P system based model of passenger flow in public transportation systems: a case study of Prague Metro⋆

Zbyněk Janoška and Jiří Dvorský

Department of Geoinformatics, Faculty of Science, Palacký University Olomouc,
17. listopadu 50, 771 46, Olomouc, Czech Republic
zbynek.janoska@centrum.cz, jiri.dvorsky@upol.cz

Abstract. P systems are branch of bio-inspired computing, which takes inspiration from the structure and functioning of a living cell. Current research of P systems focuses mainly on their computational power, applications include biochemical and ecological modeling. In this paper we propose model of passenger flow in metro based on P systems. This model focuses on detailed description of passenger behavior, while retaining simple and robust in description of vehicle flow. Formal description of model is given and simulations using Prague Metro as an example with real traffic flow data from 2008 are presented. Some open problems are discussed and further directions of research are suggested.

Keywords: P systems, passenger flow, Prague Metro, transportation simulation

1   Introduction

P systems were introduced by Păun [8] as s computing model mimicking structure and behavior of a living cell. Păun based theoretical model of P systems on observation, that processes, taking place inside a living cell can be understood as a computation. This theoretical approach to computing is called membrane computing, while formalized mathematical models from this family are called P systems.

P systems consist of three essential parts - membrane structure, multisets of objects and rules, which process them. The structure of P systems is directly inspired by nature, where membranes delimit different compartments of a cell and their purpose is both protection of content of these compartments and also serve as a transportation channel. Different sets of objects are present in these compartments - they are called multisets, because great numbers of each element are assumed inside a living cell. These objects are processed by a set of rules,

---

which take the form similar to a chemical reaction: $a \rightarrow b$, where a and b are multisets of objects. When a rule is applied, all objects on the left side of a rule are removed and objects on the right side of a rule are introduced into a system. Application of rules is exhaustive, maximally parallel and non-deterministic.

For detailed description of P systems please consult [3], complete biography on P systems is available from [13].

Most research in the field of P systems focuses on computational power (e.g.[9, 11, 12]), applications are restricted mostly to biochemical [1] or ecological [2] topics. In our research, we focus on application of P systems to vehicular traffic flow phenomena [4]. In this paper we propose model for passenger flow simulation in public transportation networks. The model focuses on detailed description of passenger behavior both inside and outside the vehicle; vehicle movement is described briefly and not modeled in detail.

Aim of the model is to accurately estimate numbers of passengers, who use the transportation system at given time. At every time step, numbers of passengers waiting at stations and numbers of passengers in trains are available. This information can be used to examine the performance of the transportation system and the occupancy of trains, for example in case, when a great number of passengers is introduced into a system, or when the schedule of trains is changed. Model can also be used in evacuation studies, when numbers of people in different parts of the transportation system are needed. On the other hand, this model is not suitable for examination of changes of behavior of passengers. It might be of interest to know, how long are passengers willing to wait for a delayed train, or if they rather wait for next train in case, that the current train is almost fully occupied. This kind of research questions requires agent-based modelling, where passengers will make decisions. P systems do not allow decision making – behavior of passengers is described using predefined set of rules, which do not change during the computation, and therefore passengers can not react to ongoing changes within the system.

Performance of a model is shown on an example of Prague Metro, which is simple network of three intersecting lines. Traffic flow data from 2008 are used in simulations.

The paper is structured as follows: in section 2, a formal description of model is given, in section 3, performance of model is examined, section 4 focuses on some problems, which were encountered during the analysis and future directions of research are suggested. Section 5 contains short conclusion.

2   Model description

There exist several levels of traffic modeling, ranging from macro-models, describing traffic flow only in terms of populations of object, to micro-models, where every individual object in the system is examined in detail [5]. For purposes of passenger flow simulation, mezo-scale models are recommended [10]. In mezo-models, some parts of system are described in detail, while description of other parts is rather laconic. Such a model is suitable to focusing on certain part

of traffic flow phenomena, while retaining robust and computationally efficient. In similar manner, proposed model is designed to capture detailed behavior of passengers, while flow of vehicles is brief and simplistic.

Real world system consists of several components, which must be represented in terms of P systems. Metro stations are considered as membranes. Network of stations is represented as a graph, similar to neural P systems [6]. Metro trains are considered membranes, but unlike classical membranes in P systems, they are mobile - their position in the system changes as the system evolves. This evolution is handled by a set of rules [7]. Finally, the passengers are represented as objects. Their behavior follows set of rules, which change according to the position of passengers (inside train or at station).

We believe this representation is very expressive – it is easy to see metro stations as membranes, which are entered and left by passengers – objects. Vehicles serve as membranes too – their function, same as function of living membranes – is to protect its content and serve as a mean of selective transportation (not all objects can enter the membrane and membrane can be entered only on specific occasions). This expressiveness (compared to description using i.e. set of differential equations) together with massive parallelism of model – are two main advantages of P systems for transportation modeling.

Formally, P systems for passenger flow simulation in public transportation networks are following construct:

$$\Pi = (O, l, syn, R), \tag{1}$$

where:

- $O = \{people_{\{a,\ldots,b\}}, empty\} | a, b \in \{l - \{train_1, \ldots, train_n\}\}\}$ is a set of objects, where *people* represents passenger and *empty* represents empty seat inside a train. To each passenger *people*, a sequence $\{a, \ldots, b\}$ is assigned. This sequence is an ordered list of all stations, which passenger visits on his route from station a (current station) to station b (end station of an individual route).
- $l = \{1, \ldots, k, tram_1, \ldots, tram_n\}$ is a set of membranes with k metro stations and n trains.
- $syn \subseteq \{(i, j, t) | i, j \in \{l - \{train_1, \ldots, train_n\}\}, i \neq j, t \in \mathbb{N}\}$ is a set of synapses, representing topology of a network. Each synapse consists of two labels of metro stations i, j and time t necessary to transport train from station i to station j.
- R is a set of rules, which describe the behavior of both membranes and objects inside them. Rules are assigned to membranes, therefore different membranes have different sets of rules and same object in two different membranes can be evolved using different sets of rules. In next section, following notation will be used: train going to station k will be denoted by $[_k \ ]_k$, hence k is label of next, not current station. Each train can be in two states - stopped or moving. Membrane polarization is used to distinguish between

the two, therefore moving train to station k is denoted as $[^+_k]^+_k$. Metro station with label m will be denoted as $(_m)_m$. Following set of rules is used to describe the evolution of the system.

1. $people_{\{a,b,...,x\}} [_a empty]^-_a \rightarrow [_a people_{\{b,...,x\}}]^-_a$ is rule describing passenger entering a train. Passenger, whose next stop is a enters a train going to station a, if there is an empty seat ($empty$) and the train is stopped (negative polarization). Once inside the train, passengers next station changes to b.

2. $[_a people_{\{a,b,...,x\}}]^-_a \rightarrow [_a people_{\{b,...,x\}}]^-_a$ is rule describing passenger staying inside a train. Passenger, whose next stop is a and who is already in a train going to a, stays inside and his next stop changes to b.

3. $[_a people_{\text{NULL}}]^-_a \rightarrow [_a empty]^-_a$ describes situation, where passenger inside a train has no next station, hence is in his final destination and leaves the system. An $empty$ object is created inside a train.

4. $[_a people_{\{b,c,...,x\}}]^-_a \rightarrow [_a empty]^-_a people_{\{b,c,...,x\}}$ describes passenger leaving the train at transfer station. If passenger, whose next stop is b, is inside train going to a, he leaves the train and $empty$ seat appears inside a train. The passenger stays at the current station.

5. $(_i [_j]^+_j)_i \xrightarrow{t} (_j [_k]^-_k)_j$ is rule describing movement of trains inside a network. Moving train in station i, whose next station is j, is moved to station j, its next station is changed to k and the train stops.

6. $(_i [_j]^-_j)_i \rightarrow (_i [_j]^+_j)_i$ changes train from stopped to moving state.

7. $(_i [_{\text{NULL}}]^-_{\text{NULL}})_i \rightarrow (_i)_i$ is rule describing situation, where train reaches its final destination (i.e. does not have next stop). Such a train is removed from the system.

8. $(_i)_i \rightarrow (_i [_a]^-_a)_i$ is rule describing generation of trains in start stations – train going to station a is created in station i. The trains is stopped, therefore passengers can enter the train immediately.

9. $(_i)_i \rightarrow (_i people_{\{a,b,...,x\}})_i$ describes arrival of people to the station i. For each passenger, who arrives at the station, a sequence of stations to visit $a, b, \ldots, x$ is generated.

3   Results

Application of model is demonstrated on example of Prague Metro. This transportation system consists of three lines, labeled A (Dejvická – Depo Hostivař), B (Zličín – Černý Most) and C (Letňany – Háje), which intersect at three stations (A–B - Můstek, A–C - Muzeum, B–C - Florenc). Prague Metro consists of 53 stations in total and approximately 1.5 millions of people are transported every day. Trains are in service from approximately 4:40 a.m. to 24 a.m. (midnight) and their frequency changes in time. In five-year periods, survey of passenger occupancy is performed with last survey taking place at 2008. Prague Public Transit Company provided detailed data about transit intensity, which were used in this case study. We selected Prague Metro for its importance as a mean of transportation and also relative simplicity of the system, which allows well-designed simulations.
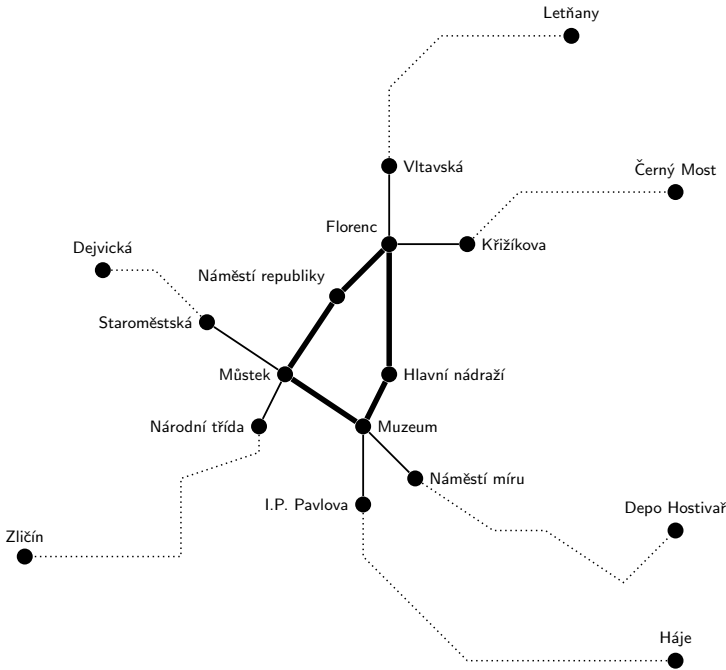
Fig. 1. Prague metro – schematic map

Current (December 2012) schedule of trains was used as basis for generation of trains at start stations. Main problem of public transport modeling is estimation of number of passengers traveling between each pair of stations. Fortunately, this information was provided by Prague Public Transit Company in form of so called Origin-Destination Matrix. Intensity of transport varies during day, which leads to problem with estimation of this intensity. Due to unknown trend of intensities during a day, the quantity was estimated using train schedule. It was assumed, that frequency of train arrivals at given station corresponds with amount of people transported. Kernel density (Epanechnikov kernel, bandwidth=50 minutes) of train frequency was estimated and used as a basis to calculate numbers of passengers entering the system at given time. Figure 2 shows estimated intensity. Values on y axis represent estimated intensity of process, generating "dots" – times of arrival of trains on x axis. We assume, that the process generating times of arrival of trains is in fact the transportation demand of passengers and therefore estimated intensity of this process can be used to calculate the numbers of passengers using the system at given time. Maximal capacity of train was set to 1363, which is occupancy of train 81-71, a standard train in Prague Metro [14]. P systems are discrete in time and one minute was set as a time unit. At every station, the number of passengers N corresponding to estimated intensity function from figure 2 was calculated each minute. Concrete number of passengers was generated from Poisson distribution with $\lambda = N$.

Currently, there are no simulators of P systems, which enable usage of rules in form, which was presented in chapter 2. Set of scripts in python was developed to perform the simulation.
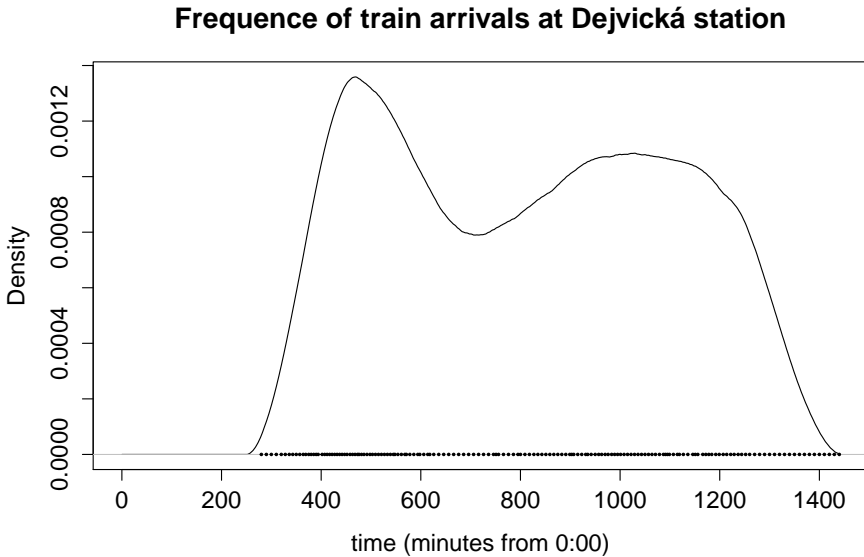
Frequence of train arrivals at Dejvická station

Fig. 2. Estimated train arrival frequency

Four interesting variants of stations were identified in the system and further examined:

1. Final stations – Dejvická, Depo Hostivař, Háje, Letňany, Zličín, Černý Most
2. Transfer stations – Můstek, Muzeum, Florenc
3. Stations between two transfer station – Náměstí Republiky and Hlavní Nádraži
4. Station next to one transfer stations – Vltavská, Křižíkova, Náměstí Míru, I.P.Pavlova, Národní Třída, Staroměstská

3.1   Final stations

Final stations are unique, because only one direction of the train line is served. Simulation showed periodic behavior, where after train departure, no passengers ale left at station (Figure 3). This means, that the transportation demand is fully served.
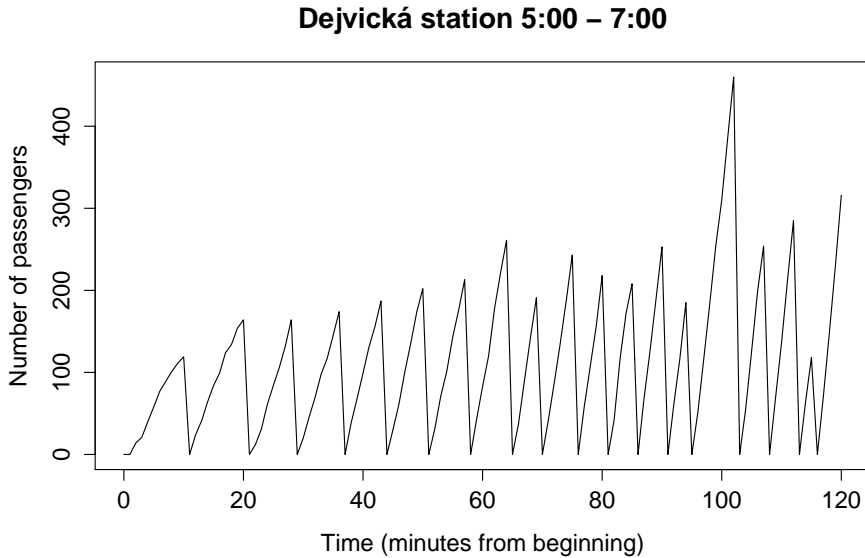
Dejvická station 5:00 – 7:00

Fig. 3. Number of passengers waiting at the station during 2 hours simulation period

3.2 Transfer stations

There are three transfer stations, where always two lines intersect. These stations are very frequently used and numbers of passengers waiting for train show also cyclic, but more irregular pattern. Moreover, at the beginning of the study period, there is elevated number of passengers waiting for the train (Figure 4). It seems, that the system is not able to handle the demand of passengers for short period of time, but later, the frequency of trains increases and the "wave" of waiting passengers is dissolved. We attribute this behavior not to design of the model, neither we think it represents real behavior of the system, but assume it is caused by incorrectly estimated passenger flow intensity. We will discuss this problem later in chapter 4.

3.3 Stations between two transfer station

Both Náměstí Republiky and Hlavní Nádraží stations show similar, but even more evident pattern as transfer stations. The periodic behavior is more regular and "peak" at the beginning of the study period is more distinctive. Due to high frequency of trains, the numbers of passengers waiting are lower than at most of the other stations (Figure 5).
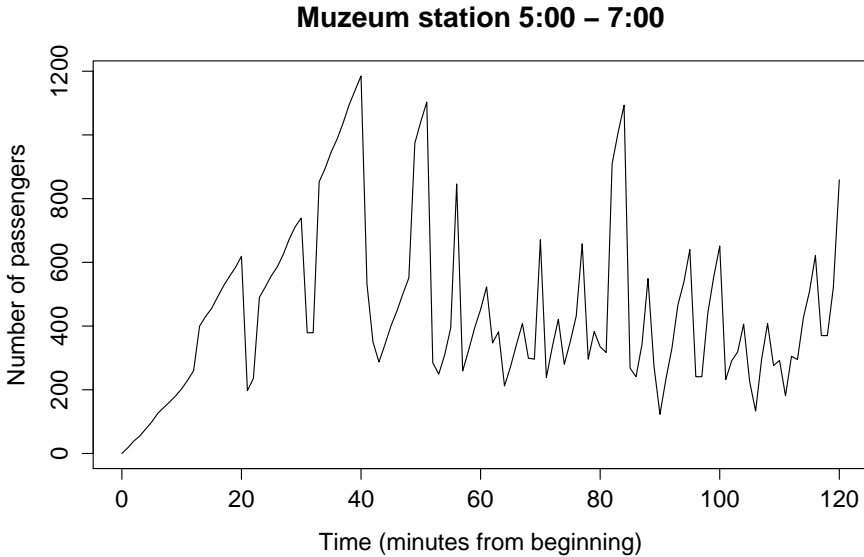
Muzeum station 5:00 – 7:00

Fig. 4. Number of passengers waiting at the station during 2 hours simulation period

3.4   Stations next to one transfer stations

A rather regular periodic pattern with two peaks can be observed at stations next to transfer stations (Figure 6). Rapid rise in number of waiting passengers was not observed at the beginning of the study period, also the number of passengers returns to values close to zero, which indicates, that the schedule is appropriate to the transport demand. The position of peaks (irregular or regularly spaced) is caused by train schedule and is not caused by stations being immediately after transfer station.

4   Discussion

Possible incorrect performance of the model can result from two types of errors: errors in design of the model and incorrect input values.

Input values of the model are traffic demand and train schedule. Numbers of passengers traveling between all pairs of stations were derived from transportation survey of passenger occupancy and should accurately describe the system. However, only sum of all passengers was available for every station, dynamics of the demand during the day is unknown.

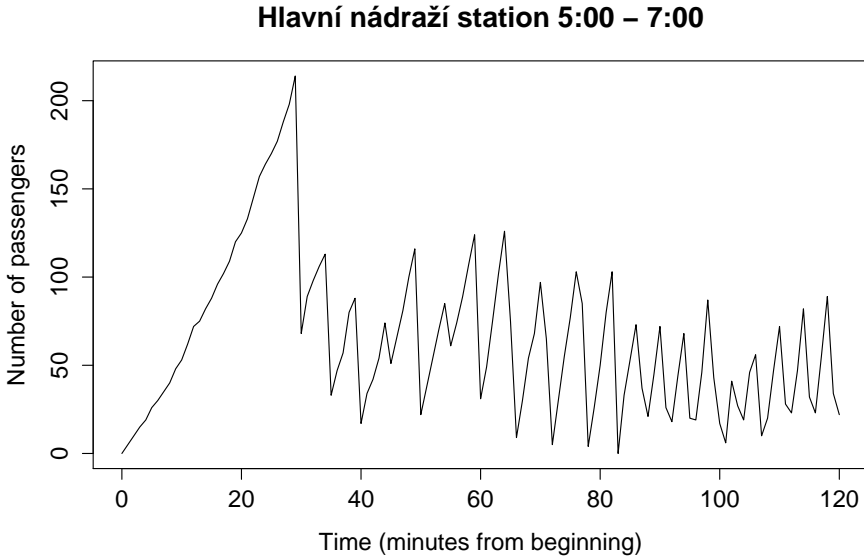Issue with elevated numbers of passengers at the beginning of the study period

Hlavní nádraží station 5:00 – 7:00

Fig. 5. Number of passengers waiting at the station during 2 hours simulation period

we attribute to incorrect estimation of passenger demand quantity. Different estimates (using different kernels and bandwidths) were examined, however none of them led to results, which both copied global trend and did not produce elevated numbers at the beginning. Correct estimation of traffic demand is crucial step and our next research will be pointed at this direction.
The movement of trains is ruled deterministically using real schedule, obtained from world wide web. In reality, this schedule will be rarely kept, therefore and arbitrary constant can be added to travel time between two consequent stations to account for train delays. This time constant should be preferably generated randomly from known distribution of time delays.

Errors in the design of the model can be represented by ommiting important rules in description of the model. Presented model focuses on more detailed description of passenger behavior, while remaining brief in description of vehicle flow.
Passengers, who exit the train are immediately removed from the system, while in real system, they remain at the station for some time. While not important for examining the capacity or occupancy of traffic system, this might be an issue in i.e. evacuation studies. Extending system by new object - passenger, who no longer participates in transportation, and extending current rules would incorporate this extension, while keeping the model robust and simple.
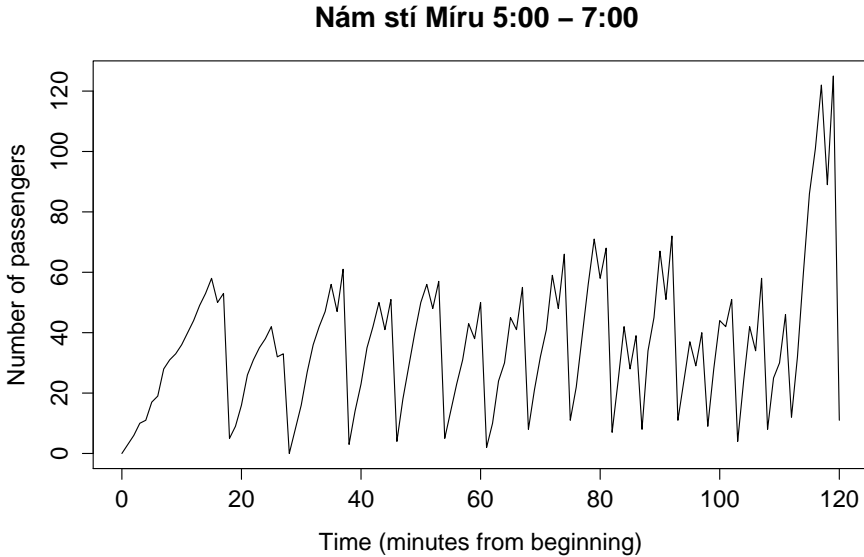
Nám stí Míru 5:00 – 7:00

Fig. 6. Number of passengers waiting at the station during 2 hours simulation period

One more possible issue, which is inherent to P systems, should be mentioned. Objects in P systems are not agents, do not posses (artificial) intelligence and do not make decisions. Their behavior is ruled by predefined set of rules, which can be probabilistic and resemble decision making, but essentially, decisions in P systems are not possible and therefore using proposed model for behavioral research of passenger choices would be problematic (i.e. research question "How long are passengers willing to wait for delayed train" is not appropriate for P systems, because requires passengers to make decisions).

The validity of model will be further examined and subjected to complex simulations. Case study only examined numbers of passengers waiting at stations, however occupancy of trains can be of interest for traffic management as well as examination of time, which is spend by certain groups of passengers in the system. Proposed model is discrete both in time and processing units (every passenger is considered an individual element in the system), therefore can be suitable for more sophisticated simulation. Research questions, which will be explored in the future are: How long time delays of trains are still manageable and which length of delays will cause the system to collapse? In the case of change of travel behavior of passengers, which changes could be made to increase the effectivenes of the system? If an increased number of passengers is introduced to the system (i.e. 1000 sport fans going to the game), how will the system respond?

5 Summary

In this paper, a P system based model for passenger flow simulation in public transport systems was proposed. Formal description of model was given and case study using Prague Metro network as example was performed. The case study did not reveal any errors in design of the model, however it became apparent, that correct estimation of numbers of passengers using system at given time is necessary. Cyclic patterns in numbers of passengers waiting at the stations were observed. Open problems associated with usage of P systems for traffic flow simulation were discussed and directions of future research were suggested.

References

1. F. J. Romero-Campero and M. J. Pérez-Jiménez. A model of the quorum sensing system in vibrio fischeri using p systems *Artificial Life 14 (1)*. 95-109 (2008).
2. M. Cardona and M. A. Colomer and A. Margalida and A. Palau and I. Pérez-Hurtado and M. J. Pérez-Jiménez and D. Sanuy. A computational modeling for real ecosystems based on P systems. *Natural Computing 10 (1)*. 39-53 (2011).
3. G. Ciobanu and M. J. Pérez-Jiménez and Gh. Păun. *Applications of Membrane Computing*. Springer, Natural Computing Series, 2006.
4. J. Dvorský and Z. Janoška and L. Vojáček. P Systems for Traffic Flow Simulation. *Lecture Notes in Computer Science 7564*. 405-415 (2012).
5. S. P. Hoogendoorn and P. H. L. Bovy. State-of-the-art of Vehicular Traffic Flow Modelling. *Delft University of Technology*. Delft, 2001.
6. M. Ionescu and Gh. Păun and T. Yokomori. Spiking Neural P Systems. *Fundamenta Informaticea 71 (2,3)*. 279-308 (2006).
7. S.N. Krishna and Gh. Păun. P Systems with Mobile Membranes. *Kluwer Academic Publishers*. Hingham, MA, USA, 2005. 279-308 (2006).
8. Gh. Păun. Computing with Membranes. *Journal of Computer and System Sciences 61 (1)*. 108-143 (2000).
9. A. Păun and Gh. Păun. The power of communication: P systems with symport/antiport. *Journal of Computer and System Sciences 20 (3)*. 295-305 (2002).
10. S. Peeta and A. Ziliaskopoulos. Foundations of Dynamic Traffic Assignment: The Past, the Present and the Future. *Networks and Spatial Economics 1 (1/4)*. 233-266 (2001).
11. P. Sosík. The computational power of cell division in P systems: Beating down parallel computers? *Natural Computing 2 (3)*. 287-198 (2003).
12. C. Zandron and C. Ferretti and G. Mauri. Solving NP Complete Problems Using P Systems with Active Membranes. *Unconventional Models of Computation*. Springer, 2000.
13. P systems web page. `http://ppage.psystems.eu/`.
14. T. Rejdal, www.metroweb.cz. `http://www.metroweb.cz/metro/81-71/81-71.htm`