Periodica Polytechnica
Electrical Engineering
and Computer Science

58(4), pp. 175-181, 2014
DOI: 10.3311/PPee.7761
Creative Commons Attribution

RESEARCH ARTICLE

Stochastic Modeling of Electric Car Charging Station for a Taxi Fleet

Csaba Farkas[1] * / *András* Dán[2]

RECEIVED 16 OCTOBER 2014; ACCEPTED 07 JANUARY 2015

Abstract

The uninterrupted operation of electric cars requires an adequate number of chargers installed in charging stations. The incoming requests determine the required number of chargers in the station, so it is of utmost importance to have practicable car motion models for planning purposes.

In this paper a stochastic model is proposed to simulate the behavior of a taxi fleet consisting 100 identical cars. Car usage characteristics were extracted from real life measurement data and then synthetic driving cycles were generated by Monte Carlo simulation. The state of charge of every car can be monitored and the number of cars that have to wait for being charged can be determined using the proposed algorithm. Parameters can be varied freely, so the model is adequate for modeling different datasets.

Keywords
electric car · charging station · Monte Carlo simulation · stochastic model

1 Introduction

Electric vehicles (EVs) require recharging, just as cars with internal combustion engines need to be refuelled. However, EVs take longer time to be recharged, so the planning of a charging station requires precise models in order to avoid unwanted and annoying queues and waiting at the very station itself. A deterministic model for such planning purpose does not come into question, as the number of cars is large and they all have different characteristics: driving distances, driving style, consumption, etc. all vary from car to car. Therefore, a stochastic model is required, capable of capturing the relevant features of electric car usage and with which the required number of chargers can be determined.

In this paper we present a stochastic model applied to a taxi fleet, based on real traffic data. Car usage characteristics were extracted from measured GPS trajectories [1,2] and Monte Carlo simulation was used to generate synthetic driving cycles. The stochastic model was implemented in MATLAB.

2 The stochastic model

The core of the proposed stochastic model is a Markov-chain-like algorithm that utilizes transition matrices [3]: the probability of transition from a given state to another one is given in every simulation time step (computed using the statistics of the measurement data) and the algorithm computes the state of every car for the next time step according to the actual transition matrix.

Tab. 1. Transition matrix

state/state	move	park without charging	park and charge
move	a_{11}	a_{12}	a_{13}
park without charging	a_{21}	a_{22}	a_{23}
park and charge	a_{31}	a_{32}	a_{33}

There are three different states of a car: it can either move, park without charging or park and charge. The transition matrix for a given time interval and for a given car can be seen in Table 1.

[1] Department of Electric Power Engineering
Faculty of Electrical Engineering and Informatics,
Budapest University of Technology and Economics
H-1111 Budapest, Egry József str. 18., Hungary
* Corresponding author, e-mail: farkas.csaba@vet.bme.hu

[2] Department of Electric Power Engineering
Faculty of Electrical Engineering and Informatics,
Budapest University of Technology and Economics
H-1111 Budapest, Egry József str. 18., Hungary
e-mail: dan.andras@vet.bme.hu

For example, a_{11} is the probability that a moving car will be moving in the next time interval as well. The transition matrix is stochastic, that is, the sum of each row is 1. There are some simplifications regarding the matrix, though: it is assumed that a transition from moving or parking state to charging state never occurs by accident (after all, charging requires a specific location, namely, the charging station), so a_{13} and a_{23} both equal to 0. Similarly, a charging car never stops charging and starts moving by accident, so a_{31} and a_{32} also equal to 0. This means, that $a_{33}=1$. The rest of the matrix elements were calculated from the measurement data referenced in [1] and [2]. This matrix can be called the „initial transition matrix” (note that this matrix is time dependent: the values are determined for every time step; see Fig. 7).

These values are calculated for every time interval and for every car: this means, that the transition matrix is in fact a 4-D matrix. The first two dimensions contain the transition probabilities that can be seen in Table 1, the third dimension is time and the fourth is the no. of the car (i.e. for every single car, for every time interval there exists a unique transition matrix). The algorithm then overwrites corresponding matrix elements if it is necessary: for example, if a car starts moving, then a travelling time is associated to it (raffled by Monte Carlo simulation) and it will not stop until this travel time is reached up. This means that the elements of the transition matrix have to be overwritten: a_{11} should be 1 and the rest of the elements in the first row should be 0 as long as the car is moving. If the travel time is up, the car stops and its transition matrix for the next time step is restored to the originally determined one. Section 3 gives more details about possible car state transition changes.

2.1 Car usage characteristics

The proposed model requires several types of input data: the most important one is the distribution of travel distance and travel time. Our model utilizes travel time instead of distance, so we are going to deal with time in first place. These data were obtained from [1] and [2] by statistical analysis (again, MATLAB was used to extract the necessary statistical parameters). The database contains collected data from 10357 taxis from New York, with more than 15 million GPS coordinates. Taxis were monitored for 7 days.

The ratio of moving cars during a day can be seen on Fig. 1.

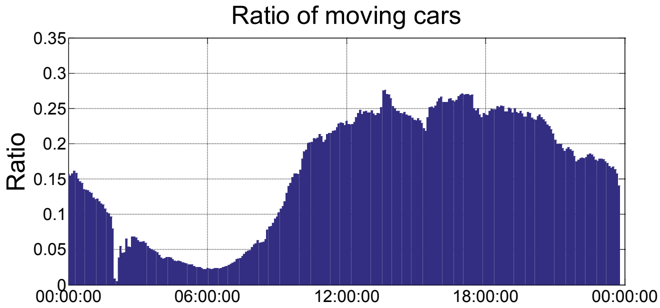

Fig. 1. Ratio of moving cars during a day

It can be seen that the ratio of moving cars is relatively small in the small hours, but it continues to increase and sets to a relatively constant value during the day to decrease again in the night. It was not taken into consideration, that the cars are operated for two shifts a day: with that consideration, the ratios are roughly the double of the ones that can be seen on Fig. 1. Similar results were obtained in the case of a Hungarian dataset composed of 50 taxis [4] (unfortunately, the number of cars and the monitored days were too few in the case of this database, so it could not be taken as a representative sample; nonetheless, the behaviour of taxis follows a similar pattern as for the New York taxis):

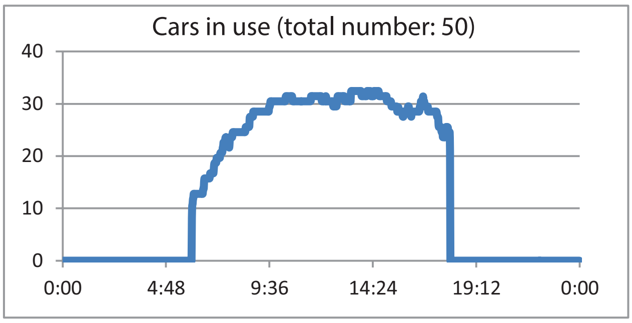

Fig. 2. Ratio of moving cars, Hungarian dataset

In this case, measurement data from only 6:00 to 18:00 was available.

2.2 Fitting of distribution functions

For Monte Carlo simulation purposes, the distribution functions of travel times and travel distances are required (our algorithm does calculations based on time, so we do not deal with travel distances much). According to references [5-18], the distribution of travelled distance is exponential or follows Lévy-distribution. About travel time distribution, the literature says almost nothing, so we tried to fit various types of distributions by using MATLAB, Wolfram Mathematica and Easy Fit StatAssist [19] (this latter allows calculations for datasets with less than 250000 points, so for the large dataset, it was not used). The following figures show the results of the fitting:

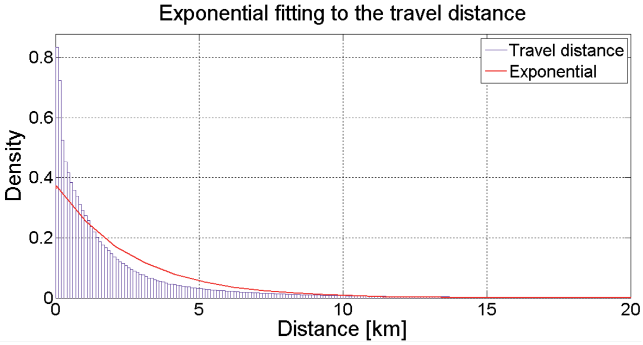

Fig. 3. Exponential fitting to the travel distance

The fitting results are the following:

```
Distribution:            Exponential
Log likelihood:          -3.05385e+006
Domain:                  0 <= y < Inf
Mean:                    2.64493
Variance:                6.99563

Parameter Estimate       Std.Err.
mu         2.64493       0.00212576

Estimated covariance of parameter estimates:
   mu
mu 4.51885e-006
```

Next, we investigated the Lévy-distribution:

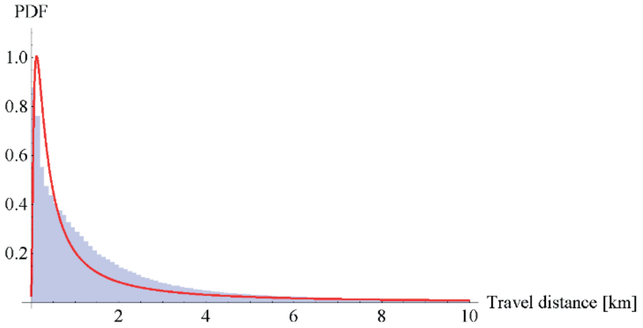

Fig. 4. Lévy-fitting to the travel distance

For travel times, we checked - among others - inverse Gaussian and lognormal distributions (the former was seen to best describe the given dataset).

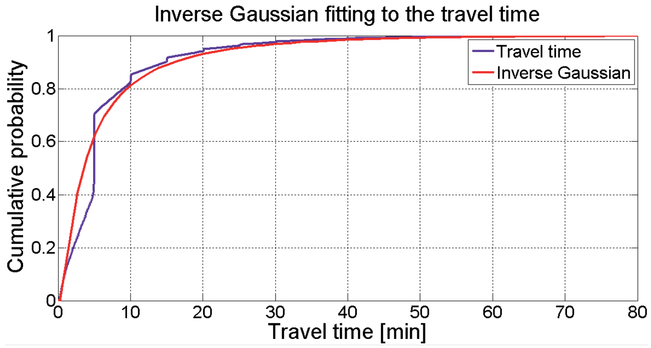

Fig. 5. Exponential fitting to the travel distence

The cumulative probabilities of these two types of distributions can be seen on Fig. 5 and Fig. 6, respectively. They fit well to the original dataset.

The parameter fitting resulted in the following values:

```
Distribution:            Inverse Gaussian
Log likelihood:          -3.14912e+006
Domain:                  0 < y < Inf
Mean:                    6.7781
Variance:                90.6083

Parameter       Estimate       Std. Err.
mu              6.7781         0.00768293
lambda          3.43681        0.00392237

Estimated covariance of parameter
estimates:
                mu                  lambda
mu        5.90274e-005        -3.87709e-010
lambda   -3.87709e-010          1.5385e-005
```

so λ=3,436, μ=6,778, where λ is the mean and μ is the shape parameter. The probability density function of the inverse Gaussian distribution is as follows:

$$f(x;\mu,\lambda)=\left(\lambda/\left(2\pi\cdot x^{3}\right)\right)\cdot\exp\left(\left(-\lambda\cdot(x-\mu)^{2}\right)/\left(2\cdot\mu^{2}\cdot x\right)\right)$$

Next we investigated the lognormal distribution:

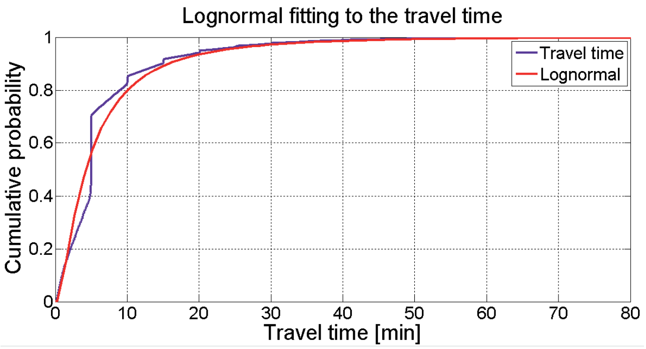

Fig. 6. CDF of lognormal fitting to travel time

The parameter fitting resulted in the following values:

```
Distribution:            Lognormal
Log likelihood:          -4.43414e+006
Domain:                  0 < y < Inf
Mean:                    7.16433
Variance:                94.9972

Parameter       Estimate       Std. Err.
mu              1.44532        0.000825951
sigma           1.02352        0.000584036

Estimated covariance of parameter
estimates:
                mu                  sigma
mu        6.82194e-007        -6.84504e-019
sigma    -6.84504e-019           3.41097e-007
```

so σ=1.02, μ=1.44 , where σ is the shape and μ is the log-scale parameter. The probability density function of the lognormal distribution is as follows:

$$f(x;\mu,\lambda)=1/\left(x\cdot\sqrt{2\pi\cdot\sigma}\right)\cdot\exp\left(-(\ln x-\mu)^{2}/\left(2\sigma^{2}\right)\right)$$

2.3 Transition probabilities

In the beginning of Section 2, we have already introduced transition matrices, as the core elements of our algorithm. The figures on Fig. 7 present the transition probabilities in a graphical way (some numerical errors can be seen in the small hours (see the pike), but it can be neglected).

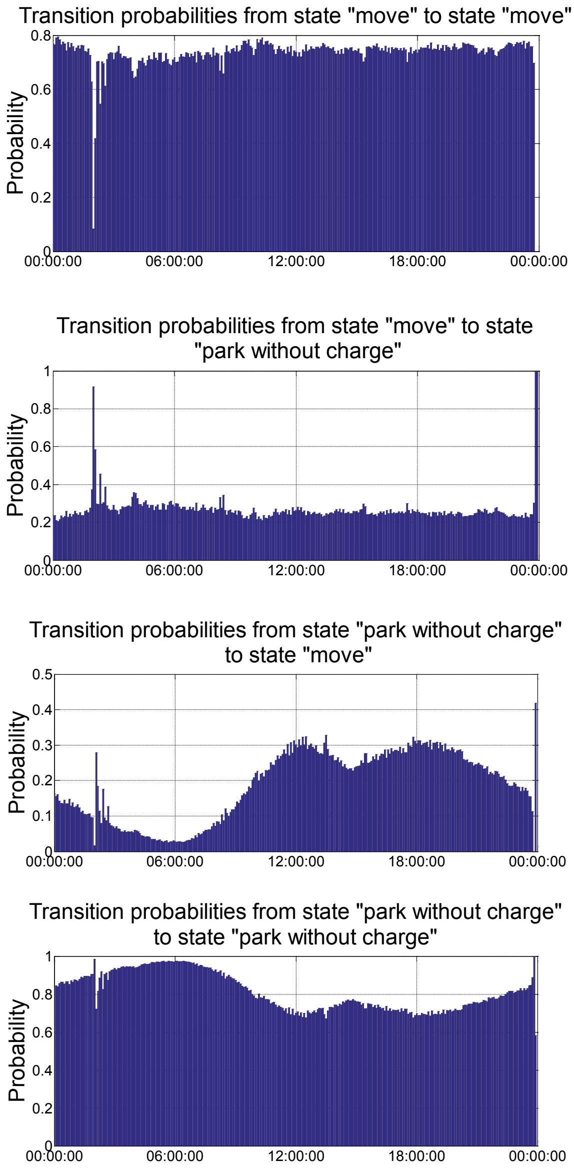

Fig. 7. Graphical representation of the transition probabilities

The transition probabilities (i.e. the matrix elements) were extracted from the raw GPS data in the following way: every car got an index of 1 if it was moving in the investigated time interval (based on similar investigations found in the literature [5-18]) a car was considered to be moving if its speed was between 1m/s and 34m/s: values out of this range were rejected) and an index of 0, if it was parking. By counting the $0 \rightarrow 0$, $0 \rightarrow 1$ $1 \rightarrow 0$ and $1 \rightarrow 1$ transitions, we could determine the matrix elements for the given time interval.

2.4 Energy consumption

The energy consumption of EVs depends on various factors, such as driving style, use of air-conditioners/heating, etc. To grasp the diversity of car consumption, a statistical approach is required once again: [20] presents consumption data for various electric vehicles and by fitting a distribution function to the dataset, we can conduct Monte Carlo simulations to obtain energy consumption values for every car and for every time interval. EasyFit StatAssist [19] was used for fitting purposes and as it can also rank the fitted distributions in goodness according to the Kolmogorov-Smirnov, Anderson-Darling and Chi-Squared tests, so we could choose from the distributions easier.

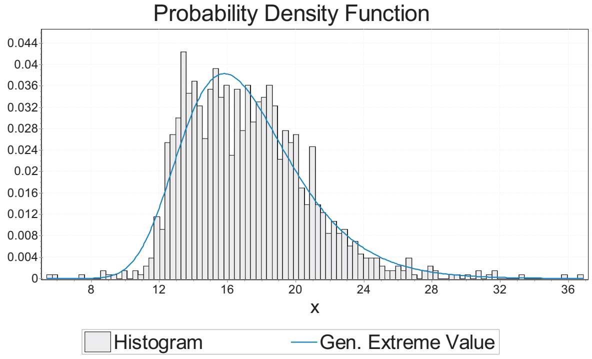

Fig. 8. Graphical representation of the transition probabilities Generalized extreme value distribution fitting to consumption

We chose generalized extreme value distribution with parameters $\xi = -0{,}06923$, $\sigma = 3{,}0{,}68$, $\mu = 15{,}603$ where μ is the location, σ is the scale and ξ is the shape parameter. The fitted curve can be seen on Fig. 8.

The consumption of cars is given in kWh/100km in [20], but our algorithm calculates consumption as the ratio of battery capacity, thus some recalculation is needed: if the capacity of the fully charged battery is *Cap* in kWh, the consumption is y in kWh and the time step is *tstep* in hours, then the consumed electricity, as the ration of battery capacity is as follows:

$$1-(Cap - y \cdot tstep)/Cap$$

that is, after simplification

$$y \cdot tstep/Cap$$

We supposed that the batteries in our simulation had 22kWh capacity (this is common for many car types, e.g. Nissan Leaf). By utilizing Monte Carlo simulation, a new consumption value was raffled for every car in every time step.

3 How the algorithm works

The algorithm determines the state of a given car in the $(i+1)^{th}$ time interval from the state in the i^{th} using the transition matrix: a random number, weighted by the transition probabilities is chosen from numbers 1-2-3 and this number represents the new state (1 is „move”, 2 is „park without charging” and 3 is „park and charge”). In queuing theory, multiplication with the transition matrix is used and in fact our algorithm works in a similar way. We have seen the structure of the transition matrix in the beginning of Section 2, now we have to describe the „constrains” that overwrite the original transition matrices when it is necessary:

- If a car has started moving, it will not stop until it has moved as much as the by Monte Carlo simulation raffled time. This means, that the fact, that the car is moving, overwrites the transition matrix of that given car: a_{11} becomes 1 and the rest of the elements in the first row will be 0 (for easier programming, the entire first column becomes 1, but it does not matter for the rest of the calculations).
- If the car has moved enough (i.e. the previously by Monte Carlo simulation raffled time is up), it stops. This means, that a_{12} has to turn to 1 and the rest of the first row elements turn to 0 (again, whole columns are changed, but it does not affect the calculations). After stopping, the car returns to the default state, so its transition matrix in the next time step is again the original, corresponding one, that we have seen in Section 2.3.
- If the state of charge (SOC) of the given car's battery decreases below a given limit (30% in our example), the car has to be recharged. This means that the transition matrix for that time interval has to be overwritten again: elements in the third column turn to 1, the rest to 0. However, the car has to get to the charging station, so its SOC decreases further (see Fig. 12).
- If the car is charging, it will not stop this process until it is fully charged. This means, that during charging, the elements in the third column remain 1, the rest is 0.
- If a car has to be recharged, the algorithm checks, whether there are any fast or slow chargers (with this priority order) available. If there is no available charger, the car has to wait, so the second column of the transition matrix is set to 1. This waiting period lasts for only one time step: in the next interval, the algorithm checks again, if there are any free chargers. If yes, the car connects there and begins charging (also, the transition matrix is overwritten again). Waiting is continued in the case no chargers are available.
- If the car has finished charging, it has to leave the station, so we set the car moving by overwriting the first column of the transition matrix with 1. After that, the car stops and returns to a default state, so the transition matrix is restored to the original, corresponding state.
- Although it was not used in our simulation, but a usage limit of the cars can also be set: we can determine from the statistics how many times a car is likely to be used during a day. If it has been used for that many times, further use is not allowed, only on the next day.

4 Simulation results

One of the numerous simulation results we obtained is presented in this section. The input data of the simulation were as follows:

- the number of cars is 100;
- the simulation time range is 3 days;
- simulation time step is 5 minutes ;
- battery capacity of all the vehicles is 22kWh;
- fast chargers recharge a fully depleted battery in 0,5 hours, slow chargers in 4 hours;
- the number of fast chargers is 20 and the number of slow chargers is 10.

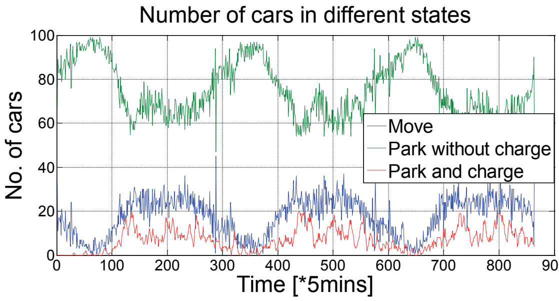

Fig. 9. Number of cars in different states

Figure 9 shows the number of cars in the different states (moving, parking without charging, parking and charging). It can be seen that the behavior of the cars shows daily periodicity and it is also obvious, that the given number of chargers is inadequate: the red curve reaches its maximum at some time instants.

Figure 10 shows the number of cars that have to wait, because there are not enough chargers in the station.

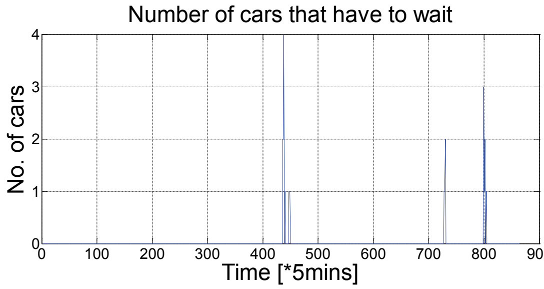

Fig. 10. Number of cars that have to wait

Waiting times also show daily periodicity. The aim of the simulations is to determine the required number of chargers so that nobody has to wait.

Figure 11 shows how the state of charge of cars (in this example we show only 3 cars) varies during the simulation. We can see that after they are fully charged, the cars leave the station, so the SOC decreases in the next time step (Fig. 12):

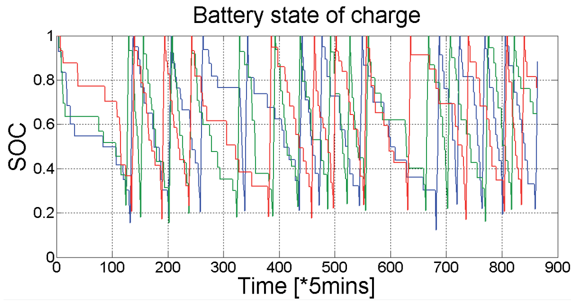

Fig. 11. The change of the state of charge of cars during the simulation

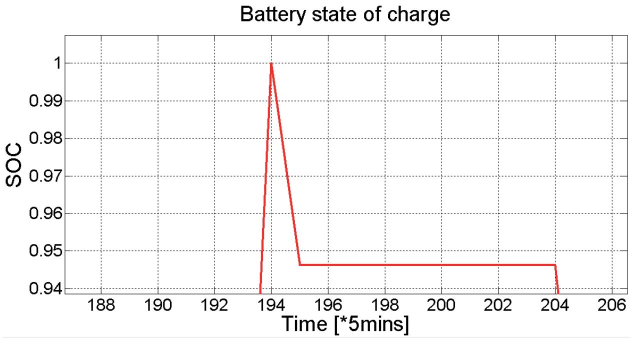

Fig. 12. Cars have to leave the charging station, so their SOC decreases

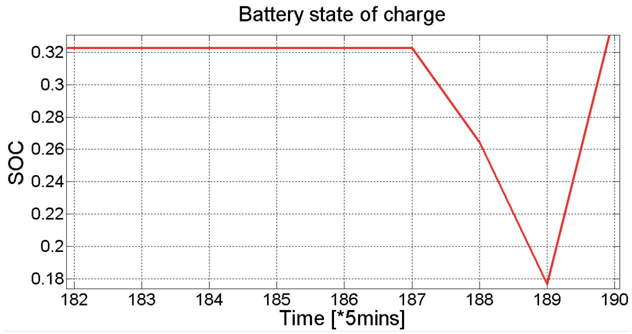

Fig. 13. After reaching the set SOC value, cars have to go back to the station to charge

Figure 13 shows that if a car's SOC falls below the given critical value, it has to go back to the charging station. Meanwhile, its SOC decreases further.

Both the leaving and the approaching of the charging station is supposed to take only one time step: the future work aims to make this assumption more realistic (i.e. it can take not only one time step to get to the station).

5 Further work

The proposed algorithm can be upgraded and we plan to do so in the future:

- Leaving and approaching the station should take more than one time step.
- The optimal number of fast and slow chargers has to be determined: given a fix cost of charger installation and an amount of money, we can set up iterative simulations to determine the optimal number of fast and slow chargers.
- Optimization for more than one charging station: cars can go to multiple locations for charging purpose.
- Further statistical analysis to determine the distribution of inter-arrival time and service time: in the view of these data we can set up queuing models (e.g. M/M/c/N or M/G/1 or G/G/1, depending on the statistical results).

6 Conclusions

Based on traffic and consumption data, we created an algorithm that utilizes transition matrices to determine a car's state transition from one time interval to the next. The model is capable of simulating the motion of a taxi fleet composed of electric cars. We can also investigate whether the number of installed chargers in the charging station for these cars is enough to meet the charging needs. The model can be developed further to be able to determine the optimal ratio of fast and slow chargers in the station taking financial aspects also into consideration.

Acknowledgement

The work reported in the paper is connected to the scientific program of the "Power Converter Optimization Technology in e-Vehicles (e-AutoTech)" project that is supported by the New Széchenyi Plan of the Hungarian Government and financed by the Research and Technology Innovation Fund. (Project ID: KMR_12-1-2012-0188).

The authors would also like to acknowledge the contribution received from Mr. Imre Orlay, technology expert of DSO ÉMÁSZ.

References

1 **Jing Y., Yu Z., Xing X., Guangzhong S.**, *Driving with knowledge from the physical world*. In. 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD'11, New York, NY, USA. (2011). URL: http://research.microsoft.com/pubs/148894/Driving%20with%20Knowledge%20from%20the%20Physical%20World-kdd11.pdf

2 **Jing Y., Yu Z., Chengyang Z., Wenlei X., Xing X., Guangzhong S., Yan H.**, *T-drive: driving directions based on taxi trajectories*. In: Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, GIS ,10, pp. 99-108. New York, NY, USA, (2010). URL: http://research.microsoft.com/pubs/138035/T-Drive-%20Driving%20Directions%20Based%20on%20Taxi%20Traces.pdf?origin=publication_detail

3 **Soares F. J., Peças Lopes J. A., Rocha Almeida P. M., Moreira C. L., Seca L.,** *A stochastic model to simulate electric vehicles motion and quantify the energy required from the grid*. 17th Power Systems Computation Conference. Stockholm, August. 22-26. (2011). URL: http://pscc.ee.ethz.ch/uploads/tx_ethpublications/fp359.pdf

4 City Taxi (Hungarian taxi company) database for 50 taxis, data from Péter Bősz, commercial and marketing director

5 **Xiaowei H., Shi A., Jian W.,** *Exploring urban taxi drivers' activity distribution based on GPS data*. Mathematical Problems in Engineering. 2014. pp. 1-13. (2014). DOI: 10.1155/2014/708482

6 **Marco V., Santi P., Carlos B.,** *Urban mobility study using taxi traces*. TDMA'11, September 18, 2011, Beijing, China. DOI: 10.1145/2030080.2030086

7 **Marco V., Santi P., Carlos B., Patrick O., Nuno F.,** *Exploratory study of urban flow using taxi traces*. 1st International Workshop on Pervasive Urban Applications (PURBA). San Francisco, 12. June 2011. URL: http://wiki.amilab.dei.uc.pt/images/0/04/Veloso_2011_PURBA.pdf

8 **Xi L., Li G., Yongxi G., Yu L.,** *Revealing daily travel patterns and city structure with taxi trip data*. submitted: 24. Oct. 2013.

9 **Jianting Z.,** *Smarter outlier detection and deeper understanding of large-scale taxi trip records: A case study of NYC*. Beijing, 18. September 2011. DOI: 10.1145/2346496.2346521

10 **Peng C., Jin X., Wong K.-C., Shi M., Liò P.** *Collective Human Mobility Pattern from Taxi Trips in Urban Area*. PLoS ONE 7(4). e34487. (2012). DOI: 10.1371/journal.pone.0034487

11 2014 Taxicab fact book. Available from: http://www.nyc.gov/html/tlc/downloads/pdf/2014_taxicab_fact_book.pdf

12 **Yu L., Chaogui K., Song G., Yu X., Yuan T.,** *Understanding intra-urban trip patterns from taxi trajectory data. Journal of Geographical Systems*. 14(4). pp. 463-483. (2012). DOI: 10.1007/s10109-012-0166-z

13 **Xiao L., Jichang Z., Li D., Ke X.,** *Unraveling the origin of exponential law in intra-urban human mobility*. Scientific Reports. 3. 2983. (2013). DOI: 10.1038/srep02983

14 **Xiao L., Xudong Z., Weifeng L., Tongyu Z., Ke X.,** *The scaling of human mobility by taxis is exponential.* Physica A: Statistical Mechanics and its Applications. 391 (5). pp. 2135-2144. (2012). DOI: 10.1016/j.physa.2011.11.035

15 **González M. C., Hidalgo, C. A., Barabási A.-L.,** *Understanding individual human mobility patterns*. Nature. 453. pp. 779-782. (2008). DOI: 10.1038/nature06958

16 **Nicholas J. S.,** *Examination of taxi travel patterns in Arlington County.* MSc thesis, Virginia Tech. URL: http://scholar.lib.vt.edu/theses/available/etd-05112012-115641/unrestricted/Nicholas_JN_T_2012.pdf

17 **Christoforou Z., Milioti C., Perperidou D., Karlafits M. G.,** *Investigation of taxi travel time characteristics.* Transportation Research Board 90th Annual Meeting, 23-27. January, 2011. DOI: 10.4399/97888548486722

18 Take charge - A roadmap to electric New York City taxis, NYC Taxi&Limousine Commission, December (2013). Available from: http://www.nyc.gov/html/tlc/downloads/pdf/electric_taxi_task_force_report_20131231.pdf

19 http://www.mathwave.com/help/easyfit/ html/tools/assist.html

20 http://www.spritmonitor.de/en/