounded solutions for transition independent decentralized MDPs', in *Proceedings of the 12th International Conference on Autonomous Agents and Multiagent Systems*, (2013).

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

[11] G. Giebel, R. Brownsword, G. Kariniotakis, M. Denhard, and C. Draxl, 'The State-Of-The-Art in Short-Term Prediction of Wind Power', Technical report, ANEMOS.plus project, (2011).

[12] R. Givan, T. Dean, and M. Greig, 'Equivalence notions and model minimization in Markov decision processes', *Artificial Intelligence*, **147**(1-2), 163–223, (2003).

[13] L. Gkatzikis, I. Koutsopoulos, and T. Salonidis, 'The Role of Aggregators in Smart Grid Demand Response Markets', *IEEE Journal on Selected Areas in Communications*, **31**(7), 1247–1257, (2013).

[14] S. W. Hadley and A. A. Tsvetkova, 'Potential impacts of plug-in hybrid electric vehicles on regional power generation', *The Electricity Journal*, **22**(10), 56–68, (2009).

[15] M. Hauskrecht, N. Meuleau, L. P. Kaelbling, and C. Boutilier, 'Hierarchical Solution of Markov Decision Processes using Macro-actions', in *Proceedings of the 14th Conference on Uncertainty in Artificial Intelligence*, pp. 220–229, (1998).

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

[17] S. Izadkhast, P. Garcia-Gonzalez, P. Frias, L. Ramirez-Elizondo, and P. Bauer, 'An aggregate model of plug-in electric vehicles including distribution network characteristics for primary frequency control', *IEEE Transactions on Power Systems*, **31**(4), 1–12, (2015).

[18] M. Kwiatkowska, G. Norman, and D. Parker, 'Symmetry Reduction for Probabilistic Model Checking', in *Computer Aided Verification*, pp. 234–248, (2006).

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

[20] L. Li, T. J. Walsh, and M. L. Littman, 'Towards a Unified Theory of State Abstraction for MDPs', in *Proceedings of the 9th International Symposium on Artificial Intelligence and Mathematics*, (2006).

[21] J. A. P. Lopes, F. J. Soares, and P. M. R. Almeida, 'Identifying management procedures to deal with connection of electric vehicles in the grid', in *IEEE PowerTech*, (2009).

[22] G. Morales-España, *Unit Commitment: Computational Performance, System Representation and Wind Uncertainty Management*, Ph.D. dissertation, Pontifical Comillas University, KTH Royal Institute of Technology and Delft University of Technology, 2014.

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

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

[25] J. Scharpff, D. M. Roijers, F. A. Oliehoek, M. T. J. Spaan, and M. M. de Weerdt, 'Solving Transition-Independent Multi-agent MDPs with Sparse Interactions', in *Proceedings of the AAAI Conference on Artificial Intelligence*, (2016).

[26] P. Ströhle, E. H. Gerding, M. M. de Weerdt, S. Stein, and V. Robu, '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*, pp. 437–444, (2014).

[27] R. S. Sutton, D. Precup, and S. Singh, 'Between MDPs and semi-MDPs: A framework for temporal abstraction in reinforcement learning', *Artificial Intelligence*, **112**(1), 181–211, (1999).

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

[29] F. Ueckerdt, R. Brecha, and G. Luderer, 'Analyzing Major Challenges of Wind and Solar Variability in Power Systems', *Renewable Energy*, **81**, 1 – 10, (2015).

[30] E. Ungar and K. Fell, 'Plug in, turn on, and load up', *IEEE Power and Energy Magazine*, **8**(3), 30–35, (2010).

[31] H. M. van den Dool, 'A New Look at Weather Forecasting through Analogues', *Monthly Weather Review*, **117**(10), 2230–2247, (1989).

[32] S. Vandael, B. Claessens, D. Ernst, T. Holvoet, and G. Deconinck, 'Reinforcement Learning of Heuristic EV Fleet Charging in a Day-Ahead Electricity Market', *IEEE Transactions on Smart Grid*, **6**(4), 1795–1805, (2015).

[33] R. A. Verzijlbergh, *The Power of Electric Vehicles - Exploring the value of flexible electricity demand in a multi-actor context*, Ph.D. dissertation, Delft University of Technology, 2013.

[34] R. A. Verzijlbergh, L. J. de Vries, and Z. Lukszo, 'Renewable Energy Sources and Responsive Demand. Do We Need Congestion Management in the Distribution Grid?', *IEEE Transactions on Power Systems*, **29**(5), 2119–2128, (2014).

[35] E. Walraven and M. T. J. Spaan, 'Planning under Uncertainty with Weighted State Scenarios', in *Proceedings of the 31st Conference on Uncertainty in Artificial Intelligence*, (2015).

[36] H. L. S. Younes and R. G. Simmons, 'Solving Generalized Semi-Markov Decision Processes using Continuous Phase-Type Distributions', *Proceedings of the 19th National Conference on Artificial Intelligence*, 742–747, (2004).

[37] T. Zhang, W. Chen, Z. Han, and Z. Cao, '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, (2014).