Reinforcement Learning Portfolio Optimization of Electric Vehicle Virtual Power Plants

Master Thesis

Author: Tobias Richter (Student ID: 999999999)
Supervisor: Univ.-Prof. Dr. Wolfgang Ketter
Co-Supervisor: Max Mustermann

Department of Information Systems for Sustainable Society
Faculty of Management, Economics and Social Sciences
University of Cologne

February 3, 2019

Eidesstattliche Versicherung

Hiermit versichere ich an Eides statt, dass ich die vorliegende Arbeit selbstständig und ohne die Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe. Alle Stellen, die wörtlich oder sinngemäß aus veröffentlichten und nicht veröffentlichten Schriften entnommen wurden, sind als solche kenntlich gemacht. Die Arbeit ist in gleicher oder ähnlicher Form oder auszugsweise im Rahmen einer anderen Prüfung noch nicht vorgelegt worden. Ich versichere, dass die eingereichte elektronische Fassung der eingereichten Druckfassung vollständig entspricht.

Die Strafbarkeit einer falschen eidesstattlichen Versicherung ist mir bekannt, namentlich die Strafandrohung gemäß § 156 StGB bis zu drei Jahren Freiheitsstrafe oder Geldstrafe bei vorsätzlicher Begehung der Tat bzw. gemäß § 161 Abs. 1 StGB bis zu einem Jahr Freiheitsstrafe oder Geldstrafe bei fahrlässiger Begehung.

Tobias Richter

Köln, den xx.xx.20xx

Contents

1 Introduction (10%) 1

1.1 Research Motivation 1

1.2 Research Question 1

1.3 Relevance 1

2 Related Literature (10%) 2

2.1 Smart Charging and Balancing the Electric Grid with EV Fleets . 2

2.2 Reinforcement Learning in Smart Grids 5

3 Theoretical Background (10%) 8

3.1 Electricity Markets 8

3.1.1 Balancing Market 8

3.1.2 Spot Market 8

3.2 Reinforcement Learning 8

3.2.1 Notation 8

3.2.2 Introduction 8

3.2.3 Markov Decision Processes 9

3.2.4 Tabular Methods 10

3.2.5 Function Approximation 10

3.2.6 Exploitation-Exploration Trade-off 10

3.2.7 Deep Reinforcement Learning 10

4 Empirical Setting / Data (10%) 10

4.1 Carsharing Fleets of Electric Vehicles 10

4.1.1 Raw Data 10

4.1.2 Preprocessing Steps 13

4.2 Electricity Markets Data 13

4.2.1 Secondary Operating Reserve Market 13

4.2.2 Intraday Continuous Spot Market 13

5 Model: FleetRL (20%) 13

5.1 Information Assumptions 13

5.2 Mobility Demand & Clearing Price Prediction 13

5.3 Reinforcement Learning Approach 13

5.4 Bidding Strategy 13

5.5 Dispatch Heuristic / Algorithm 13

6 Evaluation (30%) 13
6.1 Event-based Simulation 13
6.2 Benchmark: Ad-hoc Strategies 13
6.3 FleetRL 13
6.4 Sensitivity Analysis: Prediction Accuracy 13
6.5 Sensitivity Analysis: Infrastructure Changes 13
6.6 Sensitivity Analysis: Bidding Strategy 13

7 Discussion (5%) 13
7.1 Generalizability 13
7.2 Future Electricity Landscape 13
7.3 Limitations 13

8 Conclusion (5%) 13
8.1 Contribution 13
8.2 Future Research 13

9 References 13

List of Figures

1 Markov Decision Process 10

List of Tables

1 Raw Car2Go Trip Data from Stuttgart 11

1 Introduction (10%)

1.1 Research Motivation

- (Lopes et al., 2011)

1.2 Research Question

1.3 Relevance

2 Related Literature (10%)

2.1 Smart Charging and Balancing the Electric Grid with EV Fleets

The increasing penetration of EVs has a substantial effect on electricity consumption patterns. During charging periods, power flows and grid losses increase considerably and challenge the grid. Operators have to reinforce the grid to ensure that transformers and substations do not get overloaded (Sioshansi, 2012; Lopes et al., 2011). Loading multiple EVs in the same neighborhood, or worse, whole EV fleets at once, stress the grid. In these cases, even brown- or blackouts are possible (Kim et al., 2012). Despite these challenges, it is possible to postpone the physical reinforcement by adopting smart charging strategies. In smart charging, EVs get charged when the grid is less congested to achieve more grid stability. Smart charging reduces peaks in electricity demand, called *Peak Cutting* and complement the grid in times of low demand, called *Valley Filling*. Smart charging has been researched thoroughly in the IS literature, in the following we will outline some of the most important contributions.

Valogianni et al. (2014) find that using intelligent agents to schedule EV charging substantially reshapes the energy demand and reduces peak demand without violating individual household preferences. Moreover, they show that the proposed smart charging behavior reduced average energy prices and thus economically benefit households. In another study Kara et al. (2015) investigate the effect of smart charging on public charging stations in California. Controlling for arrival and departure times, the authors present beneficial results for the distribution system operator (DSO) and the owners of EVs. A price reduction in energy bills and a peak load reduction could be determined. An extension of the smart charging concept is Vehicle-to-Grid (V2G). When equipped with V2G devices, EVs can discharge their batteries back into the grid. Several authors researched this technology in respect to grid stabilization effects and arbitrage possibilities. Schill (2011) find that EVs can be beneficial for average consumer electricity prices. Excess EV battery capacity can be used to charge in off-peak hours and discharge in peak hours when the prices are higher. These arbitrage possibilities reverse welfare effects of generators and increase general overall welfare and consumer surplus. Tomić and Kempton (2007) show that the arbitrage opportunities are especially prominent when a high variability in electricity prices on the target electricity market exists. The authors state that short intervals between the contract of sale and the physical delivery of electricity increase arbitrage benefits. Consequently, ancillary service markets, like frequency control and operating reserve markets, are attractive for smart charging.

Peterson et al. (2010) investigate energy arbitrage profitability with V2G in the light of battery depreciation costs in the US. Their results indicate that large-scale use of EV batteries for grid storage does not yield enough profits to incentivize EV owners to participate in V2G activities. Considering battery depreciation cost, they arrive at an annual profit of only 6$ - 72$ per EV. Brandt et al. (2017) evaluated a business model for parking garage operators operating on the German frequency regulation market. When taking infrastructure costs and battery depreciation costs into account, they concluded that the proposed vehicle-grid integration is not profitable. Even with generous assumptions about EV adoption rates in Germany and altered auction mechanisms they arrived at negative profits. Kahlen et al. (2017) used EV fleets to offer balancing services to the grid. Evaluating the impact of V2G in their model the authors conclude that V2G would only be profitable if reserve power prices would be twice as high. Considering the results from the studies mentioned above this research does not include V2G is the model, only marginal profits are expected.

In order to maximize profits, it is essential for market participants to develop good bidding strategies. Successful bidding strategies to jointly participate in multiple markets have been developed, e.g., by Mashhour and Moghaddas-Tafreshi (2011). The authors use stationary battery storage to participate in the spinning reserve market and the day-ahead market at the same time. They developed a non-equilibrium model, which solves the presented mixed-integer program with Genetic Programming (GP). Contrarily, we use a model-free RL agent that learns an optimal policy (i.e., a trading strategy) from actions it takes in the environment (i.e., bidding on electricity markets). Using a model-free approach is especially beneficial for us, since additional unknown variables and constraints (i.e., customer mobility demand), make it complicated to formulate a mathematical model.

He et al. (2016) have done similar research to Mashhour and Moghaddas-Tafreshi (2011). The authors additionally incorporate battery life cycle in their profit maximization model, which proves to be a decisive factor. In contrast to the authors, we jointly participated in the secondary operating reserve and spot market with the *non-stationary* storage of EV batteries. Because shared EVs have to satisfy mobility demands, they have to be charged in any case, which allows us to safely exclude battery deprecation from our model. Further, we chose the intraday market over the day-ahead market, as it has the lowest reaction time of the spot markets, and thus potentially offers higher profits (Tomić & Kempton, 2007).

Previous studies often assume that car owners or households can directly trade on electricity markets. In reality, this is not possible due to the minimum capacity requirements of the markets, that single EVs do not meet. For example, the German Control Reserve Market (GCRM) has a minimum trading capacity of 1MW to 5MW, depending on the specific market. In order to reach the minimum capacity, over 200 EVs would need to be connected to the grid via a standard 4.6kW charging station at the same time. Ketter et al. (2013) introduced the notion of electricity brokers, aggregators that act on behalf of a group of individuals or households to participate in electricity markets. Brandt et al. (2017) and Kahlen et al. (2014) successfully showed in their studies that electricity brokers could overcome the capacity issues by aggregating EV batteries. In addition to electricity brokers, we apply the concept of Virtual Power Plants (VPPs). VPPs are flexible portfolios of distributed energy resources, which are presented with a single load profile to the system operator, making them eligible for market participation and ancillary service provisioning (Pudjianto et al., 2007). Hence, VPPs allow providing regulation capacity to the market without knowing which exact sources provide the promised capacity until the delivery time (Kahlen et al., 2017). This concept is specially useful when dealing with EV fleets: VPPs enable carsharing providers to issue bids and asks based on an estimate of available fleet capacity, without knowing beforehand which exact EVs will provide the capacity at the time of delivery. Based on the battery charge and the availability EVs, an intelligent agent will decide in real-time which vehicles provide the capacity.

Carsharing providers manage large EV fleets, which makes it possible for them to use the presented concepts as a viable business extension. Free float carsharing is a popular concept where which allows cars to be picked up and parked everywhere, and the customers are billed is by the minute. Free float carsharing offers more flexibility to its users, saves resources and reduces carbon emissions (Firnkorn & Müller, 2015). Most previous studies concerned with using EVs for electricity trading, assumed that trips are fixed and known in advance, e.g., in Tomić and Kempton (2007). The free float concept adds uncertainty and nondeterministic behavior, which make predictions about future rentals a complex issue.

Kahlen et al. (2017) showed that it is possible to use free float carsharing fleets as VPPs to profitably offer balancing services to the grid. The authors compared cases from three different cities across Europe and the US. They used an event-based simulation, bootstrapped with real-world carsharing and secondary operating reserve market data from the respective cities, to arrive at their results. A central dilemma within this research is to decide whether an EV should be committed to a VPP or to be free for rent, in the core a classification problem. Since rental profits are considerably higher than profits from electricity trading, it is crucial not to allocate an EV to a VPP when it could have been rented out otherwise. To deal with the asymmetric payoff, Kahlen et al. use stratified sampling in their classifier. This method gives rental misclassifications higher weights, which reduces the likelihood of EVs to participate in VPP activities. The authors use a Random Forest regression model to predict the available balancing capacity, which they use to place bids and asks on the market. An agent predicts available capacity on an aggregated fleet level, in order to leverage risk-pooling effects. Only at the delivery time the agent decides which EVs will provide the regulation capacity based on the likelihood that the vehicle is rented out and on the expected benefits of the EV. In a similar study, the authors showed that carsharing companies could participate in day-ahead markets for arbitrage purposes (Kahlen et al., 2018). In this paper, the authors use a time-series model to predict available trading capacity, due to less time between commitment and delivery. Another central problem for the carsharing provider is that committed trades which can not be fulfilled result in substantial penalties from the system operator or electricity exchange. In other words, it should be avoided at all costs, that the fleet commits to buy any amount of electricity, for which it does not have enough available EVs to charge it at the delivery time. To address this issue, the authors develop a mean asymmetric weighted (MAW) objective function. They use it for their time-series based prediction model, to penalize committing an EV to VPP when it would have been rented out otherwise. Because of the two issues mentioned above, Kahlen et al. (2018) can only make very conservative estimations and commitments of overall available trading capacity, which results in a high amount of foregone profits. This effect is especially prominent when participating in the secondary operating reserve market since commitments have to be made one week in advance when mobility demands are still uncertain. Kahlen et al. (2017) state that in 42% to 80% of the cases EVs are *not* committed to a VPP when it would have been profitable to do so.

This research is proposing a solution, in which the EV fleet participates in the balancing market and intraday market simultaneously. With this approach, we align the potentially higher profits on the balancing markets, with more accurate capacity predictions for intraday markets (Tomić & Kempton, 2007). This research followed Kahlen et al. (2017) with, who proposed to work on a combination of multiple markets in the future.

2.2 Reinforcement Learning in Smart Grids

Previous research showed that intelligent agents equipped with Reinforcement Learning methods can successfully take action in the smart grid. The following chapter outlines different research approaches of RL in the domain of smart grids. For a more thorough description, mathematical formulations and common issues of RL refer to Chapter 3.2.

Reddy and Veloso (2011a, 2011b) use autonomous broker agents to buy and sell electricity from DER on a proposed *Tariff Market*. The agents use Markov Decision Processes (MDPs) and RL to learn pricing strategies to profitably participate in the Tariff Market. To control for a large number of possible states in the domain, the authors used *Q-Learning* with derived state space features. Based on descriptive statistics, they define derived price and market participant features. By engaging with its environment, the agent learns an optional sequence of actions (policy) based on the state of the agent. Peters et al. (2013) build on that work and further enhance the method, by using function approximation. Function approximation allows to efficiently learn strategies over large state spaces, by deriving a function that describes the states instead of defining discrete states. By using this technique, the agent can adapt to arbitrary economic signals from its environment, which resulted in better performance than previous approaches. Moreover, the authors applied feature selection and regularization methods to explore the agent's adaption process. These methods are particularly beneficial in smart markets because market design, structures, and conditions might change in the future; hence intelligent agents should be able to adapt to it (Peters et al., 2013).

Vandael et al. (2015) facilitate learned EV fleet charging behavior to optimally purchase electricity on the day-ahead market. Similar to Kahlen et al. (2018) the problem is framed from the viewpoint of the aggregator which tries to define a cost-effective day-ahead charging plan in the absence of knowing EV charging parameters, like departure time. A crucial point of the study is weighting low charging prices against imbalance costs that have to be paid when an excessive or insufficient amount of electricity is bought from the market. Contrarily, Kahlen et al. (2018) don not consider imbalance cost in their model and avoid them by sacrificing EV mobility in order to balance the market. Vandael et al. (2015) use a *fitted Q Iteration* to control for continuous variables in their state and action space. In order to achieve faster convergence, they additionally optimize the *temperature step* parameter of the Boltzmann exploration probability.

Dusparic et al. (2013) proposed a multi-agent approach for residential demand response. The authors investigated a setting where 9 EVs were connected to the same transformer. The RL agents learned to charge at minimal costs, without overloading the transformer. Dusparic et al. (2013) utilized *W-Learning* to learn multiple policies (i.e., objectives like ensuring minimum battery charged or ensuring charging at low costs) at the same time. Taylor et al. (2014) extended this research by employing Transfer Learning and *Distributed W-Learning* to achieve communication between the learning processes of the agents in a multi-objective, multi-agent setting. Dauer et al. (2013) proposed a market-based EV fleet charging solution. The authors introduced a double-auction call market where agents trade the available transformer capacity, complying with the minimum required State of Charge (SoC). The participating EV agents autonomously learn their bidding strategy with standard *Q-Learning* and discrete state and action spaces. Giorgio et al. (2013) presented a multi-agent solution to minimizing charging costs of EVs, which required neither prior knowledge of electricity prices nor future price predictions. Similar to the previous study the authors employed standard *Q-Learning* and the ϵ-greedy approach for action selection. Vaya et al. (2014) also proposed a multi-agent approach, in which the individual EVs are agents that actively place bids in the spot market. Again, the agents use *Q-Learning*, with an ϵ-greedy policy to learn their optimal bidding strategy. The strategy relies on the agents willingness-to-pay which depends on the urgency to charge. State variables like SoC, time of departure and price development on the market determine the urgency to charge. The authors compared this approach with a centralized aggregator-based approach which they developed in another paper (Vaya & Andersson, 2015). Compared to the centralized approach, where the aggregator manages charging and places bids for the whole fleet, the multi-agent approach caused slightly more costs but solved scalability and privacy problems. Shi and Wong (2011) consider a V2G control problem, while assuming real-time pricing. The authors proposed an online learning algorithm which they modeled as a discrete-time MDP and solved through *Q-Learning*. The algorithm controls the V2G actions of the EV and can react to real-time price signals of the market. In this single-agent approach, the action space only compromises charging, discharging and regulatory actions, which makes it relatively easy to learn an optimal policy. Chis et al. (2016) looked at reducing the costs of charging for a single EV using known day-ahead prices and predicted next-day prices. A Bayesian ANN was employed for prediction and *fitted Q-Learning* was used to learn daily charging levels. In their research, the authors used function approximation and batch reinforcement learning, an offline, model-free learning method. Ko et al. (2018) proposed a centralized controller for managing V2G activities in multiple microgrids. The proposed method considers mobility and electricity demands of microgrids, as well as SoC of the EVs. The authors formulated a MDP with discrete state and action spaces and use standard *Q-Learning* with ϵ-greedy policy to derive an optimal charging policy, which takes microgrid autonomy and electricity prices into account.

It should be noted that advanced RL methods and techniques are not the only solutions for problems in the smart grid. Often basic algorithms and heuristics are good enough to solve the task. Despite that, this paper considers RL an optimal fit for the design of our proposed intelligent agent. Given the ability to learn user behavior (e.g., mobility demand) and the flexibility to adapt to the environment (e.g., electricity prices) RL methods are a promising way of solving complex challenges in smart grids (Vázquez-Canteli & Nagy, 2019).

3 Theoretical Background (10%)

3.1 Electricity Markets

3.1.1 Balancing Market

3.1.2 Spot Market

3.2 Reinforcement Learning

3.2.1 Notation

The input to the network $x \in \mathbb{R}^D$ is fed to the first residual layer to get the activation $y = x + \sigma(wx + b) \in \mathbb{R}^D$ with $w \in \mathbb{R}^{D \times D}$, and $b \in \mathbb{R}^D$ the weights and bias of the layer.

3.2.2 Introduction

(Vázquez-Canteli & Nagy, 2019).

Elements of Reinforcement Learning

1. Policy
    - Agent behavior at a given time
    - Mapping states to actions
    - Function or Lookup table
    - Sufficient to determine behavior
    - Policies may be stochastic, give probabilities for each action
2. Reward signal
    - Goal of the RL problem
    - Numeric signal the environment sends to the agent
    - Agents objective is to maximize the reward signal on the long run

   - Reward signal primary reason to change the policy: Low reward following an action of the policy may result in changing the policy to select another action
   - Rewards determine the immediate desirability of a state
   - Reward signals can be stochastic functions of the state and the actions

3. Value function

   - Value of a state is the total amount of reward an agent can expect to accumulate over the future, starting from that state
   - Values indicate the long-term desirability of states, taking future states and their rewards into account.
   - We seek actions cause states of highest value, because these actions obtain the greatest amount of reward in int long run.
   - Values must be estimated and re-estimated over the agents lifetime.
   - Efficiently estimating values is the most important component.

4. Model of the environment

   - Model allows inferences to be made about how the environment will behave. E.g., Given state and action the model predicts the next state and next reward.
   - Model-based methods are used for **Planning**: Deciding on a course of action by considering possible future situations before they happen.
     - **Control**: Model-free methods are simpler methods, what are explicitly trial-and-error learners

Planning vs. Control

On-Policy vs. Off-Policy

Policy Iteration

Value Iteration

3.2.3 Markov Decision Processes

- Classical formulation of sequential decision making
- Actions influence immediate rewards, but also future rewards
- Trade-off between immediate and delayed reward

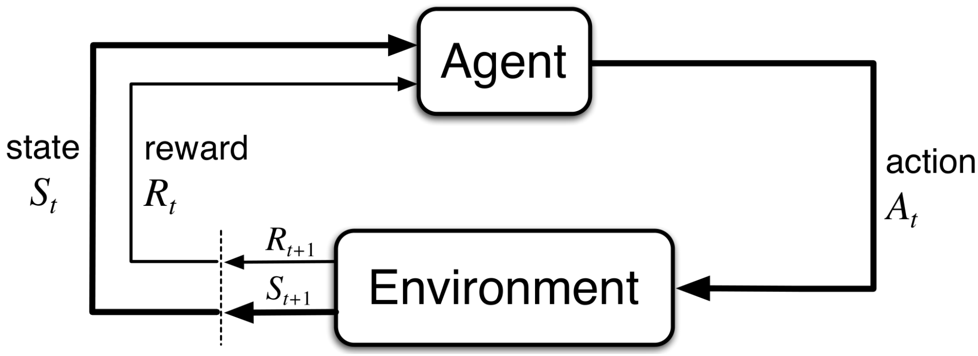

Figure 1: The agent-environment interaction in a Markov decision process (Sutton & Barto, 2018)

3.2.4 Tabular Methods

Dynamic Programming

Monte-Carlo Methods

Temporal-Difference Learning

1. TD Prediction
2. SARSA: On-policy TD Control
3. Q-learning: Off-policy TD Control

Q-Learning

3.2.5 Function Approximation

3.2.6 Exploitation-Exploration Trade-off

ϵ-greedy Method

3.2.7 Deep Reinforcement Learning

4 Empirical Setting / Data (10%)

4.1 Carsharing Fleets of Electric Vehicles

4.1.1 Raw Data

The dataset consists of 500 EVs in Stuttgart. As displayed in Table 1, the data contain spatio-temporal attributes, such as timestamp, coordinates, and address of the EVs. Additionally, status attributes of the interior and exterior are given, the relative state of charge and information whether the EV is plugged into one of the 200 charging stations in Stuttgart.

Table 1: Raw Car2Go Trip Data from Stuttgart

Number Plate	Latitude	Longitude	Street	Zip Code	Engine Type
S-GO2471	9.19121	48.68895	Parkplatz Flughafen	70692	electric
S-GO2471	9.15922	48.78848	Salzmannweg 3	70192	electric
S-GO2471	9.17496	48.74928	Felix-Dahn-Str.45	70597	electric
S-GO2471	9.17496	48.74928	Felix-Dahn-Str.45	70597	electric
S-GO2471	9.17496	48.74928	Felix-Dahn-Str.45	70597	electric
Number Plate	**Interior**	**Exterior**	**Timestamp**	**Charging**	**State of Charge**
S-GO2471	good	good	22.12.2017 20:10	no	94
S-GO2471	good	good	24.12.2017 23:05	no	72
S-GO2471	good	good	26.12.2017 00:40	yes	81
S-GO2471	good	good	26.12.2017 00:45	yes	83
S-GO2471	good	good	26.12.2017 00:50	yes	84

4.1.2 Preprocessing Steps

4.2 Electricity Markets Data

4.2.1 Secondary Operating Reserve Market

4.2.2 Intraday Continuous Spot Market

5 Model: FleetRL (20%)

5.1 Information Assumptions

5.2 Mobility Demand & Clearing Price Prediction

5.3 Reinforcement Learning Approach

5.4 Bidding Strategy

5.5 Dispatch Heuristic / Algorithm

6 Evaluation (30%)

6.1 Event-based Simulation

6.2 Benchmark: Ad-hoc Strategies

6.3 FleetRL

6.4 Sensitivity Analysis: Prediction Accuracy

6.5 Sensitivity Analysis: Infrastructure Changes

6.6 Sensitivity Analysis: Bidding Strategy

7 Discussion (5%)

7.1 Generalizability

7.2 Future Electricity Landscape

7.3 Limitations

8 Conclusion (5%)

8.1 Contribution

8.2 Future Research

9 References

vehicle-grid integration: Evidence from germany. *Transportation Research Part D: Transport and Environment*, 488-504. Retrieved from https://doi.org/10.1016/j.trd.2016.11.017 doi: 10.1016/j.trd.2016.11.017

Chis, A., Lunden, J., & Koivunen, V. (2016). Reinforcement learning-based plug-in electric vehicle charging with forecasted price. *IEEE Transactions on Vehicular Technology*. Retrieved from https://doi.org/10.1109/tvt.2016.2603536 doi: 10.1109/tvt.2016.2603536

Dauer, D., Flath, C. M., Strohle, P., & Weinhardt, C. (2013). Market-based ev charging coordination. In *Ieee/wic/acm international joint conferences on web intelligence (wi) and intelligent agent technologies (iat).* Retrieved from https://doi.org/10.1109/wi-iat.2013.97 doi: 10.1109/wi-iat.2013.97

Dusparic, I., Harris, C., Marinescu, A., Cahill, V., & Clarke, S. (2013). Multi-agent residential demand response based on load forecasting. In *Ieee conference on technologies for sustainability (sustech).* Retrieved from https://doi.org/10.1109/sustech.2013.6617303 doi: 10.1109/sustech.2013.6617303

Firnkorn, J., & Müller, M. (2015). Free-floating electric carsharing-fleets in smart cities: the dawning of a post-private car era in urban environments? *Environmental Science & Policy*, 30-40. Retrieved from https://doi.org/10.1016/j.envsci.2014.09.005 doi: 10.1016/j.envsci.2014.09.005

Giorgio, A. D., Liberati, F., & Pietrabissa, A. (2013, 12). On-board stochastic control of electric vehicle recharging. In *52nd ieee conference on decision and control.* Retrieved from https://doi.org/10.1109/cdc.2013.6760789 doi: 10.1109/cdc.2013.6760789

He, G., Chen, Q., Kang, C., Pinson, P., & Xia, Q. (2016). Optimal bidding strategy of battery storage in power markets considering performance-based regulation and battery cycle life. *IEEE Transactions on Smart Grid*, 2359-2367. Retrieved from https://doi.org/10.1109/tsg.2015.2424314 doi: 10.1109/tsg.2015.2424314

Kahlen, M., Ketter, W., & Gupta, A. (2017). Fleetpower: Creating virtual power plants in sustainable smart electricity markets.

Kahlen, M., Ketter, W., & van Dalen, J. (2014). Balancing with electric vehicles: a profitable business model.

Kahlen, M., Ketter, W., & van Dalen, J. (2018). Electric vehicle virtual power plant dilemma: Grid balancing versus customer mobility.

Kara, E. C., Macdonald, J. S., Black, D., Bérges, M., Hug, G., & Kiliccote, S. (2015). Estimating the benefits of electric vehicle smart charging at non-residential locations: a data-driven approach. *Applied Energy*, 515-525. Retrieved from `https://doi.org/10.1016/j.apenergy.2015.05.072` doi: 10.1016/j.apenergy.2015.05.072

Ketter, W., Collins, J., & Reddy, P. (2013). Power tac: a competitive economic simulation of the smart grid. *Energy Economics*, 262-270. Retrieved from `https://doi.org/10.1016/j.eneco.2013.04.015` doi: 10.1016/j.eneco.2013.04.015

Kim, E. L., Tabors, R. D., Stoddard, R. B., & Allmendinger, T. E. (2012). Carbitrage: Utility integration of electric vehicles and the smart grid. *The Electricity Journal*, 16-23. Retrieved from `https://doi.org/10.1016/j.tej.2012.02.002` doi: 10.1016/j.tej.2012.02.002

Ko, H., Pack, S., & Leung, V. C. M. (2018). Mobility-aware vehicle-to-grid control algorithm in microgrids. *IEEE Transactions on Intelligent Transportation Systems*. Retrieved from `https://doi.org/10.1109/tits.2018.2816935` doi: 10.1109/tits.2018.2816935

Lopes, J. A. P., Soares, F. J., & Almeida, P. M. R. (2011). Integration of electric vehicles in the electric power system. *Proceedings of the IEEE*, 168-183. Retrieved from `https://doi.org/10.1109/jproc.2010.2066250` doi: 10.1109/jproc.2010.2066250

Mashhour, E., & Moghaddas-Tafreshi, S. M. (2011). Bidding strategy of virtual power plant for participating in energy and spinning reserve markets-part II: Numerical analysis. *IEEE Transactions on Power Systems*, 957-964. Retrieved from `https://doi.org/10.1109/tpwrs.2010.2070883` doi: 10.1109/tpwrs.2010.2070883

Peters, M., Ketter, W., Saar-Tsechansky, M., & Collins, J. (2013). A reinforcement learning approach to autonomous decision-making in smart electricity markets. *Machine learning*, 5–39.

Peterson, S. B., Whitacre, J., & Apt, J. (2010). The economics of using plug-in hybrid electric vehicle battery packs for grid storage. *Journal of Power Sources*, 2377-2384. Retrieved from `https://doi.org/10.1016/j.jpowsour.2009.09.070` doi: 10.1016/j.jpowsour.2009.09.070

Pudjianto, D., Ramsay, C., & Strbac, G. (2007). Virtual power plant and system integration of distributed energy resources. *IET Renewable Power Generation*,

10. Retrieved from https://doi.org/10.1049/iet-rpg:20060023 doi: 10.1049/iet-rpg:20060023

Reddy, P. P., & Veloso, M. M. (2011a). Learned behaviors of multiple autonomous agents in smart grid markets. In *Aaai.*

Reddy, P. P., & Veloso, M. M. (2011b). Strategy learning for autonomous agents in smart grid markets. In *Ijcai proceedings-international joint conference on artificial intelligence.*

Schill, W.-P. (2011). Electric vehicles in imperfect electricity markets: the case of germany. *Energy Policy*, 6178-6189. Retrieved from https://doi.org/10.1016/j.enpol.2011.07.018 doi: 10.1016/j.enpol.2011.07.018

Shi, W., & Wong, V. W. (2011). Real-time vehicle-to-grid control algorithm under price uncertainty. In *Ieee international conference on smart grid communications (smartgridcomm).* Retrieved from https://doi.org/10.1109/smartgridcomm.2011.6102330 doi: 10.1109/smartgridcomm.2011.6102330

Sioshansi, R. (2012). The impacts of electricity tariffs on plug-in hybrid electric vehicle charging, costs, and emissions. *Operations Research*, 506-516. Retrieved from https://doi.org/10.1287/opre.1120.1038 doi: 10.1287/opre.1120.1038

Sutton, R. S., & Barto, A. G. (2018). *Reinforcement learning: An introduction.* MIT press.

Taylor, A., Dusparic, I., Galvan-Lopez, E., Clarke, S., & Cahill, V. (2014). Accelerating learning in multi-objective systems through transfer learning. In *International joint conference on neural networks (ijcnn).* Retrieved from https://doi.org/10.1109/ijcnn.2014.6889438 doi: 10.1109/ijcnn.2014.6889438

Tomić, J., & Kempton, W. (2007). Using fleets of electric-drive vehicles for grid support. *Journal of Power Sources*, 459-468. Retrieved from https://doi.org/10.1016/j.jpowsour.2007.03.010 doi: 10.1016/j.jpowsour.2007.03.010

Valogianni, K., Ketter, W., Collins, J., & Zhdanov, D. (2014). Effective management of electric vehicle storage using smart charging. In *Aaai* (pp. 472–478).

Vandael, S., Claessens, B., Ernst, D., Holvoet, T., & Deconinck, G. (2015). Reinforcement learning of heuristic ev fleet charging in a day-ahead electricity market. *IEEE Transactions on Smart Grid*, 1795-1805. Retrieved from https://doi.org/10.1109/tsg.2015.2393059 doi: 10.1109/tsg.2015.2393059

Vaya, M. G., & Andersson, G. (2015). Optimal bidding strategy of a plug-in electric vehicle aggregator in day-ahead electricity markets under uncertainty. *IEEE Transactions on Power Systems*. Retrieved from https://doi.org/10.1109/tpwrs.2014.2363159 doi: 10.1109/tpwrs.2014.2363159

Vaya, M. G., Rosello, L. B., & Andersson, G. (2014). Optimal bidding of plug-in electric vehicles in a market-based control setup. In *Power systems computation conference.* Retrieved from https://doi.org/10.1109/pscc.2014.7038108 doi: 10.1109/pscc.2014.7038108

Vázquez-Canteli, J. R., & Nagy, Z. (2019). Reinforcement learning for demand response: a review of algorithms and modeling techniques. *Applied Energy*, 1072-1089. Retrieved from https://doi.org/10.1016/j.apenergy.2018.11.002 doi: 10.1016/j.apenergy.2018.11.002