Application of the Theory of Absorbing Markov Processes, for Estimating the Load of Road Sections and Charging Stations, for Electric Car Transport

László Hanka

Óbuda University, Bánki Donát Faculty of Mechanical and Safety Engineering, Institute of Natural Science and Basic Subjects, Népszínház u. 8, H-1081 Budapest, Hungary, hanka.laszlo@bgk.uni-obuda.hu

University of Public Service, Faculty of Military Science and Officer Training, Department of Natural Science hanka.laszlo@uni-nke.hu

Abstract: Due to the short range of pure electric vehicles, considering a long trip, the number and the locations of electric charging stations, especially the distance between consecutive charging stations, is a basic question, because the discharged battery must be recharged quite frequently. Today, the network of electric charging stations is unsatisfactory. During the construction of new and expansion of existing networks, the load of roads and the utilization of existing or imagined charging stations must be taken into account. For the estimation of these features, a perfect mathematical tool, is the theory of absorbing Markov-chains. In this article a possible mathematical model, using Markov-chains, extended by the logit regression model will be presented.

Keywords: electric car; charging station; absorbing Markov process; logistic model

1 Introduction

The long-distance route for an electric vehicle requires lots of possibilities for recharging its battery, because, considering recent technology, the range of an electric car is not satisfactory for long-distance transport. The battery must be recharged frequently, therefore along roads, for example at road junctions lots of new charging stations must be built. The location, the number, the density, and the capacity of these charging stations are basic questions, because on the one hand, there is an expectation from the perspective of the drivers, on the other hand providing charging service for vehicles is a business enterprise for providers. These characteristic values and the utilization of charging stations strongly depend on the load of road sections in a road network.

In this article a mathematical model will be proposed, that can be applied to estimating the above-mentioned characteristic values. The point is, that this problem can be characterized by probabilities since in this process there are lots of random variables, for example, how does a driver select a new road section at a road junction, how many charging stations must be used for reaching the destination, etc.

Two mathematical tools, will be presented for examining the above-mentioned problem. On one hand, the logistic regression/decision-making model will be applied for giving probabilities that are associated with the road sections and on the other hand, it will be proven, that for estimating some fundamental quantity, the theory of absorbing Markov-chains can be used efficiently.

In the litarature several concepts can be found related this question. First, this problem is considered as an optimization problem [1] [2]. For optimization authors propose for example the very efficient genetic algorithm. Second, several other perspectives can be emphasized, like topography of the road network, and the battery lifetime [3]. Modelling the traffic of vehicles by graphs is commonly applied tool in the litarture [4-6]. And finally, the probabilistic point of view is also arises, the Monte-Carlo simulation process can be a possible mathematical tool too [7].

In this paper the author proposes a different procedure for modelling the traffic. Instead of handling it as an optimization problem a probabilistic model is presented, which is based on the very efficient and widely applied theory of Markov-process, that is combined with a probabilistic model. The advantage of this procedure is that the Markov chain is basicly independent to the specific probabilistic model. In the article one probabilistic approach is presented but any other method can be chosen by the user, it doesn't affect the structure of the Markov process. Furthermore this method requires a small amount of calculations, even if the network is great, only powers of matrices and some eigenvectors must be determined.

The rest of the paper is organised as follows: In Section 2, a simple example is examined, illustrating basic problems and challenges, in Section 3, one possible probabilistic model is proposed for choosing road sections, in Sections 4-6, the proposed application of the theory of absorbing Markov-chains is demonstrated, in these sections the construction procedure of Markov matrices and numerical results also can be found. In Section 7, the proposed algorithm is applied for a more general road network which has more terminal points. Finally, in Section 8, conclusions are summarized.

2 A Case Study

As an introduction, the road network, depicted in figure 1. will be studied in detail. As it can be seen, there is one initial point, node $\{1\}$, and one terminal point/destination, which is node $\{6\}$. This graph illustrates the road network by edges, which can be considered as road sections, and nodes that illustrate road junctions. This graph is called transition diagram in the theory of Markov-chains. Charging stations can be at road junctions, or along any road sections.

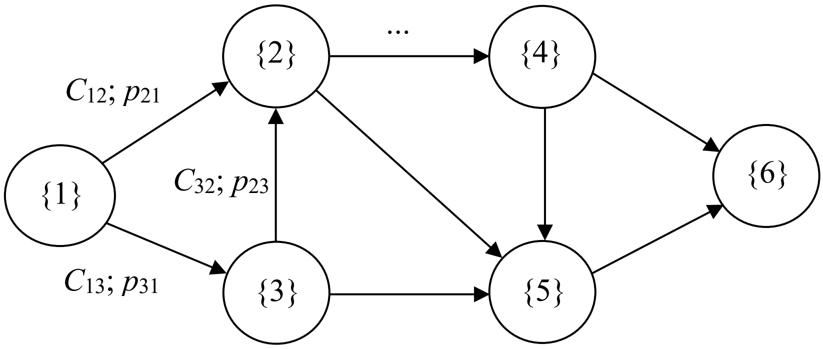

Figure 1
Transition diagram: The graph of a road network, with edges between nodes $\{j\} \to \{i\}$ that are possible routes, the cost of the road sections (C_{ji}) and transition probabilities (p_{ij})

The point is that at some nodes in this example at nodes {2}, {3} and {4}, there is more than one possibility for the further route, the driver has to select one following road section, by some probability, therefore, from the initial point {1} to destination {6} there are several possibilities for choosing a route. The list below gives every possible route:

R_1: $\{1\} \to \{2\} \to \{4\} \to \{6\}$

R_2: $\{1\} \to \{2\} \to \{5\} \to \{6\}$

R_3: $\{1\} \to \{3\} \to \{5\} \to \{6\}$

R_4: $\{1\} \to \{2\} \to \{4\} \to \{5\} \to \{6\}$

R_5: $\{1\} \to \{3\} \to \{2\} \to \{4\} \to \{6\}$

R_6: $\{1\} \to \{3\} \to \{2\} \to \{5\} \to \{6\}$

R_7: $\{1\} \to \{3\} \to \{2\} \to \{4\} \to \{5\} \to \{6\}$

Considering the choice of possible routes, several questions arise. Since a probability model will be presented, the most important and basic problem is the probability that can be attached to one specific route, if at one node in the road network, there is a possibility for choosing a further road section, because it is a fundamental decision-making process for the driver. In other words, if from the node $\{j\}$ there is a direct route to node $\{i\}$, these are consecutive nodes in the network, then the transition probability p_{ij} must be determined by a plausible and reasonable method.

In the following section, it will be demonstrated in detail, that the first reasonable assumption is that this transition probability depends on the "cost" of the road section, which is denoted by C_{ji} költségtől. Several different definitions can be given for this "cost" it depends on expectations, demands, drivers, etc. In this article, we assume that the cost C_{ji} is proportional to the time that is required for the travel along the road $\{j\} \to \{i\}$, but instead of time that quantity could be the fuel consumption too. If the length of the road section is S_{ji}, then the cost of the road section can be defined by the following formula

$$C_{ji} = \frac{S_{ji}}{v} + T \qquad (1)$$

where v is the average velocity of the vehicle considering the whole route, and T is the average charging time of the accumulator. This quantity is naturally taken into account only in nodes {2,3,4,5} and we assume that at node {1} the vehicle starts with a fully charged battery. This cost must be used if the probability of the choice of the road section is determined. The cost is the basis of the definition of the probability distribution which is defined by the logistic decision-making model. This model will be presented in Section 2 in detail.

3 The Logistic Model of Transition Probabilities

The logistic probability model [8] [9] first of all defines the ratio of two probabilities, the quantity that is called "odds", and not the probability itself. This model is widely used in decision-making processes, like the process that is being studied in this article. This model can also be used for giving a probability distribution, the method is as follows.

First of all assume, that there are only two choices. The probability of choosing option one is p, and naturally in this case the probability of option two is $1 - p$. According to the logistic model, the natural logarithm of the ratio of these probabilities "the odds", are approximated by a first-order polynomial by the following formula:

$$\ln\left(\frac{p}{1-p}\right) = a + bx =: \beta(x) \qquad (2)$$

where x can be an "explanatory variable" in a linear regression model. This definition is the reason for the "logistic regression" name. For simplicity, we introduce a notation β for the right-hand side, which is, by assumption, a linear function of the cost, defined in formula (1). Solving this equation for p, the following probability distribution is obtained for the two option case

$$p = p(\beta) = \frac{e^{\beta}}{1+e^{\beta}}; \quad 1-p = \frac{1}{1+e^{\beta}}; \tag{3}$$

Thanks to the shape of the graph of the function defined by (3) the obtained curve is called a sigmoid curve (Figure 2).

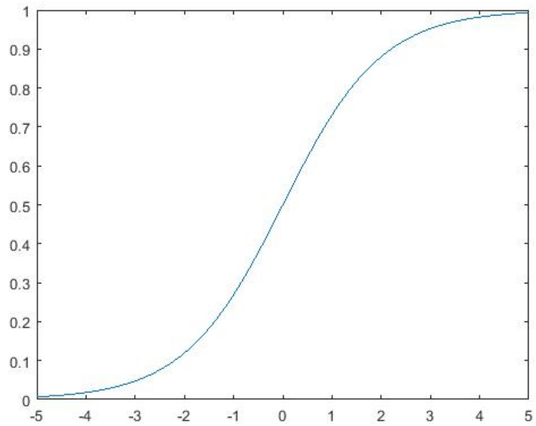

Figure 2
The graph of the function defined by (3), the "sigmoid curve"

The obvious and clear advantage of the application of formula (3) is that independent of the value of β, the codomain of the function is always the interval [0,1] therefore, for any β the value of the function can be considered as a probability.

In practice, in real life generally, there are not only two options for choice but more than two. Consequently, the probability distribution that is defined by formula (3) must be given in a more general form. Assume that there are K options, and one option must be selected. The question is, how can be modified the previously presented simple procedure?

The process is the following [8]. One specific option is chosen as a reference, it can be any, for example, let it be the choice K, and every odds is defined as a ratio of option p_r and p_K, for every possible $r = 1, 2, 3,..., K-1$, similarly to (2):

$$\ln\left(\frac{p_1}{p_K}\right) = \beta_1; \ \ln\left(\frac{p_2}{p_K}\right) = \beta_2; \ .... \ \ln\left(\frac{p_{K-1}}{p_K}\right) = \beta_{K-1}; \tag{4}$$

from which it follows that,

$$p_1 = p_K e^{\beta_1}; \ p_2 = p_K e^{\beta_2}; \ .... \ p_{K-1} = p_K e^{\beta_{K-1}}; \tag{5}$$

and naturally,

$$p_K = 1 - \sum_{r=1}^{K-1} p_r = 1 - \sum_{r=1}^{K-1} p_K e^{\beta_r} = 1 - p_K \sum_{r=1}^{K-1} e^{\beta_r}; \tag{6}$$

from where by some rearrangement, the following probability distribution is obtained,

$$p_K = \frac{1}{1 + \sum_{r=1}^{K-1} e^{\beta_r}}; \quad p_k = \frac{e^{\beta_k}}{1 + \sum_{r=1}^{K-1} e^{\beta_r}}; \; k = 1, 2, ..., K-1. \tag{7}$$

Formula (7) defines the probability distribution of a decision-making process for K options. The only disadvantage of this result is that the distribution in this form seems pretty strange because for option K it provides a different formula as if option K would be specific. Since in general there is no discrimination between options, giving this distribution in a symmetric form would be expedient. Using a linear parameter transformation, the distribution can be modified. Introducing parameters γ_r according to formulas as follows:

$$\beta_r = \gamma_r - \gamma_K; \; r = 1, 2, ..., K-1; \;\; \beta_K = \gamma_K - \gamma_K = 0; \tag{8}$$

instead of parameters β_r the distribution can be given using "new" parameters. On the one hand

$$\begin{aligned} p_k &= \frac{e^{\beta_k}}{1 + \sum_{r=1}^{K-1} e^{\beta_r}} = \frac{e^{\gamma_k - \gamma_K}}{1 + \sum_{r=1}^{K-1} e^{\gamma_r - \gamma_K}} = \frac{e^{-\gamma_K} e^{\gamma_k}}{1 + e^{-\gamma_K} \sum_{r=1}^{K-1} e^{\gamma_r}} = \\ &= \frac{e^{\gamma_k}}{e^{\gamma_K} + \sum_{r=1}^{K-1} e^{\gamma_r}} = \frac{e^{\gamma_k}}{\sum_{r=1}^{K} e^{\gamma_r}}; \; k = 1, 2, ..., K-1. \end{aligned} \tag{9}$$

and on the other hand,

$$\begin{aligned} p_K &= \frac{1}{1 + \sum_{r=1}^{K-1} e^{\beta_r}} = \frac{1}{1 + \sum_{r=1}^{K-1} e^{\gamma_r - \gamma_K}} = \frac{1}{1 + e^{-\gamma_K} \sum_{r=1}^{K-1} e^{\gamma_r}} = \\ &= \frac{e^{\gamma_K}}{e^{\gamma_K} + \sum_{r=1}^{K-1} e^{\gamma_r}} = \frac{e^{\gamma_K}}{\sum_{r=1}^{K} e^{\gamma_r}}; \end{aligned} \tag{10}$$

Consequently, a symmetric form of the probability distribution in a decision-making model is obtained for K choices, according to the logistic probability model. Summarizing results, the distribution is as follows

$$p_k = \frac{e^{\gamma_k}}{\sum_{r=1}^{K} e^{\gamma_r}}; \; k = 1,2,...,K. \tag{11}$$

This distribution will be applied when the traffic of electric vehicles will be examined. Naturally, the basic question is the definition of parameters γ_r. A plausible and reasonable definition can be the following

$$\gamma_r = -C_r; \; r = 1,2,...,K. \tag{12}$$

where C_r is the cost of route r. Using this definition, the probability distribution for the problem, that is being examined in this article, is the following [8] [9]:

$$p_k = \frac{e^{-C_k}}{\sum_{r=1}^{K} e^{-C_r}}; \; k = 1,2,...,K. \tag{13}$$

The obvious benefit of the application of this model is that thanks to the negative exponent, the exponential function is strictly decreasing, the probability of the choice of "expensive" routes is highly, "exponentially" reduced, and conversely, the probability of the choice of "cheap" routes are highlighted, emphasizing the chance of inexpensive road sections. In the following section, this probability model will be used for modeling the traffic using Markov-chains.

4 Modelling the Traffic by Markov-Chains

Considering the road network, depicted in Figure 1, basically two problems must be examined. The first and basic question is the probability of the event that a vehicle is at one specific node (road junction). In mathematics, in the theory of Markov-chains these are "states" (p_i, i = 1,2,...,6). These probabilities that form a discrete distribution are usually summarized in a state vector

$$\boldsymbol{\pi}_n = \left[p_1^{(n)}, p_2^{(n)}, p_3^{(n)}, p_4^{(n)}, p_5^{(n)}, p_6^{(n)} \right]^T; \; n = 0,1,2,... \tag{14}$$

The notation "$\boldsymbol{\pi}$" is commonly used in the theory. The superscript n refers to the fact, that in every junction/node, there is a challenge for the driver, a new road section must be selected, and after every step, the probability distribution changes.

In formula (14) *N* highlights that the given distribution is valid after *n* steps. Finally *T* superscript stands for the transpose of the vector because the vector must be a column vector.

The initial state vector, the initial probability distribution in case, depicted in Figure 1. is obviously the following:

$$\boldsymbol{\pi}_0 = [1,0,0,0,0,0]^T ; \tag{15}$$

The second fundamental question is the probability of the choice of the route $\{j\} \rightarrow \{i\}$ which is p_{ij} if {j} and {i} are consecutive nodes along the rote in this order. Considering the order of subscripts, the explanation of the notation is the following. This probability, p_{ij} is a conditional probability, the probability of the event that the vehicle is at the node $\{i\}$, using the language of the theory of Markov-chains [10-13], it is in the state $\{i\}$ now, assuming that its previous state was the node $\{j\}$. This probability, which is called transition probability, is denoted by the symbol $P(i \mid j)$ in probability theory, for which the notation p_{ij} is only a simplification.

The matrix, that contains every transition probability for the road network, depicted in Figure 1., which is called the "transition probability matrix" [10, 13], is the following:

$$\mathbf{A} = \overset{\text{From}}{\left[\begin{array}{ccccc:c} 0 & 0 & 0 & 0 & 0 & 0 \\ p_{21} & 0 & p_{23} & 0 & 0 & 0 \\ p_{31} & 0 & 0 & 0 & 0 & 0 \\ 0 & p_{42} & 0 & 0 & 0 & 0 \\ 0 & p_{52} & p_{53} & p_{54} & 0 & 0 \\ \hdashline 0 & 0 & 0 & p_{64} & 1 & 1 \end{array}\right]}; \quad \text{to} \tag{16}$$

Considering the structure of the matrix the role of rows and columns must be clarified. Column (*j*) is the former state the "old" node/junction/position of the vehicle, where the vehicle comes from, and row (*i*) is the latter state, the "new" node/junction/position where the vehicle goes to. Every transition probability p_{ij} is obviously zero if there is no immediate connection between junctions $\{j\}$ and $\{i\}$, in other words, there is no edge in the graph between nodes $\{j\}$ and $\{i\}$. Furthermore, it is also obvious that $p_{65} = 1$ since from state {5} there is only one route to state {6}, there is no possibility for choice between nodes {5} and {6}, therefore, the probability of selecting route $\{5\} \rightarrow \{6\}$ is 1. Finally, since the goal is reaching node {6} the vehicle remains in this state by probability $p_{66} = 1$. This state plays a particular role because this state can be accessed from every other state, but from this state, no other nodes can be accessed. If a vehicle reaches this state it remains in that state. This kind of state is called an "absorbing state". It must be emphasized that despite that $p_{65} = 1$ state {5} is not absorbing, because there is a transition from the state $\{5\}$ to state $\{6\}$. According to this observation, the examined stochastic process is an absorbing Markov-chain. The commonly applied partition [11] [12] can be seen in the transition probability matrix (16).

The matrix **A**, defined by the formula (16) is a column-stochastic Markov-matrix, which means that the sum of entries in every column sum up to one

$$\sum_{i=1}^{6} p_{ij} = 1; \; j = 1,2,...,6. \tag{17}$$

According to the theory of Markov-chains [10-13], the state vectors after some steps can be calculated by the following matrix-vector products

$$\boldsymbol{\pi}_1 = \mathbf{A}\boldsymbol{\pi}_0; \; \boldsymbol{\pi}_2 = \mathbf{A}\boldsymbol{\pi}_1 = \mathbf{A}^2\boldsymbol{\pi}_0; \; \boldsymbol{\pi}_3 = \mathbf{A}\boldsymbol{\pi}_2 = \mathbf{A}^3\boldsymbol{\pi}_0;....;\boldsymbol{\pi}_n = \mathbf{A}\boldsymbol{\pi}_{n-1} = \mathbf{A}^n\boldsymbol{\pi}_0;.... \tag{18}$$

Considering (18) it is clear, that state vectors basically depend on transition probabilities. Our goal in the following section, using the logistic probability model presented in section 2., determining transition probabilities so that answers can be given to the questions asked in the introduction.

(In some literature the role of rows and columns are interchanged in the matrix (16). In this case, the state vector (14) must be a row vector, the state transition matrix is row-stochastic and multiplications in (18) are in the opposite order $\boldsymbol{\pi}_n = \boldsymbol{\pi}_{n-1}\mathbf{A}$, etc. We use the above-given definition and consequences exclusively in this article.)

5 Determination of Transition Probabilities

In this section transition probabilities will be calculated [1] [2], which are entries of matrix **A** in (16), using the logistic decision-making model presented in Section 2, for the road network, illustrated in Figure 1. Probabilities will be calculated using the formula (13) for every node, considering reasonable modifications. The most important consequence of Section 2. is that for a route $\{j\} \rightarrow \{i\}$ the transition probability is proportional to $exp(-C_{ij})$, so the only remaining problem is finding the normalizing factor such that (17) would be fulfilled. If the question is the transition probability p_{ij} from the state $\{j\}$ to state $\{i\}$ then the answer will be the following:

$$p_{ij} = e^{-C_{ji}} \frac{\sum_{R \in \Omega_{i6}} e^{-C_{i6}^R}}{\sum_{R \in \Omega_{j6}} e^{-C_{j6}^R}}; \tag{19}$$

where Ω_{i6} denotes the set of every possible route from the state $\{i\}$ to state $\{6\}$.

Simply to say, the transition probability is computed by multiplying the exponential factor by a normalization factor. This normalization factor is simply a ratio of the sum of weight factors considering every route from the node $\{i\}$ to the node $\{6\}$ and the sum of weight factors considering every route from the node $\{j\}$ to node $\{6\}$. It will be clear, that these probabilities form a discrete distribution in every column.

Illustrating the procedure, we determine the distribution in the first column of matrix **A** in detail. Using the formula (19) the following is obtained

$$p_{21} = e^{-C_{12}} \frac{\sum_{R\in\Omega_{26}} e^{-C_{26}^R}}{\sum_{R\in\Omega_{16}} e^{-C_{16}^R}} = \quad (20)$$

$$= e^{-C_{12}} \frac{e^{-C_{246}} + e^{-C_{256}} + e^{-C_{2456}}}{e^{-C_{1246}} + e^{-C_{1256}} + e^{-C_{12456}} + e^{-C_{1356}} + e^{-C_{13246}} + e^{-C_{13256}} + e^{-C_{132456}}};$$

where the cost $C_{j...6}$ in the exponent of the exponential function denotes the total cost of the route $\{j\} \to ... \to \{6\}$ no matter how long it is. To make it clear assume first, that for every route $\{j\} \to \{i\}$ the cost C_{ji} is one unit. In this case, it is clear, that the total cost is proportional to the length of the route. For example for the route $\{2\} \to \{4\} \to \{6\}$ the cost is 2 units, for the route $\{1\} \to \{3\} \to \{2\} \to \{4\} \to \{6\}$ the total cost is 4 units, etc.

For the sake of simpler formulas, the notation $\alpha = exp(-1)$ is introduced. In this case, the formula (20) is equivalent to the following

$$p_{21} = \alpha \frac{\alpha^2 + \alpha^2 + \alpha^3}{\alpha^3 + \alpha^3 + \alpha^4 + \alpha^3 + \alpha^4 + \alpha^4 + \alpha^5} = \frac{2\alpha^3 + \alpha^4}{3\alpha^3 + 3\alpha^4 + \alpha^5}; \quad (21)$$

The other non-zero probability in the first column can be obtained by a similar procedure

$$p_{31} = \alpha \frac{\alpha^3 + \alpha^3 + \alpha^4 + \alpha^2}{\alpha^3 + \alpha^3 + \alpha^4 + \alpha^3 + \alpha^4 + \alpha^4 + \alpha^5} = \frac{\alpha^3 + 2\alpha^4 + \alpha^5}{3\alpha^3 + 3\alpha^4 + \alpha^5}; \quad (22)$$

It is clear that the requirement $p_{21} + p_{31} = 1$ is fulfilled, so a probability distribution is obtained.

Repeating the previous process, every probability distribution can be computed in columns of matrix **A**. The results are as follows

$$p_{42} = \frac{\alpha^2 + \alpha^3}{2\alpha^2 + \alpha^3}; \qquad p_{23} = \frac{2\alpha^3 + \alpha^4}{\alpha^2 + 2\alpha^3 + \alpha^4}; \qquad p_{54} = \frac{\alpha^2}{\alpha + \alpha^2};$$
$$p_{52} = \frac{\alpha^2}{2\alpha^2 + \alpha^3}; \qquad p_{53} = \frac{\alpha^2}{\alpha^2 + 2\alpha^3 + \alpha^4}; \qquad p_{64} = \frac{\alpha}{\alpha + \alpha^2}; \tag{23}$$

It is clear that in every column a discrete distribution is obtained, in other words, the matrix **A** is column-stochastic. Therefore the Markov-matrix **A** has been constructed. Using the above-given definition of α, the Markov-matrix, filled up with numerical data is the following

$$\mathbf{A} = \left[\begin{array}{ccccc|c} 0 & 0 & 0 & 0 & 0 & 0 \\ 0.5586 & 0 & 0.4656 & 0 & 0 & 0 \\ 0.4414 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0.5777 & 0 & 0 & 0 & 0 \\ 0 & 0.4223 & 0.5344 & 0.2689 & 0 & 0 \\ \hline 0 & 0 & 0 & 0.7311 & 1 & 1 \end{array}\right]; \tag{24}$$

The commonly applied and suitable partition of matrix **A** has been preserved. Compare matrix (16) and (24). The benefit of this partition will be clarified in the following section. In Section 5 the application of Markov-matrix will be presented, and it will be demonstrated that the theory can be efficiently applied to the examination of public transport.

6 Properties of Absorbing Markov-Chains

Considering (18) and (24) the probability of any state can be calculated for every integer n. The probability distribution of states can be seen below for n = 1, 2, 3, etc.

$$\boldsymbol{\pi}_1 = \mathbf{A}\boldsymbol{\pi}_0 = \begin{bmatrix} 0 \\ 0.5586 \\ 0.4414 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \; \boldsymbol{\pi}_2 = \mathbf{A}^2\boldsymbol{\pi}_0 = \begin{bmatrix} 0 \\ 0.2055 \\ 0 \\ 0.3227 \\ 0.4718 \\ 0 \end{bmatrix}; \; \boldsymbol{\pi}_3 = \mathbf{A}^3\boldsymbol{\pi}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0.1187 \\ 0.1736 \\ 0.7077 \end{bmatrix};....;etc. \tag{25}$$

A bit more interesting question is what the powers of matrix **A** look like. The powers of **A** can be seen below for every positive integer exponent.

$$\mathbf{A}^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0.2055 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.3227 & 0 & 0.2689 & 0 & 0 & 0 \\ 0.4718 & 0.1554 & 0.1966 & 0 & 0 & 0 \\ 0 & 0.8446 & 0.5344 & 1 & 1 & 1 \end{bmatrix}; \ \mathbf{A}^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.1187 & 0 & 0 & 0 & 0 & 0 \\ 0.1736 & 0 & 0.0723 & 0 & 0 & 0 \\ 0.7077 & 1 & 0.9277 & 1 & 1 & 1 \end{bmatrix};$$
$$\mathbf{A}^4 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0.0319 & 0 & 0 & 0 & 0 & 0 \\ 0.9681 & 1 & 1 & 1 & 1 & 1 \end{bmatrix}; \quad \mathbf{A}^5 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 \end{bmatrix} = \mathbf{A}^n, \ n > 5; \quad (26)$$

The result must be underlined, that for $n \geq 5$ the nth power of **A** stays the same, in other words, it becomes stable, and it won't change if the exponent increases. A "stable matrix" or a "limit matrix" [3-6] is obtained, which is generally defined by the following limit:

$$\lim_{n\to\infty} \boldsymbol{A}^n \quad (27)$$

In this specific illustrating example the limit is reached if $n = 5$. This value depends on the shape and structure of the transition diagram, it can be less and also greater. This stable matrix can be computed in general, the form of the limit matrix can be given by the suitable partition that has already been applied earlier in (16) and (24). In general, the partitioned form of **A** is the following:

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{Q}_{(n-k)\times(n-k)} & \boldsymbol{0}_{(n-k)\times k} \\ \boldsymbol{R}_{k\times(n-k)} & \boldsymbol{I}_{k\times k} \end{bmatrix} \quad (28)$$

in which the meaning of partitions is clear according to (16) és (24), but we emphasize that **0** is the zero matrix, and **I** is the identity matrix, k is the number of absorbing states (in the examined example $k = 1$), and n is the total number of states (in this case $n = 6$). Using this partition calculations that are necessary for powers of **A** and the stable matrix can be carried out easily. For some specific exponents, and for any n integer exponent the power of **A** can be found below:

$$\boldsymbol{A}^2 = \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}^2 & \boldsymbol{0} \\ \boldsymbol{RQ} + \boldsymbol{R} & \boldsymbol{I} \end{bmatrix}$$
$$\boldsymbol{A}^3 = \begin{bmatrix} \boldsymbol{Q}^2 & \boldsymbol{0} \\ \boldsymbol{RQ} + \boldsymbol{R} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}^3 & \boldsymbol{0} \\ \boldsymbol{RQ}^2 + \boldsymbol{RQ} + \boldsymbol{R} & \boldsymbol{I} \end{bmatrix}$$
$$\boldsymbol{A}^4 = \begin{bmatrix} \boldsymbol{Q}^3 & \boldsymbol{0} \\ \boldsymbol{RQ}^2 + \boldsymbol{RQ} + \boldsymbol{R} & \boldsymbol{I} \end{bmatrix} \begin{bmatrix} \boldsymbol{Q} & \boldsymbol{0} \\ \boldsymbol{R} & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}^4 & \boldsymbol{0} \\ \boldsymbol{RQ}^3 + \boldsymbol{RQ}^2 + \boldsymbol{RQ} + \boldsymbol{R} & \boldsymbol{I} \end{bmatrix}$$

…

$$\boldsymbol{A}^n = \begin{bmatrix} \boldsymbol{Q}^n & \boldsymbol{0} \\ \boldsymbol{R}\boldsymbol{Q}^{n-1}+\ldots+\boldsymbol{R}\boldsymbol{Q}^3+\boldsymbol{R}\boldsymbol{Q}^2+\boldsymbol{R}\boldsymbol{Q}+\boldsymbol{R} & \boldsymbol{I} \end{bmatrix} = \begin{bmatrix} \boldsymbol{Q}^n & \boldsymbol{0} \\ \boldsymbol{R}(\boldsymbol{Q}^{n-1}+\ldots+\boldsymbol{Q}^3+\boldsymbol{Q}^2+\boldsymbol{Q}+\boldsymbol{I}) & \boldsymbol{I} \end{bmatrix} \quad (29)$$

The stable matrix is obtained if $n \to \infty$. Since matrix **Q** is one partition of the Markov-matrix, every entry is less than 1 (see 24), therefore the sequence of powers of matrix **Q** tends to the zero matrix, according to the properties of the geometric sequence. The sum in the lower left corner, using again the properties of the geometric series, can be given in the following simple form:

$$\boldsymbol{I}+\boldsymbol{Q}+\boldsymbol{Q}^2+\boldsymbol{Q}^3+\ldots+\boldsymbol{Q}^{n-1}+\ldots=(\boldsymbol{I}-\boldsymbol{Q})^{-1} \quad (30)$$

which sum exists if in some positive integer power of **Q,** every entry is less than 1 [10, 11, 13]. In this example, it is valid for **Q** itself. Summarizing observations, the stable matrix can be given in a general form:

$$\lim_{n\to\infty} \boldsymbol{A}^n = \begin{bmatrix} \boldsymbol{0} & \boldsymbol{0} \\ \boldsymbol{R}(\boldsymbol{I}-\boldsymbol{Q})^{-1} & \boldsymbol{I} \end{bmatrix} \quad (31)$$

Using this limit matrix, on the basis of the theory of Markov-chains, the traffic, the public transport can be characterized numerically. The load of road sections and the utility of charging stations can be estimated by probabilities.

The justification of this statement is, that some parts and partitions of the limit matrix have fundamental meaning [10-13].

1) The entry in the ith row and jth column of the matrix $(\mathbf{I} - \mathbf{Q})^{-1}$ is the expected value of the random variable that the node $\{i\}$ is reached exactly from the node $\{j\}$, in other words, the average number of vehicles along the road section $\{j\} \to \{i\}$.
2) The sum of columns of the matrix $(\mathbf{I} - \mathbf{Q})^{-1}$ which is the row vector $\mathbf{1}^T(\mathbf{I} - \mathbf{Q})^{-1}$ has also fundamental meaning. The jth coordinate of this vector is the average number of steps, in other words, the expected value of steps from the state $\{j\}$ to the absorbing state, the average number of road sections for a vehicle from the node $\{j\}$ to the destination.
3) The (i, j) entry in the matrix $\mathbf{R}(\mathbf{I} - \mathbf{Q})^{-1}$ is the probability of the event, that the absorbing state $\{i\}$ is reached through the state $\{j\}$. In other words, the probability of the event that the vehicle reaches the destination $\{i\}$ and before it attains node $\{j\}$, for example, charges its battery at that road junction.

For the road network, depicted in Figure 1, and for the transition probability matrix that is given by (24) these matrices are as follows:

$$(\boldsymbol{I}-\boldsymbol{Q})^{-1}=\begin{bmatrix}1 & 0 & 0 & 0 & 0\\ 0.7641 & 1 & 0.4656 & 0 & 0\\ 0.4414 & 0 & 1 & 0 & 0\\ 0.4414 & 0.5777 & 0.2689 & 1 & 0\\ 0.6773 & 0.5777 & 0.8034 & 0.2689 & 1\end{bmatrix} \quad (32)$$

In this matrix for example the entry (2,3) is 0.4656, which means that from the node {3} to node {2} the expected number of selecting the road section is 0.4656, etc. These numbers characterize the whole road network, and the load of various road sections can be compared. The indirect consequence of these data is that the utilization of charging stations at the nodes can be concluded. Summing columns of the previous matrix, the following row vector is obtained:

$$\mathbf{1}^T(\boldsymbol{I}-\boldsymbol{Q})^{-1}=[3.3242 \quad 2.1554 \quad 2.5379 \quad 1.2689 \quad 1.0000] \quad (33)$$

As we described earlier, the meaning of these vector components is the average number of steps from a specific node to the destination. For example, it is obviously 1 at the node {5} because from this node a vehicle can only go to node {6} and the "distance" is only one step. But the from the node {2} the average number of steps to node {6} is 2.1554.

Finally the result:

$$\boldsymbol{R}(\boldsymbol{I}-\boldsymbol{Q})^{-1}=[1.0000 \quad 1.0000 \quad 1.0000 \quad 1.0000 \quad 1.0000] \quad (34)$$

is trivial in this case, because there is only one destination/target in this network, therefore independent of nodes, the vehicle reaches node {6} by probability 1.

7 A Case Study for Two Absorbing States

Illustrating the proposed method, we present one more example, which is a bit more general than the previously studied example. Let the road network be the graph depicted in Figure 3:

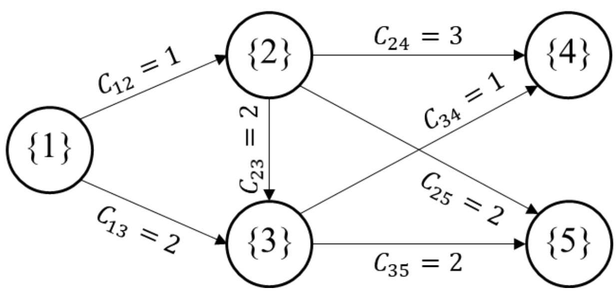

Figure 3

Transition diagram: The graph of a road network, with one initial point, two absorbing states and different raod costs

The obvious difference between this second illustrating example and the previous one, on the one hand is that in this road network there are two terminal nodes, in other word two absorbing states {4} and {5}, and on the other hand, road costs are not units and are not the same. In this section the application of the proposed mathematical tools will be presented for this more general case.

The column stochastic Markov-matrix for this case, is as follows and the necessary partition is also marked:

$$\mathbf{A}=\begin{bmatrix}\mathbf{Q}_{3\times3} & \mathbf{0}_{3\times2}\\ \mathbf{R}_{2\times3} & \mathbf{I}_{2\times2}\end{bmatrix}=\left[\begin{array}{ccc:cc} 0 & 0 & 0 & 0 & 0\\ p_{21} & 0 & 0 & 0 & 0\\ p_{31} & p_{32} & 0 & 0 & 0\\ \hdashline 0 & p_{42} & p_{43} & 1 & 0\\ 0 & p_{52} & p_{53} & 0 & 1\end{array}\right] \tag{35}$$

Due to the fact that in this network there are 5 nodes, the transition probability matrix is a 5×5 matrix. Furthermore, since in this network there are two absorbing states, $k = 2$, in the partition of the matrix the identity matrix in the lower right corner is a 2×2 matrix, therefore matrix **Q** in the upper left corner is a 3×3 matrix.

The probability distributions in the columns of **A** can be given by following the same logic. Using the same α notation for exp(–1), columns of **A** are as follows:

$$\begin{cases} p_{21}=\dfrac{\alpha^5+2\alpha^4+\alpha^3}{\alpha^5+3\alpha^4+2\alpha^3};\\ p_{31}=\dfrac{\alpha^4+\alpha^3}{\alpha^5+3\alpha^4+2\alpha^3};\end{cases}\quad \begin{cases} p_{32}=\dfrac{\alpha^4+\alpha^3}{\alpha^4+2\alpha^3+\alpha^2};\\ p_{42}=\dfrac{\alpha^3}{\alpha^4+2\alpha^3+\alpha^2};\\ p_{52}=\dfrac{\alpha^2}{\alpha^4+2\alpha^3+\alpha^2};\end{cases}\quad \begin{cases} p_{43}=\dfrac{\alpha}{\alpha+\alpha^2};\\ p_{53}=\dfrac{\alpha^2}{\alpha+\alpha^2};\end{cases} \tag{36}$$

since for example,

$$p_{21}=e^{-C_{12}}\frac{\sum_{R\in\Omega_{2v}} e^{-C_{2v}^R}}{\sum_{R\in\Omega_{1v}} e^{-C_{1v}^R}}= \tag{37}$$
$$=e^{-C_{12}}\frac{e^{-C_{24}}+e^{-C_{234}}+e^{-C_{25}}+e^{-C_{235}}}{e^{-C_{124}}+e^{-C_{125}}+e^{-C_{1234}}+e^{-C_{1235}}+e^{-C_{134}}+e^{-C_{135}}}=\frac{\alpha^5+2\alpha^4+\alpha^3}{\alpha^5+3\alpha^4+2\alpha^3};$$

because $C_{1235} = 1 + 2 + 2 = 5$, etc. It is clear that these columns sum up to one, so the matrix is indeed column-stochastic. Markov-matrix filled up with numerical values can be seen below

$$\boldsymbol{A} = \left[\begin{array}{ccc|cc} 0 & 0 & 0 & 0 & 0 \\ 0.5777 & 0 & 0 & 0 & 0 \\ 0.4223 & 0.2689 & 0 & 0 & 0 \\ 0 & 0.1966 & 0.7311 & 1.0000 & 0 \\ \hline 0 & 0.5344 & 0.2689 & 0 & 1.0000 \end{array}\right] \quad (38)$$

State vectors,

$$\boldsymbol{\pi}_0 = \begin{bmatrix} 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix}; \quad \boldsymbol{\pi}_1 = \boldsymbol{A}\boldsymbol{\pi}_0 = \begin{bmatrix} 0 \\ 0.5777 \\ 0.4223 \\ 0 \\ 0 \end{bmatrix}; \quad \boldsymbol{\pi}_2 = \boldsymbol{A}\boldsymbol{\pi}_1 = \boldsymbol{A}^2\boldsymbol{\pi}_0 = \begin{bmatrix} 0 \\ 0 \\ 0.1554 \\ 0.4223 \\ 0.4223 \end{bmatrix};$$

$$\boldsymbol{\pi}_3 = \boldsymbol{A}\boldsymbol{\pi}_2 = \boldsymbol{A}^3\boldsymbol{\pi}_0 = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0{,}5359 \\ 0{,}4641 \end{bmatrix}; etc. \quad (39)$$

and powers of **A** for different exponents are illustrated below,

$$\boldsymbol{A} = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0.5777 & 0 & 0 & 0 & 0 \\ 0.4223 & 0.2689 & 0 & 0 & 0 \\ 0 & 0.1966 & 0.7311 & 1 & 0 \\ 0 & 0.5344 & 0.2689 & 0 & 1 \end{bmatrix};$$

$$\boldsymbol{A}^2 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.1544 & 0 & 0 & 0 & 0 \\ 0.4223 & 0.3932 & 0.7311 & 1 & 0 \\ 0.4223 & 0.6068 & 0.2689 & 0 & 1 \end{bmatrix};$$

$$\boldsymbol{A}^3 = \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 \\ 0.5359 & 0.3932 & 0.7311 & 1 & 0 \\ 0.4641 & 0.6068 & 0.2689 & 0 & 1 \end{bmatrix} = \boldsymbol{A}^4 = \boldsymbol{A}^5 \ldots = \lim_{n\to\infty} \boldsymbol{A}^n \quad (40)$$

The first observation is that the stable matrix is reached for $n = 3$. Finally, the partitions of the stable matrix must be examined.

1) The expected number of "road section choices" can be seen in the matrix $(\mathbf{I} - \mathbf{Q})^{-1}$:

$$(\boldsymbol{I} - \boldsymbol{Q})^{-1} = \begin{bmatrix} 1 & 0 & 0 \\ 0.5777 & 1 & 0 \\ 0.5777 & 0.2689 & 1 \end{bmatrix} \quad (41)$$

2) The average number of steps to absorbing states are provided by the row vector $\mathbf{1}^T(\mathbf{I} - \mathbf{Q})^{-1}$:

$$\mathbf{1}^T(\boldsymbol{I} - \boldsymbol{Q})^{-1} = [2.1554 \quad 1.2689 \quad 1] \quad (42)$$

3) The probability of the event that one absorbing state is reached from a specific node, is given by the matrix $\mathbf{R}(\mathbf{I}-\mathbf{Q})^{-1}$:

$$\boldsymbol{R}(\boldsymbol{I}-\boldsymbol{Q})^{-1}=\begin{bmatrix}0.5359 & 0.3932 & 0.7311\\ 0.4641 & 0.6068 & 0.2689\end{bmatrix} \tag{43}$$

In this case, a great and basic difference can be realized between the first and the second example. In this matrix there are two rows, and the reason is clear, there are two absorbing states, so from any state both absorbing states can be reached, there are two options for any intermediate state, the probability for reaching one or the other absorbing state can be found in the matrix. It must be emphasised, that in all columns, probabilities form a distribution!

Conclusions

In this work, a method has been presented on how the theory of absorbing Markov-chains can be used for modeling the traffic of vehicles along a given road network. The basis of this mathematical model is the logistic regression model, which is a sophisticated method for giving a probability of any decision in a decision-making process. The possible application of this logistic model is reasonable, because along a road network in every road junction, there is an expectation against the driver for choosing a following road section, so in every node, there is a decision-making process.

It has been shown that the combination of these two theories can be used for estimating the load of road sections, and the utilization of charging stations, because these quantities can be characterized by data that can be found in the limit of the Markov-matrix. The process was illustrated by two examples. The first example, contained only one initial point and one destination with the same and unit road costs. In the second example, there is more than one destination in the road network, the road costs are different and not units.

Further study will possibly include modifying the number of initial points and for a greater network/graph, an efficient and simpler method must be developed for giving the Markov-matrix.

Calculations have been performed using the software MATLAB.

References

[1] Lech-Nikonczuk: The method of route optimization of electric vehicle, Procedia Computer Science Volume 207, 2022, pp. 4454-4462

[2] Luque at al.: Optimization of the powertrain of electric vehicles for a given route, Transportation Research Procedia Volume 58, 2021, pp. 246-253

[3] Perger-Auer: Energy efficient route planning for electric vehicles with special consideration of the topography and battery lifetime, Energy Efficiency volume 13, 2020, pp. 1705-1726

[4] Pokorádi: Graph model-based analysis of technical systems, IOP Conf. Ser.: Mater. Sci. Eng. 393 012007, 2018

[5] Péter-Szabó: A new network model for the analysis of air traffic network, Periodica Polytechnica Transportation Engineering, 40(1), pp. 39-44

[6] Péter-Szauter-Rozsás-Lakatos: Integrated application of network traffic and intelligent driver models in the test laboratory analysis of autonomous vehicles and electric vehicles, International Journal of Heavy Vehicle Systems Vol. 27, Nos. 1-2, 2020, pp. 227-245

[7] Pokorádi: Monte-Carlo Simulation-Based Accessibility Analysis of Temporal Systems, Symmetry 2022, *14*(5), p. 983

[8] Wackerly: Mathematical Statistics with Applications, Brooks/Cole, USA, 2008

[9] Bijma: An Introduction to Mathematical Statistics, Amsterdam University Press, 2017

[10] Karlin - Taylor: An introduction to stochastic modeling, Academic Press, London, 1998

[11] Ross: Stochastic processes, John Wiley and sons, USA, 1996

[12] Knill: Probability and Stochastic Processes with Applications, Overseas Press, India, 2006

[13] Hillier - Lieberman: Introduction to Operations Research, McGraw-Hill, USA, 2001