Bachelor Project

Czech
Technical
University
in Prague

F3

Faculty of Electrical Engineering
Department of Cybernetics

Data-Driven Sizing of Electric Vehicle Charging Stations

Vojtěch Jeřábek

Supervisor: Ing. Martin Schaefer
Field of study: Open Informatics
Subfield: Informatics and Computer Science
August 2020

BACHELOR'S THESIS ASSIGNMENT

I. Personal and study details

Student's name: **Jeřábek Vojtěch** Personal ID number: **474470**

Faculty / Institute: **Faculty of Electrical Engineering**

Department / Institute: **Department of Cybernetics**

Study program: **Open Informatics**

Branch of study: **Computer and Information Science**

II. Bachelor's thesis details

Bachelor's thesis title in English:

Data-Driven Sizing of Electric Vehicle Charging Stations

Bachelor's thesis title in Czech:

Optimalizace nabíjecích kapacit pro elektromobily

Guidelines:

The electric vehicles are becoming part of large company fleets. The companies are investing in installing charging stations in their facilities.
The historical fleet operation data can be analyzed to make more informed decisions about the sizing of the charging stations.
1. Research the related problems to multiple charging stations sizing.
2. Propose a usage of the detailed fleet operation data to optimize the size of charging stations.
3. Propose several hypothetical scenarios of fleet electrification process and evaluate the charging infrastructure sizing on these scenarios by simulation of charging station utilization.

Bibliography / sources:

[1] Yang, J., Dong, J. and Hu, L., 2017. A data-driven optimization-based approach for siting and sizing of electric taxi charging stations. Transportation Research Part C: Emerging Technologies, 77, pp.462-477.
[2] Mir Hassani, S.A. and Ebrazi, R., 2012. A flexible reformulation of the refueling station location problem. Transportation Science, 47(4), pp.617-628.
[3] Lam, A.Y., Leung, Y.W. and Chu, X., 2014. Electric vehicle charging station placement: Formulation, complexity, and solutions. IEEE Transactions on Smart Grid, 5(6), pp.2846-2856.

Name and workplace of bachelor's thesis supervisor:

Ing. Martin Schaefer, Artificial Intelligence Center, FEE

Name and workplace of second bachelor's thesis supervisor or consultant:

Date of bachelor's thesis assignment: **07.01.2020** Deadline for bachelor thesis submission: **14.08.2020**

Assignment valid until: **30.09.2021**

Ing. Martin Schaefer Supervisor's signature	doc. Ing. Tomáš Svoboda, Ph.D. Head of department's signature	prof. Mgr. Petr Páta, Ph.D. Dean's signature

III. Assignment receipt

The student acknowledges that the bachelor's thesis is an individual work. The student must produce his thesis without the assistance of others, with the exception of provided consultations. Within the bachelor's thesis, the author must state the names of consultants and include a list of references.

Date of assignment receipt	Student's signature

Acknowledgements

I would like to thank my supervisor Ing. Martin Schaefer for the patient guidance, helpful consultations, and useful comments. Anytime a problem occurred, he was always ready to help. He supported me and was always willing to help. In addition, I would like to thank my family and friends for their patience and constant support during my studies.

Declaration

I declare that the presented work was developed independently and that I have listed all sources of information used within it in accordance with the methodical instructions for observing the ethical principles in the preparation of university theses.

Prague, 11. August 2020

Abstract

Electric vehicles are becoming part of a large company fleets. More often, charging stations are installed into their facilities. This bachelor thesis deals with historical fleet data usage to make decisions about the sizing of the charging stations. In our thesis, we research related publications. On their basis, we predict an optimization of the sizing of charging stations in company's facility. The optimization method is tested for a different number of available chargers. Last, we predict two possible scenarios of fleet transition from vehicles with combustion engines to electric vehicles. We test the sizing of stations on unknown charging demand. Results show approximately a 10% higher success rate of optimized setup in comparison with uniform setup. The optimized setup handles unknown demand well, and its performance is only 3% lower than the maximal possible.

Keywords: optimization, electric vehicle, charging station, GPS trajectory data, transportation

Supervisor: Ing. Martin Schaefer
Department of Computer Science,
Faculty of Electrical Engineering,
Czech Technical University in Prague

Abstrakt

Elektrická vozidla se stávají součástí velkých flotil firemních vozidel. Stále častěji jsou do firemních areálů instalovány nabíjecí stanice. Tato bakalářská práce se zabývá využitím historických dat firemní flotily pro určení vhodné velikosti nabíjecích stanic. V naší práci prozkoumáváme existující publikace zabývající se podobnou tématikou. Na jejich základě navrhujeme metodu pro optimalizaci rozmístění nabíječek pro elektromobily mezi jednotlivé nabíjecí stanice v rámci jednoho areálu. Tuto optimalizační metodu testujeme pro různé množství stanic dostupných k rozmístění. Poslední části je návrh dvou možných scénářů, podle kterých by probíhal přechod flotily z vozidel se spalovacím motorem na elektrická vozidla. Rozmístění stanic testujeme na neznámé nabíjecí poptávce. Výsledky ukazují o přibližně 10% větší úspěsnost optimalizovaného rozmístění v porovnání s rozmístěním rovnoměrným. Optimalizované řešení nemá problém s neznámou poptávkou a rozdíl mezi jeho výsledky a maximálním možným výsledkem není větší než 3%.

Klíčová slova: optimalizace, elektrická vozidla, nabíjecí stanice, GPS trajektorie, doprava

Překlad názvu: Optimalizace nabíjecích kapacit pro elektromobily

Contents

Figures

Tables

Chapter 1

Introduction

Electromobility, specifically electric vehicles (EVs) have been on the rise during recent years because they are one of the possible alternatives to vehicles with combustion engines. Although it is not the goal of this thesis to decide which alternative to combustion engines is the right one, EVs have the potential to decrease our dependency on non-renewable resources and create a more sustainable environment.

As part of the transformation from combustion engines to electrical engines, changes in infrastructure are inevitable. With an increasing number of EVs, charging stations will have to increase the number of available chargers. If the charging station wants to be operated efficiently, the number of chargers should not be random. Using traffic data, we can estimate charging demand and offer a sizing solution for multiple charging stations with respect to budget limitation.

1.1 Goals and Motivation

Despite having a promising future, electric vehicles still have some disadvantages over combustion engines. EVs have a significantly lower range in

comparison with combustion engines. With 450 kilometers vs. 700 kilometers [2], and their range is more volatile to external influences, such as freeze or high temperatures [3]. Low range combined with infrequent charging infrastructure can lead to range-anxiety [4], fear that the vehicle has insufficient range to reach its destination. Range-anxiety and high price are considered to be one of the major barriers to large-scale adoption of EVs. While it takes approximately 5 minutes to fuel tank with petrol, even the fastest chargers will need at least 30 minutes [5] to charge a battery up to 80%. If a vehicle is charged from standard electricity output at home, charging time increases up to approximately 5 hours.

With different recharging habits and increased numbers of EVs emerges a problem with infrastructure. In places with available space, positioning of charging stations can be tailored to the needs of EVs. The situation will be far more complicated in cities and industrial facilities where almost every bit of free space is taken by existing infrastructure. In those places charging infrastructure will have to adapt to existing parking lots and fuel stations. In the case of industrial areas, the problem is slightly different. The owner or property manager is responsible for the placement of charging stations in those areas. They provide recharging services available for their in-house vehicles only. To achieve effective and economical service, optimization of charging stations capacity will be needed. Most scenarios of charging demand do not have linear traffic at charging stations. In cases with irregular traffic, charging demand distinctly differs during the day. Full coverage of such demand would require a large number of charging stations and thus become economically unbearable. Planners of recharging infrastructure have to find the optimal trade-off between charging demand satisfaction during rush hours and expenses on charging stations.

Figure 1.1: Chevrolet Volt plug-in hybrid charging in Fremont, California[1]

In this thesis, we set three goals. First is researching the related problems to multiple charging stations sizing. As a second goal, we predict optimizing charging stations size with the usage of detailed fleet operation data. Last is to introduce several hypothetical scenarios of fleet electrification process and evaluate the charging infrastructure sizing on these scenarios by simulation of charging station utilization. We can estimate charging demand and distribute available chargers among charging stations inside facility based on traffic data from potential charging stations. Based on daytime charging demand is changing according to traffic stereotypes such as rush hours, weekends, etc. Optimized capacities of charging stations provide the most effective distribution of a limited number of chargers.

[1]Picture is from Wikimedia Commons, url:`https://commons.wikimedia.org/wiki/File:Volt_charging_station.jpg`, published: 2012

1.2 Terminology

List of terms that are used in thesis. Terms are explained and put into context for better understanding of thesis.

- EV charger - Device used for charging of EVs, usually placed at designated parking space. Only one car at a time can be charged.
- charging station - Place with multiple EV chargers, for example parking lot.
- capacity of charging station - Number of chargers placed at charging station. Capacity should be large enough to meet most of the charging demand.
- charging demand - Number of cars that want to charge at charging station. If charging station has sufficient number of EV chargers the demand is satisfied otherwise not. Changing demand is changing during time. Demand is defined at the exact time.
- distribution of chargers - Placement of chargers among charging stations.
- industrial facility or facility - Area with several charging stations. Only one provider of recharging is present in the area. The provider manages all of the charging stations. Our optimization method distributes chargers between charging stations inside one facility. For example industrial facility with parking lots for employees. Only vehicles owned by the recharging provider are allowed to recharge at these charging stations.

1.3 Problem

Our problem is data-driven sizing of electric vehicle charging stations. We estimate charging demand from GPS traces of the company car fleet. We have to prepare an optimization method that distributes available chargers between possible charging station locations. Our solution is limited by existing infrastructure because we are building charging stations in an industrial area

with a high density of buildings. We can only choose from predefined places, usually parking lots. Traffic in the area of interest corresponds to the usual working hours, which means most of the cars are accumulated in parking lots between 8:00-15:00.

Experimenting with different number of available chargers is used to prove benefits of solution based on relevant data. Optimized setups are compared with simple setups to show the importance of optimization. Simple setups distribute chargers between charging station uniformly.

1.4 Thesis Structure

The remainder of this work is organized as follows. In Chapter 2, related literature is reviewed, summarised, and compared with specifics of our problem. In Chapter 3, the solution is introduced, and used data are described. Chapter 4 describes specifics of implementation and presents experiments with capacity and fleet. Chapter 5 concludes thesis with summary of results, possible improvements and future work on presented solutions.

Chapter 2

Related Work

Since introduction of EVs there is strong demand for finding ideal sizing and location of charging stations. Naturally, the academic community responded with a number of research projects evolving around electromobility. In 2018 there were 661 research studies focused on charging station of EVs [1]. According to [1], among those 661 publications, 119 are focused on sizing and localization of charging stations. This chapter wants to give readers an overview of existing scientific publications and describe different approaches. From Figure 2.1 is clear that publications focused on charging stations of electric vehicles are created almost all over the world. In countries like Iran [6] and China [7], studies are motivated by government goal to improve air quality. Research of related work shows that despite all publications having a similar goal of finding optimal size and location of charging stations, presented solutions are very different. We are using a division scheme from [1]. Related work is divided by those criteria: data source, target criteria, optimization process, and region of study.

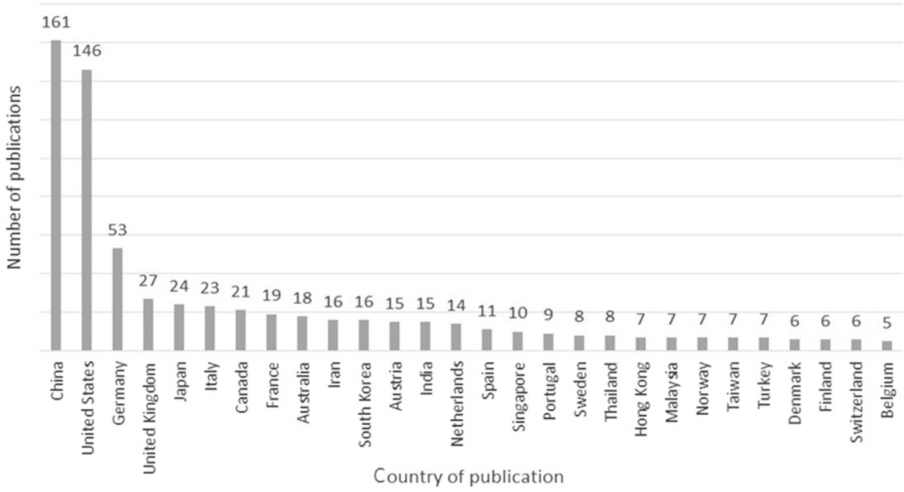

Figure 2.1: Number of publications from different countries [1]

2.1 Data Source

Input data are used for selecting the potential destination of charging stations and their size. Data can be retrieved from different sources. Those sources are divided into five categories: statistics, traffic data of EVs or at charging stations, travel data of combustion engine vehicles, surveys, and simulation or test data.

With statistics about wealth, age, population, income, number of cars, or occupation, we can estimate charging demand. For example, in [8], wealth, and population size in various areas of Hong Kong are used to predict the EV penetration rate. Charging demand is then calculated according to this information. In [9], authors combined sociodemographics, family size, ownership, and number of vehicles to characterize different user groups. According to [1], one statistic value for a group of users constrains the models, as they assume the same conditions without detailed spatial distinction.

Another source of data is traffic data of current EV and occupancy data of charging stations. Because it is data from EVs, we do not expect a change in driving patterns like in case of combustion engine car data. Analysis of

EV data is done in paper [10], and EVs can be used for modeling traffic on agent-based methods with modeling of their routes. But with occupancy data of charging stations, we know about arriving and departing cars from stations, but we have zero information about their trip.

GPS data from combustion engine vehicles (CEVs) offer better insight into the number of trips taken because CEVs still represent a majority of cars with 97% market share [11]. Publications based on CEVs data expect that driving patterns will not change if the transition to EVs is made. This approach is used in [12] and [13].

Surveys are another possibility mentioned in [1], but they seem obsolete in comparison with other methods of collecting input data. Questioning the drivers about their driving style, origin, and destination with diary is used in [14]. Drivers may share their expectations about charging stations' locations for further charging station location analysis.

2.2 Target Criteria

According to [1] goal to achieve optimal size and location of charging stations is shared among reviewed publications, but they differ in criteria by which the models determine the locations of stations. Different approaches are a result of various input data. Examples of selected criteria are demand density, the distance between user and charging station. Dwelling duration is time spent stationed at a potential charging station that can be used to determine charging demand. Traffic density criteria can be used to determine the number of cars per route. Other criteria are financial costs such as installation and operation costs of stations, energy consumption, or walking distances from charging stations. The most common criteria for sizing and locating of stations are demand density and financial costs with 26 and 28 out of 119 reviewed cases.

2.3 Optimization Process

According to [15] and [1] optimization algorithms used in publications can be categorized as genetic algorithms, particle swarm optimization, and integer linear programming. Based on information from [15] generic algorithm is an evolutionary algorithm that uses techniques inspired by natural evolution, such as inheritance, selection, mutation, and crossover, to find solutions to optimization problems.

Particle swarm optimization tries to improve a possible solution in every iteration with respect to given parameters. It represents problem as space with particles flying through it. Those particles are individual solutions to the problem. Each particle moves in the direction of its best known solution but also the course of the best known solution inside the searched space. The swarm of particles should move in the direction of the best solution. In comparison with other optimization methods, particle swarm optimization is easy to implement, converges faster, and can obtain an optimal solution with higher possibility. The optimization method is inspired by social behavior in bird flock [15].

Linear programming is a mathematical optimization program that aims to achieve the best outcome with non-integer variables. The objective of linear programming is to find the values of the linear equality or inequality constraints that maximize or minimize the linear function. Linear integer programming is a term in which objective functions and constraints are linear (integers). Papers [16], [8] and [17] are all using linear programming with some minor modifications. The linear programming problems are generally NP-hard.

2.4 Region of Study

At least a partial transition to electromobility is inevitable in every country. That is why we can see interest in finding optimal sizing and location of charging stations almost all around the world, as shown in Figure 2.1. Area

of interest in which studies try to place charging stations scales from parts of districts up to whole countries. In most of the region of the case of study is a city, city district, or metropolitan area [1]. For example [12], uses Changsha's city in China, and [13] is locating charging stations in the same country in the city of Shenzhen. In [17] locations for charging stations are optimized in Canadian province Quebec and in US state California. The majority of studies test their performance on an area with high population density. Rural regions are not represented.

We want to mention specific of our problem that varies from those presented in this chapter. Used data are from parking lots around multiple industrial factories that are located inside shared grounds. The researched area is not bigger than a few blocks of a city. Up to our best knowledge, there is not a publication focused on the area of similar size. The closest to our problem are publications focused on single city districts, but they do not use demand models similar to ours. Because our demand model is created from traffic at the workplace, it is affected by rush hours during the start and end of work.

Chapter 3

Methodology

The methodology chapter introduces the theoretical solution to our problem. First, we present solution design and describe each part of the process that leads to optimizing the sizing of charging stations. Second, we explain how we created the charging demand model and how we use it in simulation. Next, we formulate and solve the optimization problem. Then we explain how we evaluate the charging setup predicted by optimization with the help of simulation. Last, we interpret the simulation results, the conclusions drawn from results, and how they can be used for optimal sizing.

3.1 Solution Design

Our thesis aims to analyze historical fleet operation data to make more informed decisions about the sizing of the charging station. In the following section, we plan to introduce each part of our solution, which ultimately leads to charging stations' optimal size in cases with known demand. We plan to simulate charging demand on various configurations of charging stations' sizing. We expect that the traffic after the transition to EVs will be the same as before. For purposes of the thesis, we consider each stop of EV longer than 5 minutes as an opportunity for recharging. This means we consider

each stop in historical traffic data as a demand for charging. This demand represents the demand model. The optimization of charging stations' sizing is based on demand data.

The first step is to create a demand model. The demand model is created from historical data and replicates requests of cars from car fleet to charge at charging stations during some timespan. The charging station either has a free charger or not. This model can be created from GPS traces, a log of arriving and departing vehicles from parking lots or entrance gate. The next step is the optimization of charging stations' size for the given demand. We have multiple charging stations in the area of interest, and the budget for chargers is limited. We want to base the distribution of chargers between stations on an informed decision. The optimization method uses the demand model as input and finds the stations' optimal setup with respect to the efficiency of charging. Found setup is optimal only for a given demand. Then simulation starts testing found setup with different demand models to evaluate the flexibility of the setup. We expect real-world demand to be similar to the one predicted from historical data. The simulation creates a virtual representation of our area of interest with charging stations of various capacity placed over the area. After the simulation, results such as the number of charged cars, time spent charging, the amount of charged electricity in kWh and other statistics, are presented. The results of optimization can then be interpreted to the administrator of the charging stations site.

3.2 Demand Model

The demand model simulates EVs' behavior in the are of interest, how often do EVs need to charge, and how long does charging take. Each EV of the fleet has an individual charging demand due to a different volume of trips or trip length. The demand model provides information about the car, charging station, start, and end of charging. The simulation uses the model as input and evaluates whether each car can be charged or not. There are different sources from which the demand model can be created. For example, GPS traces, register of arriving and departing vehicles from entrance gate or parking lot or CCTV footage.

We predicate both demand models for optimization and evaluation of simulation from historical data. Those models are created from different parts of historical data to guarantee uninfluenced results. A similar approach is presented in the Related Work (Chapter 2). The approach is used in publications [12] and [13].

3.3 Optimization Problem

Our goal is to optimize charging stations capacity. With the proper distribution of charging stations, we should be able to **maximize the amount of distributed electricity**.

We will now introduce mathematical formulation and formalization of this problem which will be referred as Charging Stations Sizing Problem (CSSP).

Now follows description of instance of CSSP problem. CSSP finds optimal placement of P chargers for a set of locations L with size n. Placement of chargers is an assignment of m chargers to each location from L, where function $f(L, D)$ defines how much electricity is charged during demand D and configuration of locations L.

$$\underset{L}{\operatorname{argmax}}\, f(L, D)$$

subject to

$$\sum_{i=1}^{n} l_i = P$$

- P - Number of charging stations that can be placed to locations L.
- n - Number of available locations L.
- $L = \{l_1, ..., l_n\}$ - Set of numbers representing amount of charging stations placed at location i if $i \in \langle 1, n \rangle$.
- D - Demand data is a set of pairs with arrival and departure timestamps, each charging station has its own set of pairs.

- $f(L, D)$ - Function simulates how much energy (in kWh) is distributed with given demand D and configuration L.

We solve the optimization problem for a known charging demand. We plan to use optimal sizing found for known demand on cases with unknown demand. We estimate that with enough data for the demand model, we can make a versatile setup that performs successfully on unknown demand.

3.4 Vehicle-Charger Allocation

We simulate arrivals and departures of cars from several parking lots inside one facility. For example, we have an area with a factory, an office building and two parking lots. Those parking lots will be considered in simulation as charging stations with assigned capacity (number of chargers). Each arriving car is assigned to one charging station inside the area. At the charging station, vehicles are assigned to free chargers. Chargers allocation to cars that are requesting charting is based on First Come, First Served (FCFS) strategy. We have selected this strategy because it reflects real-world behavior at gas stations or parking lots. Other allocation strategies would require reservation of chargers and would give us options to decide who and when will be charging. We only choose how many chargers are going to be at charging stations. Either all or a selected number of cars are considered as EVs. Input data contain information about car ID, charging point ID, arrival timestamp, departure timestamp, and type of vehicle.

Evaluation is based on a loop that goes through demand data one row at a time. Each row represents a charging request from one car. A charging request is assigned to a specified charging station. If the charging station has a free charger, the request is considered as satisfied if the charger is not available; the request is marked as not-satisfied.

With every arriving vehicle at the charging station, simulation verifies if a free charger is available. The timestamp of arrival is checked with the departure timestamp of cars that are currently being charged. If any car is considered as departed, its former charger waits for the next car.

Overall statistics such as the total number of cars, number of satisfied cars, total kWh charged, total time spent charging, and others are calculated at the end of the evaluation.

For a better understanding of simulation, we have decided to write simple pseudocode of simulation described above.

Algorithm 1: Simulation

```
input : Traffic data, Capacity
initiateChargingPoints(Capacity);
for chargingRequest in Traffic do
    removeExpiredTimestamps(chargingRequest →
     departureTimestamp);
    if charger avail. at chargingRequest → chargingPointID then
        startCharging;
        updateStatistics;
    else
        updateStatistics;
    end
end
```

3.5 Results Interpretation

After we find the optimal setup for a given demand, we experiment with various demand models to prove the solution's robustness. If setup performance is solid on several demand models, we can expect it to perform successfully in real-life. Another criterium that we have to take into account is the electrification of the car fleet. The performance of found setup on a different number of available chargers is also an important indicator of the solution's versatility. The ability to compare the versatility of found solutions should help us to choose the best possible setup.

Chapter 4

Experiments

In this chapter, we describe our solution based on theory in Methodology (Chapter 3). First, we explain how the input data are preprocessed to a usable format. Then, we introduce the implementation of our solution, what coding language is used, and other details. The last part are experiments itself. In each experiment section, we describe how the experiment is done, what we expect to discover, and what experiment result means. These experiments are divided into two parts. The first part compares the performance of optimization with various budget sizes. The second part predicts two possible scenarios of car fleet transfer to electromobility.

4.1 Input Data

In this section, we describe input data used as parameters of simulation and procedures of acquiring them. We have received input data from an unnamed company, one of the largest employers in the Czech Republic. Data were collected during several months among the car fleet of the company. The input is GPS traces of car traffic, which must first be converted to compatible form. We use SQL scripts to convert the database of traces into final input. The second part of the input is information about the capacity charging

stations. Each charging station has assigned its amount of chargers. The number of chargers is either predefined or predicted by the optimization method.

4.1.1 Preprocessing Data

We have been working with GPS data from the company car fleet. Each record consists of car ID, latitude, longitude, and timestamp. GPS coordinates are recorded approximately every 5 seconds. We have received traffic data from March 2019 to September 2019.

Based on timestamp and location, we have detected stops in trajectories of each car. If a car does not move for a period of 5 minutes, the stop is considered as a potential for charging. Because of company policy regarding GPS tracking, trips are only recorded if they are work-related. This means we have to filter out stops that did not start and ended at the same place because private trip has been taken meanwhile.

To those stops left, the nearest potential charging point is assigned within the range of 200 meters. In Figure 4.1 are pictured locations of stops that are colored according to their nearest charging station. Each color on the map represents one charging station. Converted data have following structure car ID, charging point ID, timestamp of arrival at a charging point, timestamp of departure from charging point, and optional information whether it is EV or vehicle with a combustion engine. The data are sorted by arrival timestamp, so simulation browses through them chronologically. From converted data, we can calculate charging time and usage of each charging point.

4.1.2 Traffic Data

From the observation of data, we can clearly see that behavior in this working environment corresponds to the usual working hours. Most of the cars arrive between 8:00 and 9:00 and leave around 15:00. This can be seen in Figure 4.2, which represents the number of cars moving through all parking lots during

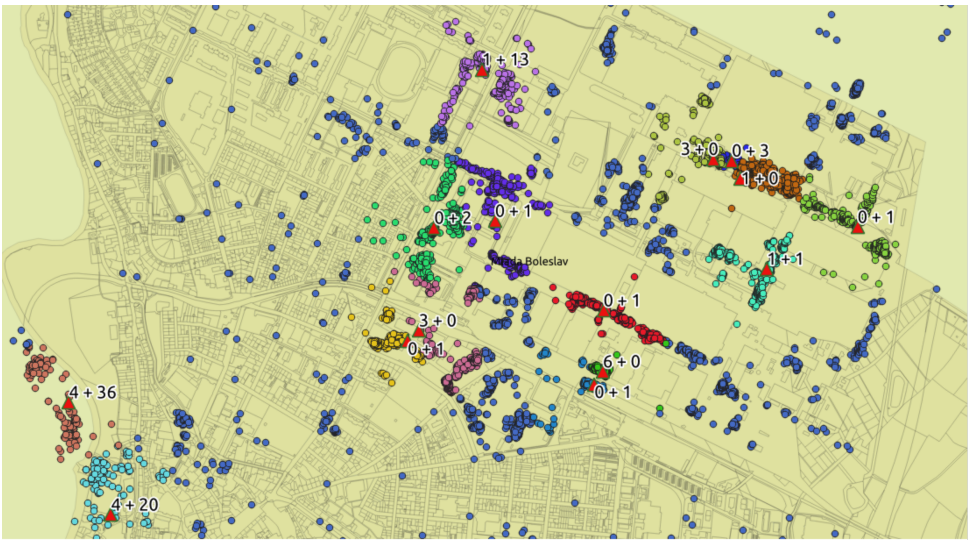

Figure 4.1: Map of industrial facility with stops assigned to the nearest charging point

the month of May displayed as one day. After 6:00 cars start accumulating at parking lots and then after 12:00 they start slowly departing.

Cars that arrive or depart during nights and weekends are negligible. Traffic on different parking lots/potential charging slots is consistent during different days of the workweek. This can easily be explained by the fact that employees are usually assigned to one workplace. The Figure 4.3 represents daily summary of traffic between stated dates. From Figure 4.3 is clearly visible low volume of cars arriving and departing during weekends. Traffic data containing weekends can be later excluded from dataset because they are irrelevant to optimization results.

4.1.3 Capacity Data

The second part of input data is a number of chargers assigned to each charging point. Capacity determines how many cars can be charged at the same time on one charging point. These numbers are either given by default or calculated using the optimization method. Capacity is usually limited by

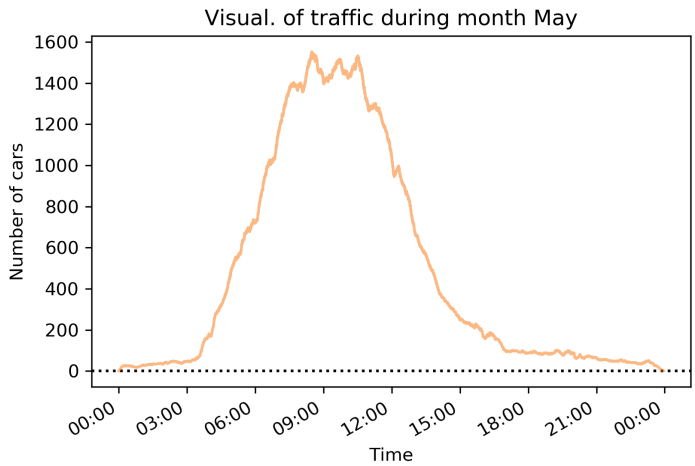

Figure 4.2: Graph with summary of month May displayed as one day

budget restrictions.

4.2 Implementation

From the solution design introduced in the Methodology (Chapter 3), it is clear that object-oriented language is the most suitable for our charging setup evaluator. We are using Python because it is versatile, easy to use, and fast to develop. The only disadvantage is its speed limitation, Python is considered as one of the slow ones among the most used programming languages. This might be a problem in the optimization section of our solution, but slower computational time is negligible because result is not needed in real-time. Thanks to Jupyter Notebook, we have live code, visualization, and explanatory text in the same file. All of our experiments are executed in notebook in which we import our code. For better visualization and management of data, we use Pandas, Numpy, and Matplotlib packages. The code consists of seven Python classes and one notebook.

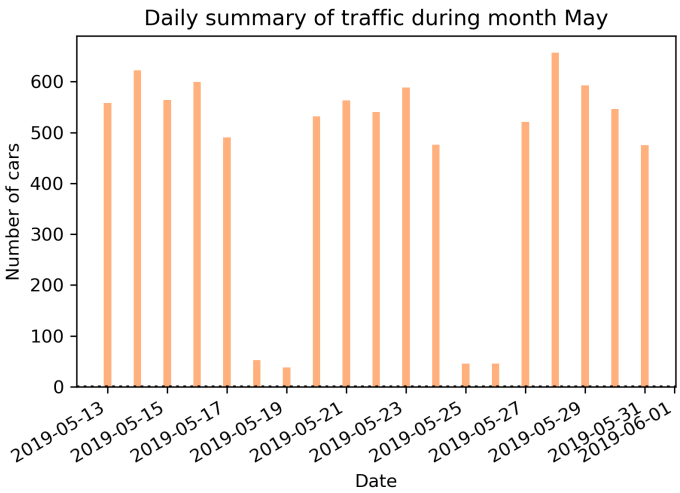

Figure 4.3: Bar of arriving and departing vehicles between Monday - 13.5.2019 and Saturday - 1.6.2019

4.3 Results

In the first part of experiments we aim to show the influence of budget size on the amount of delivered electricity to vehicles. Budget size is the number of available chargers. We start with the default number of chargers, and the number of chargers will be gradually lowered in each experiment.

4.3.1 Original capacity

In this experiment, we use the default number of chargers. The default number of chargers is 134, which is a number provided by the data source. Provider expects that 134 chargers will be enough to satisfy charging demand safely. We are going to verify that in the following experiment. Those chargers are divided between charging stations in different setups. We want to see the difference between the performance of the three setups. The first setup is a simple one where chargers are divided uniformly between charging stations. The second setup is optimized on the training dataset. The third setup is optimized on the testing dataset to give us an idea about maximum possible

performance. This means that the third setup is overfitted. The dataset contains 28 workweeks of charging demand. Four random workweeks are always selected as a testing dataset, and 24 workweeks are used as a training dataset.

We use the k-fold cross-validation method, where $n = 28$ and $k = 4$. Sequent workweeks are randomly shuffled for purposes of cross-validation. For each fold of cross-validation, all three of the setups mentioned above are created and evaluated. The measured criterium is the amount of charged electricity in kWh during a simulation. As input to the simulation is used k-number of workweeks of charging demand from demand model.

Trained setup performs almost as well as an overfitted setup (see Figure 4.4). This means that training and testing data do not contain any anomalies. Also, we do not have to worry about the overfitting of setup on training data and then to perform poorly on the testing set. The difference between overfitted and trained setup results is smaller than one percent. The uniform setup shows a decrease of charged electricity in comparison with a setup based on an informed decision. The Figure 4.4 shows a similar level of charged electricity in each of the seven datasets. This is a good sign that our optimization method has stable performance and can not be influenced by the dataset.

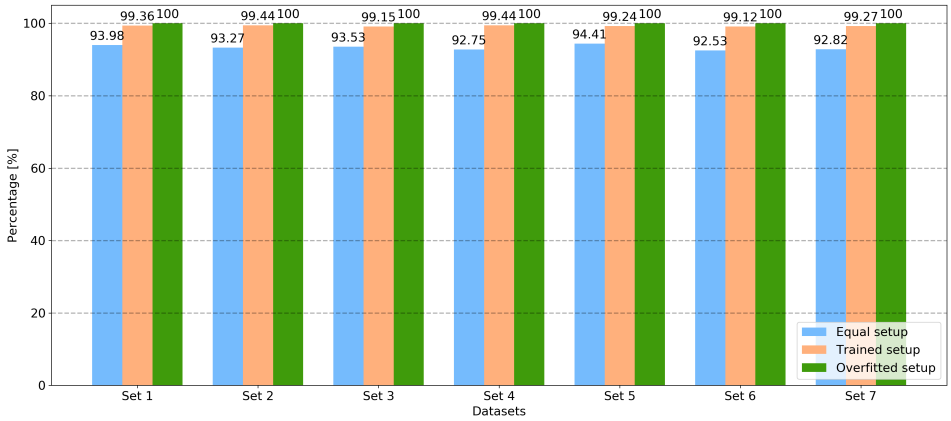

Figure 4.4: Performance of default budget size during cross-validation

4.3.2 Decreased Capacities

The following experiments are conducted to show the impact of decreased budget on the performance of the optimized setup. We are using 80%, 65%, and 50% share of original capacity dedicated to charging stations by the contracting authorities. We want to see if tradeoff between additional chargers and additional performance is directly proportional or not. The last monitored criterium is the non-versatility of optimal setup on testing data. The optimal setup of each capacity is compared with a setup where the same amount of charges is distributed uniformly between charging stations and with overfitted setup. We expect that the difference between uniform and optimized setup will grow with the increasing budget reduction.

All experiments are executed using k-fold cross-validation with the same parameters as the previous one with default capacity. We want to compare the performance of all capacities in Figure 4.5. Each capacity calculates average delivered electricity in kWh for all three setups. The biggest possible amount of delivered electricity represents overfitted setup with 100% of original capacity. The rest of the results unwinds from it. From Figure 4.5 is visible decrease of charged electricity with every decline of budget size, but the tradeoff is not directly proportional. With half of the original budget size, we can still deliver 75.83% of the electricity that can be delivered. Figure 4.6, show that the setup is trained correctly and thus do not perform poorly on testing data. That can be seen from the small difference between the performance of the overfitted and trained setup. Additional Figures A.1 and A.2 in the Appendix A contain results of setups with 65% and 80% of the original size.

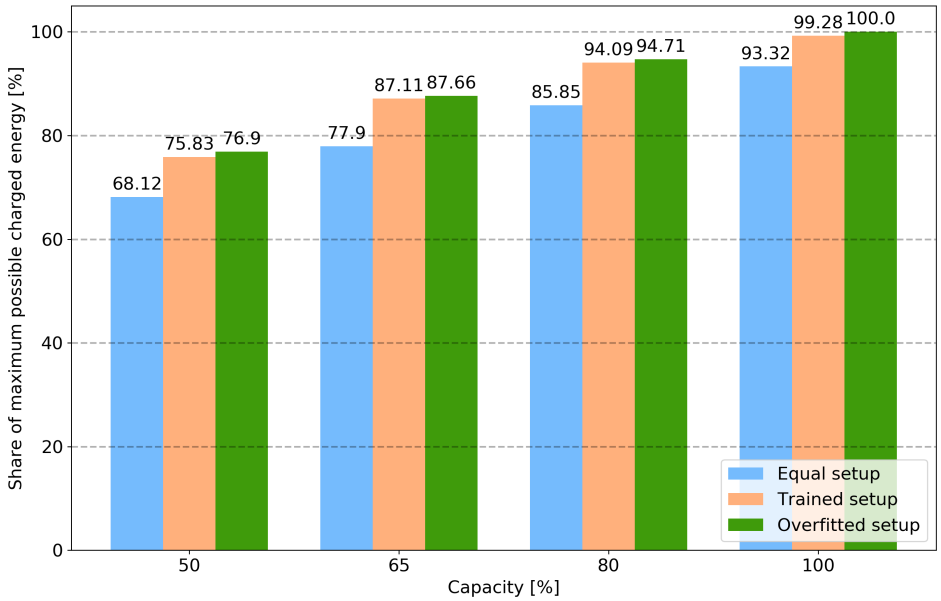

Figure 4.5: Comparison of performance of various budget sizes

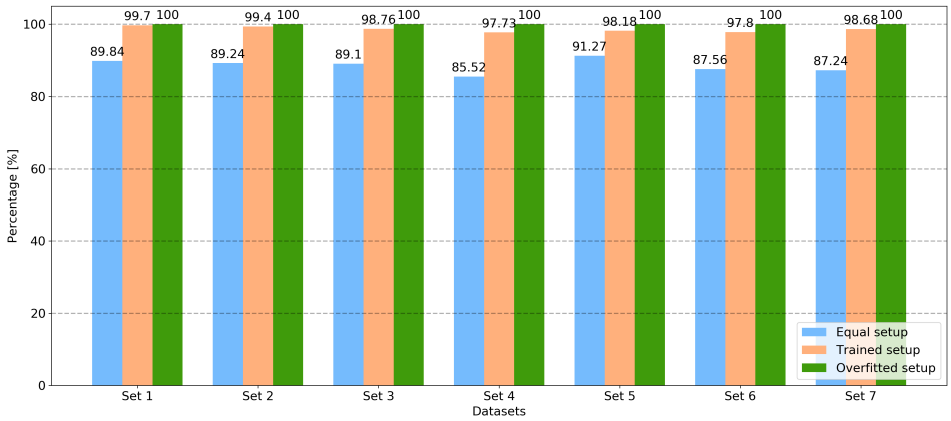

Figure 4.6: Performance of 50% of budget size during cross-validation

4.3.3 EV Transfer Scenarios

The second part of the experiments is a prediction of hypothetical scenarios of the fleet electrification process. Our goal is to evaluate the utilization of charging stations during these scenarios, compare utilization, and determine scenarios with the best utilization. We predict two possible scenarios based on available data and the expected behavior of car users. We always expect that at least 50% of the fleet consists of EVs. The first scenario is a random selection of 50% of the fleet and considering it as EVs. This case represents a situation where individuals choose between EV and combustion engine vehicles. The assignment of EVs is not based on an informed decision but imitates personal preferences. The second scenario is based on statistics about the stop length of each car. We consider each stop as an opportunity for charging. We expect 50% of the fleet with the longest average stops to be EVs. This case can occur if authority assigning the vehicles makes the transition using an informed decision.

For purposes of evaluation of transfer scenarios, we are uniformly distributing chargers among charging stations with default budget size. Scenarios are evaluated with a demand model that represents charging demand from March 2019 to September 2019.

A comparison of the two scenarios is in Table 4.1. Scenario with random selection has a slightly bigger total share of charged cars but has a lower total amount of charged electricity and time spent charging. While the difference between the percentage of charged vehicles is approximately 3%, the difference between charged power is around 10%. A scenario where electric vehicles will be assigned according to stop length offer better utilization of charging stations.

	Random 50%	**Selected 50%**
Charging time	1240 days, 12:52:55	1562 days, 16:26:12
Charged electricity [MWh]	506.14	637.58
Charged cars/total cars	32432/33200	31478/33200
Charged cars [%]	97.69	94.82

Table 4.1: Comparison of transfer scenarios

Chapter 5

Conclusion

This thesis explores one of the possible ways of data-driven sizing of electric vehicle charging stations. We have researched problems related to multiple charging stations sizing. Then we analyzed available input data. In our case, we have GPS traces of the car fleet. After that, we predicted our solution based on knowledge from related publications. That is how we created Charging Stations Sizing Problem or CSSP, an optimization method that uses fleet operation data to optimize charging stations' size. Specific aspects of our thesis are the ownership of charging stations, car fleet, and area of interest where optimization occurs. The area of interest is an industrial facility with a single owner who manages car fleet and charging stations. That is the reason why we maximize the utilization of the station and not the profits of charging. The last part of the thesis is a prediction of scenarios of fleet electrification process and evaluation of those scenarios. We predicted two possible scenarios.

Results show that with the use of optimization of charging stations sizing, we can achieve better utilization. To prove the benefits of informed decision, we have prepared a comparison with uniformly distributed setups and setups of different budget sizes. Experiments with varying sizes of budget show us the trade-off between additional charging stations and charged electricity. If optimization is used, we can achieve 75% of maximal possible utilization with only half the budget. Results of transfer scenarios show that scenario, where

the supervisor manages transfer to electric vehicles, leads to a larger amount of charged electricity and better utilization of stations. Electric vehicles are assigned to employees with a greater charging demand.

Our thesis can be used as a tool for companies that are interested in a transition to EVs. This tool gives them an idea of how the transition would affect their car fleet. A comparison of different budget sizes would provide a useful overview for management. An approximate estimate about the size of the initial investment into transition to EVs can be made from the performance of various budget sizes. The only thing required for EV transition analysis would be GPS traces or other data about traffic in the facility of the company.

In future work, we can add more features to our optimization model, thus creating a more advanced solution. In the Related Work (Chapter 2), we saw that similar solutions are modeling each car's trips to estimate battery capacity and thus simulate a more realistic recharging process. With data about speed and traveled distance of each car, we have a foundation to prepare a similar traffic model. Another future to be added later is the possibility of redirecting EV to the next nearest charging station. Redirecting vehicles may lead to better utilization of the whole set of charging stations. The next shortcoming to mention is the lack of transfer scenarios. Preparing additional transfer scenarios is another goal for future work.

In conclusion, the thesis provides a method for the optimization of charging stations' size. This method is based on historical fleet data analysis and designed for usage inside one area of interest. Results show that optimization of sizing can help to achieve more effective and economical service of charging stations.

Bibliography

[1] R. Pagany, L. R. Camargo, and W. Dorner, "A review of spatial localization methodologies for the electric vehicle charging infrastructure," *International Journal of Sustainable Transportation*, vol. 13, no. 6, pp. 433–449, 2019.

[2] S. Huntingford, "What car? real range: which electric car can go farthest in the real world?." `https://bit.ly/30VqV6Z`, 2019. Accessed: 2020-07-10.

[3] U. E. Department, "Maximizing electric cars' range in extreme temperatures." `https://bit.ly/3gWgJAx`, 2019. Accessed: 2020-05-08.

[4] N. Rauh, T. Franke, and J. F. Krems, "Understanding the impact of electric vehicle driving experience on range anxiety," *Human factors*, vol. 57, no. 1, pp. 177–187, 2015.

[5] S. Guinn, "Level 1 vs level 2 ev charging stations." `https://www.clippercreek.com/level-1-level-2-charging-stations/`, 2017. Accessed: 2020-03-25.

[6] A. Rajabi-Ghahnavieh and P. Sadeghi-Barzani, "Optimal zonal fast-charging station placement considering urban traffic circulation," *IEEE Transactions on Vehicular Technology*, vol. 66, no. 1, pp. 45–56, 2017.

[7] Yaping Gao and Xiaochun Lu, "Analysis of the electric vehicle charging station's layout in beijing based on improved traveling simulation model," in *2015 International Conference on Logistics, Informatics and Service Sciences (LISS)*, pp. 1–6, 2015.

[8] A. Y. Lam, Y.-W. Leung, and X. Chu, "Electric vehicle charging station placement: Formulation, complexity, and solutions," *IEEE Transactions on Smart Grid*, vol. 5, no. 6, pp. 2846–2856, 2014.

[9] A. Namdeo, A. Tiwary, and R. Dziurla, "Spatial planning of public charging points using multi-dimensional analysis of early adopters of electric vehicles for a city region," *Technological Forecasting and Social Change*, vol. 89, pp. 188 – 200, 2014.

[10] Y. Ahn and H. Yeo, "An analytical planning model to estimate the optimal density of charging stations for electric vehicles," *PLOS ONE*, vol. 10, pp. 1–26, 11 2015.

[11] E. E. Agency, "Electric vehicles as a proportion of the total fleet." `https://www.eea.europa.eu/data-and-maps/indicators/proportion-of-vehicle-fleet-meeting-4/assessment-4`, 2019. Accessed: 2020-04-13.

[12] J. Yang, J. Dong, and L. Hu, "A data-driven optimization-based approach for siting and sizing of electric taxi charging stations," *Transportation Research Part C: Emerging Technologies*, vol. 77, pp. 462–477, 2017.

[13] W. Tu, Q. Li, Z. Fang, S.-l. Shaw, B. Zhou, and X. Chang, "Optimizing the locations of electric taxi charging stations: A spatial–temporal demand coverage approach," *Transportation Research Part C: Emerging Technologies*, vol. 65, pp. 172–189, 2016.

[14] L. Zhang, B. Shaffer, T. Brown, and G. S. Samuelsen], "The optimization of dc fast charging deployment in california," *Applied Energy*, vol. 157, pp. 111 – 122, 2015.

[15] H. Shareef, M. M. Islam, and A. Mohamed, "A review of the stage-of-the-art charging technologies, placement methodologies, and impacts of electric vehicles," *Renewable and Sustainable Energy Reviews*, vol. 64, pp. 403 – 420, 2016.

[16] S. MirHassani and R. Ebrazi, "A flexible reformulation of the refueling station location problem," *Transportation Science*, vol. 47, no. 4, pp. 617–628, 2013.

[17] M. F. Anjos, B. Gendron, and M. Joyce-Moniz, "Increasing electric vehicle adoption through the optimal deployment of fast-charging stations for local and long-distance travel," *European Journal of Operational Research*, vol. 285, no. 1, pp. 263 – 278, 2020.

Appendix A

Additional Results

In this appendix, we present the rest of the results of experiments with different capacities from Experiments (Chapter 4).

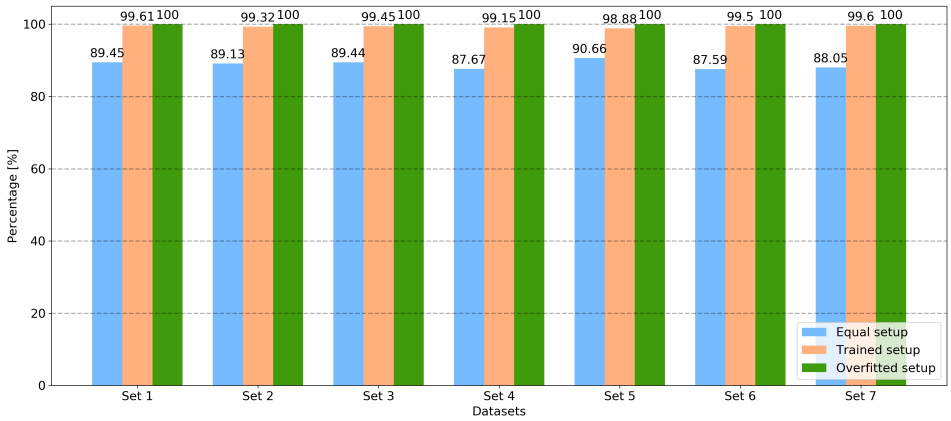

Figure A.1: Performance of 65% of budget size during cross-validation

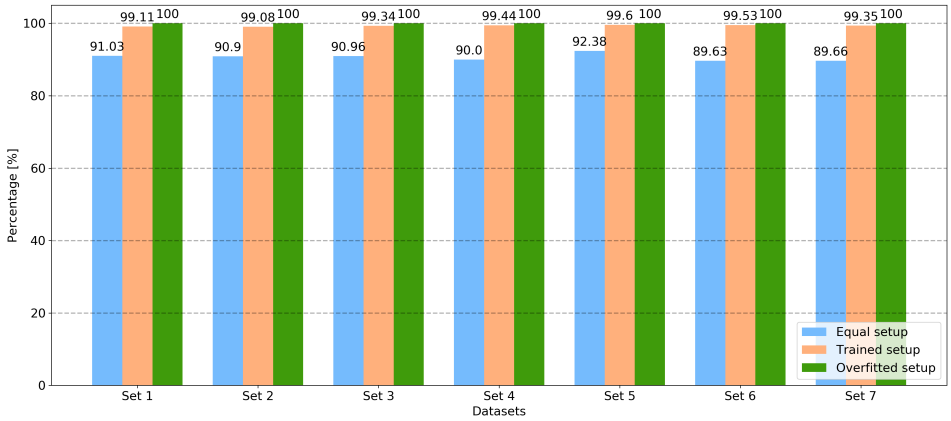

Figure A.2: Performance of 80% of budget size during cross-validation

Appendix B

User Guide

In this appendix is structure of attached files together with desription of each part. Because of the data provider's policy, we are providing only sample of input data. Contact thesis supervisor Ing. Martin Schaefer for information about the rest of the input data.

```
Jerabek_thesis_attachment
├─ code
│  └─ my-sim
│     ├─ input
│     └─ ntb.ipynb
├─ images
└─ thesis_text
```

my-sim Contains source files of our Python project ***my-sim*** used for evaluation of sizing.

ntb.ipynb Jupyter Notebook containing traffic data analysis and experiments. Notebook is fully run-able.

input Traffic data used as input to notebook.

images Images used in thesis and images that are result of previous traffic data analysis.

thesis_text Contains PDF version of this thesis.