Integrated Approach to Building a Microscopic City Model

Ilia Ludan
Samara National Research University
Samara, Russian Federation
ludanilya@gmail.com

Evgeny Maiorov
Samara National Research University
Samara, Russian Federation
benjamin1437@mail.ru

Johan David Motta Santana
Samara National Research University
Samara, Russian Federation
johanmotta10@gmail.com

Oleg Saprykin
Samara National Research University
Samara, Russian Federation
saprykinon@ssau.ru

Abstract

Due to the rapidly growing number of urban residents, a number of changes take place that affect the socio-economic processes of the city and the state generally. To solve the problems associated with transport systems, we can use special software for modeling, due to inability to influence the system in real life. However, the main issue that researches face during the creating of transportation models is inaccessibility of citizens' mobility data. The paper is devoted to developing methods and software of automated creation of transport microscopic models based on open data for any urban area. We propose the processing pipeline that consume the open data about city population and transport infrastructure and produce routes for agents of microscopic simulation. The developed method was implemented using the Python programming language in the Zeppelin interactive environment. The software was tested on data from the city of Samara (Russia). The obtained results show acceptability of proposed method and possible ways to improve the accuracy of model.

1 Introduction

Nowadays, countries are becoming more and more urban. The cities are growing, and the number of inhabitants is increasing. Transport systems of cities remain one of the main problems for residents. According to statistics, people spend about 3 months of their lives in traffic jams because of the poor condition of urban road networks. This fact makes us pay attention to the problem from the system's point of view. There are several modern approaches to solve this problem. The first method is to reconstruct the road network [1]. Another method to solve this problem is to create the most convenient and modern public transportation, so that the inhabitants of cities will change to a public one and no longer use their personal transport.

The investigation of transport systems is increasingly important and requires in-depth knowledge of this field. Today, there is a large amount of research devoted to this area. Discussing the transportation system, it is worth to note that it consists of several interconnected parts, therefore there are different research groups that work on specific areas related to the transport system.

The most attention should be paid to the construction of Origin-Destination (OD) matrices. OD matrix contains information about population movements in the entire city and provides capability to use it in modeling. The two most popular approaches to build OD matrices are the data driven-approach and model-driven approach.

The data-driven approach is based on the fact that the model is created from dynamic data of residents' movement in the entire city using GPS, social networks, etc. Applying GPS data to scientific research in transportation systems is very popular now. This approach is very relevant and modern, because GPS-navigation is widespread. In addition, the installation of the GPS-navigator in cars is now very common. The received data can be processed from GPS device, and as a result, an OD matrix can be created. The application of this method dramatically decreases the possibility to obtain an inadequate model. This area of research is currently relevant [2-6].

Another modern source of population movement data is information from social networks, such as Foursquare and Facebook. The method is based on the processing of geolocation data included in posts and photos. Having processed a rather large amount of data, it is possible to compile statistics showing the intensity of population movement in certain areas of the city at certain intervals. There is now a lot of research in this area as new social networks are emerging and being used [7-10].

The most accurate method for obtaining data of residents` mobility was proposed by Klaas Friso [11]. The high accuracy of the method is a result of the fact that the data is taken from the entire population using cellular communication. This technique allows for obtaining actual data, however, the processing of large datasets from different operators is very difficult. At the same time, the application of the method may be limited to the level of the Legislation of some countries, as well as the disagreement of subscribers to the processing of their personal data.

The model-driven approach is popular among the researchers too and used when GPS or mobile data are not accessible. Filippo Simini [12], who uses the universal model for mobility and migration uses gravitational method to obtain data on population movements throughout the country. This method allows tracing the trade flows in the country, as well as between countries, migration flows, and even intercity telephone calls. The gravitational method was proposed by W. Reilly [13]. The universal law of gravitation has become an idea for constructing a gravitational-type model. Applying this law to the transport system, it turns out that the objects are the points generating and absorbing traffic flows [14], and the total volume of the incoming and outgoing flow is taken as the object mass. It is possible to replace the physical distance with any costs [15] associated with the movement.

An idea for constructing the entropy model [16] was suggested by physics, and specific for the second law of thermodynamics. It is accepted by law that any closed physical system tends to achieve a stable equilibrium state, which is characterized by the maximum entropy of this system. The entropy method was first applied to the transport processes of A. Wilson [17]. The principle of the entropy method is that the system of residents' movement on the city road network has a fairly large number of uncontrolled elements. Related to the fact that our research is devoted to city transportation, it is necessary to consider the transport system as a closed one.

Michael Balmer compares two types of modeling in his work [18]: micro-simulation model and a traditional assignment model. The traditional assignment model is based on the use of OD matrices to reflect the movement of large population flows [19]. The main advantage of the micro-simulation model is that it is possible to simulate the movement of each agent individually [20].

Modern transport simulation systems allow creating the visual representation of traffic flows. Currently, there is a wide specter of transport simulation software, a few of the more popular examples are T7F/TSIS, SUMO [21], TRANSYT, VISUM, MATsim, Aimsun. They allow the verification of transport infrastructure changes, without special economic and time costs. Moreover there is research to build OD matrices from microscopic simulation outcome [22].

Despite a large number of studies in the field of transport modeling, some issues still remain. The most important issue is related to a model creating automation. This paper is devoted to developing methods and software of automated creation of transport microscopic models for any urban area based on open data. The developed system provides the ability to obtain the up-to-date information from open Internet resources, process it, and apply it for transportation model creating or updating. As an example of open resource, we consider OpenStreetMap that provides information about city infrastructure collected by volunteers. Using information from OpenStreetMap, we have defined transportation supply and demand for building a microscopic transportation model. The proposed approach allows getting models with acceptable quality using only data from public resources.

2 Transportation city model

The study proposes to use modeling for investigating capabilities to change the road network structure to increase its capacity. Modeling is used to study real objects by building models to reveal their latent characteristics, and also influence them. There are several types of modeling that allow creating a model of a specific object. Applied to natural and technical sciences, it is customary to distinguish the following types of modeling: conceptual, physical, structural-functional, mathematical, and simulation. The choice of the type of modeling depends on the goal of the research.

In our study, we use simulation, which allows us to build models describing processes, in our case, transportation processes. The basis for choosing this method is the dynamic nature of the transport system, as well as the inability to create a static model.

To date, there are a number of software applications that make it possible to create a transport model of the city at various levels of detail - from macroscopic to microscopic [23]. In this work, we use microscopic simulation, to create a microscopic model. The advantage of microscopic simulation is the ability to model individual elements, in this case, vehicles and transport infrastructure objects.

The main goal is creation of a universal method (Fig. 1), which would be suitable for modeling the transport processes of any city using open data.

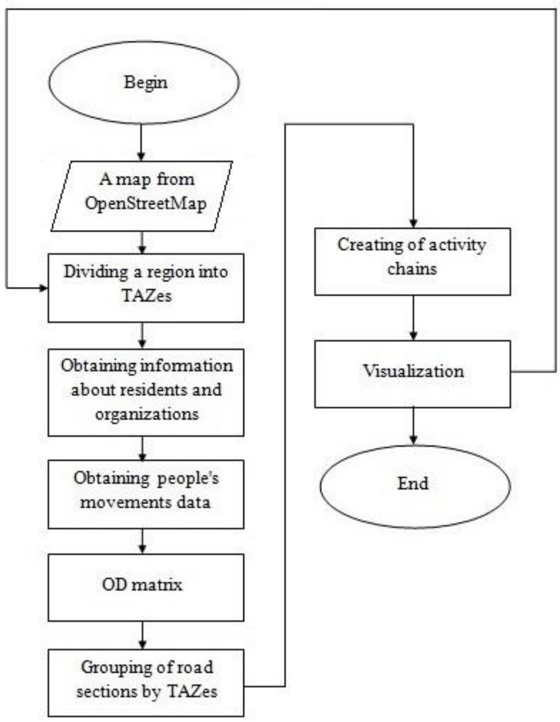

Figure 1: The method of activity chains constructing.

Creation of the most accurate model is possible by solving several problems [24]: splitting the city into traffic analysis zones (TAZ), constructing OD matrices, and building activity chains. The most accurate model of transport processes is created by dividing the city into traffic analysis zone of a certain size, within which the citizens do not use transport; however, they use public or personal transport for traveling between zones.

2.1 Traffic Analysis Zones

The division by TAZ is the foundation of correct elaboration of transport flow models [25]. Starting from these divisions, we can elaborate an image that allows the identification of several economic, political and social aspects of the city in question.

The concept of TAZ can be understood as the elementary area used for the analysis of transport models. With them, we can determine the flow of vehicles and pedestrians moving from region A to region B. At the time of the splitting of zones, the main criterion to be taken into account the points of the largest density flow, around which regions are formed, playing down such criteria as the physical size of the regions as suggested in fundamentals of transport modeling [26].

In the establishment of the Traffic Analysis Zone, it is important to choose areas where a relation between the origin and destination exists. For the determination of the scale of the transport regions it is necessary to keep in mind that the more the size of said area increases, the less effective the model will be.

On the criteria described above, the regions were extracted according to the main road arteries and natural boundaries (Fig. 2). The proposed approach allows the models performance to be more accurate and provides better possibilities for improvement. In addition, each region can be subdivided, to correspondingly make a model on this specific area. Therefore we obtain scalable model which can be adapted to solve transportation problems at different levels with less handling of data and information.

2.2 Origin-Destination Matrices

We used the gravitation model to construct the OD matrix because it has sufficient quality for urban areas and is not very sensitive to computational resources. The transport gravity model relates to the intensity of the flow. Between the total number of departures from the i-th zone and arrivals to the j-th zone and the cost of travel between zones i and j (1).

$$T_{ij} = \frac{Q_i * D_j}{c_{ij}^2} i = 1, \dots, N, j = 1, \dots, M, \qquad (1)$$

Where N is the total number of departure areas, and M is the total number of arrival areas. In this model, the distance between areas is considered the distance between the centers of these areas.

There is an opportunity to receive data from any GIS required for finding the intensity of the flow based on information about residential buildings and organizations of the city. To involve organizations into a model we need to receive geographic and attribute data about them. We also need information about residential buildings: number of stores and coordinates. With all of this information, we can compare the average number of working or living people in each organization and residential building. Coordinates will allow the location of residential buildings and organizations to be displayed on the city map.

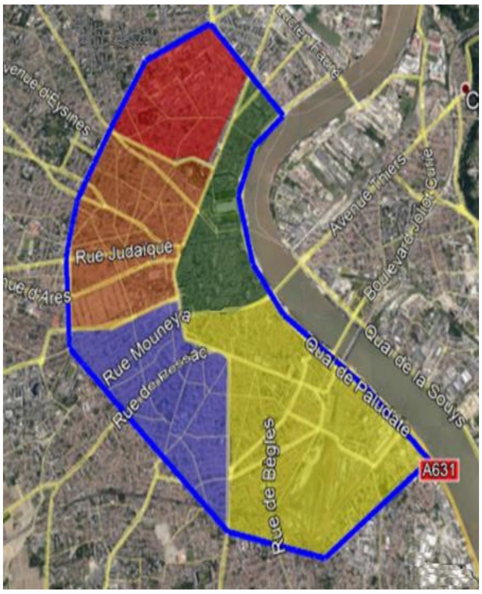

Figure 2. Splitting of some urban area on traffic analysis zones.

2.3 Method of Activity Chain Building

After dividing the city into traffic analysis zones and constructing the OD matrix, it is necessary to distribute the agents in the model according to the obtained data and to build a route for each of them. This is the process of activity chains building. To obtain them, it is necessary to determine which TAZ region each road section belongs to. For this purpose, the operation intersection is performed, which allows overlaying the two tables and intersecting the data sets. After that, agents are distributed along the road network. Start, finish and intermediate points are assigned according to the OD matrix. The last step is to combine all obtained data to visualize the simulation process, and then to compare the resulting model with the actual situation on the city's road network. If necessary, the model is corrected using additional data from the actual transport network.

3 Implementation

The automated method of microscopic transport models building was implemented using Python programming language. We have used the data for the city of Samara (Russia) for debugging the model [20].

Innovations in GIS make possible the modeling of transport processes using open data, (e.g. OpenStreetMap. Interface can be used and adapted to modeling and editing Traffic Analysis Zones with OSM SlippyMap Generator (Fig. 3)).

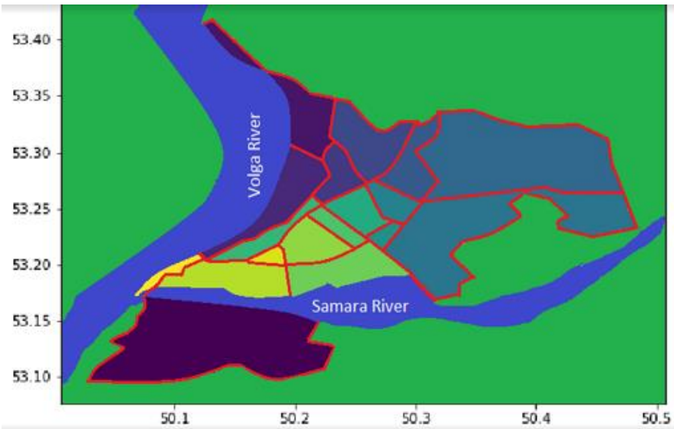

Figure 3. Splitting the city of Samara on traffic analysis zones.

For visualization, the resulting matrix was colored in characteristic colors (Fig. 4). The less the intensity between the two transport areas, the shade of the matrix cell is closer to red, and vice versa, the higher the intensity, the closer to green.

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18
1	0	1294,67	698,21	402,37	1158,95	1025,59	635,83	537,59	458,94	383,16	1052,83	857,82	863,24	810,03	754,59	612,58	322,53	164,85
2	1235,82	0	1358,4	753,27	741,53	994,8	752,51	626,02	499,02	399,01	744,9	515,82	534,16	743,52	267,29	223,63	379,85	126,27
3	761,53	1207,08	0	1142,21	594,14	736,81	692,58	869,53	561,96	652,61	636,16	456,91	514,4	500,07	378,41	327,86	275,37	116,05
4	471,07	878,31	1181,92	0	438,2	582,35	402,33	617,91	520,22	792,73	434,83	482,52	364,21	361,06	214,56	195,51	155,7	88,42
5	1381,36	760,02	754,73	472,11	0	1148,82	780,52	496,62	467,55	401,81	1043,59	321,23	756,72	764,16	878,12	412,43	349,84	186,85
6	1215,08	1327,68	879,42	510,04	956,31	0	974,59	526,06	513,92	447,9	958,23	811,04	765,41	805,21	466,08	403,66	302,15	251,85
7	913,41	811,24	910,05	529,23	835,72	1053,33	0	915,72	952,3	775,59	885,41	742,91	671,84	636,42	386,5	306,42	348,93	231,14
8	545,89	923,71	1069,41	532,61	712,51	823,5	932,02	0	1142,91	896,51	867,5	629,66	492,51	501,52	396,26	315,27	367,85	225,8
9	467,32	612,04	743,56	478,18	683,62	692,83	895,21	962,61	0	977,95	800,51	487,72	395,01	583,24	234,95	322,06	552,9	273,36
10	231,76	514,67	701,9	883,64	503,3	471,92	751,5	847,26	1022,04	0	613,72	419,5	476,41	517,68	503,65	312,64	752,82	229,05
11	1142,58	572,92	634,68	351,5	985,68	853,84	827,83	524,81	374,21	421,26	0	1101,49	1123,82	1152,05	974,16	408,68	711,94	477,17
12	945,27	492,61	473,31	314,94	830,58	732,59	642,56	471,01	386,64	448,51	1003,15	0	1267,84	994,73	866,01	1062,69	879,16	587,82
13	852,02	578,02	512,58	338,02	813,05	788,02	662,93	480,66	412,23	456,02	1192,84	1252,6	0	1215,93	862,15	1023,6	1105,8	502,66
14	821,63	732,16	526,93	392,51	752,33	807,49	673,85	509,41	562,69	528,91	1082,52	1017,8	1178,41	0	732,95	963,73	986,23	456,81
15	1078,41	251,5	379,2	251,55	863,25	459,22	407,25	393,19	276,23	310,62	923,68	842,23	926,07	782,16	0	1092,85	812,34	683,86
16	630,86	213,92	338,56	179,41	438,62	409,31	349,2	327,02	332,9	421,95	998,5	975,18	944,11	834,83	1154,65	0	1023,85	912,05
17	315,12	397,41	243,03	141,52	368,96	331,83	358,51	380,41	502,73	793,66	712,51	866,81	1020,4	956,71	834,63	1064,68	0	870,51
18	179,63	115,05	108,97	85,57	193,82	210,93	202,49	244,51	274,63	230,6	469,48	576,26	512,49	477,25	647,75	893,58	883,92	0

Figure 4. Origin-Destination matrix.

To build the activity points it is necessary to separate road network sections by TAZ districts. Using the OSMnx library is a way to automatically get geodata of the city. Geodata can be obtained only by just entering the name of the city. There we can receive data about coordinates and type of organizations and residential buildings of the city of Samara in the format of GeoDataFrame (GDF). The next step of the method is to divide the city into transport areas. To implement this method, the GeoPandas library was used to perform various manipulations with a geographic database, which contains various tables with initial data: names of streets, districts, roads and so on.

After obtaining a database with the coordinates of roads and areas, it becomes necessary to determine the affiliation of roads to a certain area. In the library, there is special operations to perform various overlays of two spatial tables, as well as visualize them. In the study, by means of the sjoin operation, the traffic analysis zone (TAZ) and their roads are united.

The next step in the method of activity chain constructing is to create a file containing general data on the affiliation of all the city roads to a particular transport area. The final step of the method is the entry of all the data obtained into the SUMO program for further modeling.

The next steps are processing the received data in programs OD2TRIPS and DUAROUTER. OD2TRIPS allows defining the start and end points of a path, but it does not have the ability to compile route on real roads. In other words, program OD2TRIPS creates the starting point of the agent's movement and the end point of its path. In turn, the DUAROUTER program serves to compile a route along those roads, the data of which were obtained in the TAZ file. These programs are complementary to each other and create a well-constructed direction of traffic flows.

After using the created method of constructing activity chains and putting all the data in the SUMO program, it is possible to start modeling urban transport processes (Fig. 5). The application of this method is possible for any city, subject to obtaining data on the transport areas of the city and their residents.

Figure 5. Modeling of the transport processes of the city of Samara.

4 Conclusions

The developed method allows easily building a microscopic model of an interested urban area. However, at this moment, it has certain shortcomings and limitations. The main issue is inconsistency of the model sections with real traffic in some road network. For example, in the case of the city of Samara, most of the agents are moving along the Moskovskoye and Zavodskoye highways, ignoring roads that are no less used in real life, such as Novo-Sadovaya or Kirovskaya. There is also the problem of the emergence of new agents in the middle of the road, while they should appear near residential buildings. Generally, this issue is related to inaccuracy of open data or missing some information about transport infrastructure objects. This limitation requires manual verification and fixing the model, so it can only work automatically in the simplest cases.

Despite all the shortcomings listed above, the method of constructing activity chains completely works. Even today it can be used to create a model of transport processes in other cities. The advantage of the developed method is that only initial data is needed in order to create activity chains and OD matrices for the city (i.e. information on transport areas and number of residents of a particular transport area).

Further work is intended to reduce the scale of TAZ areas, which should lead to improvement of the model because it will reflect the urban area in more detail. Another pending task is developing a method for obtaining information about city organizations and clarifying information about residents.

Acknowledgements

This work was supported by the Ministry of Education and Science of the Russian Federation.

References

1. Bundesministeriumfür Verkehr. Bau und Wohnungswesen (BMVBW).: Neubauvon Bundesautobahnen, 2004.
2. Gasnikov, A.V., Klenov, S.L., Nurminskiy, E.A., Kholodov, Y.A., Shamrai, N.B.: Vvedeniye v matematicheskoye modelirovanie transportnykh potokov, 2010, MFTI. (In Russian).
3. Moreira-Matias, L.: Time-evolving O-D matrix estimation using high-speed GPS data streams, Expert Systems With Applications, 2016, 44: 275–288.
4. Russell, J.S., Ye, M., Anderson, B.D.O., Hmam, H., Sarunic, P.: Cooperative Localisation of a GPS-Denied UAV in 3-Dimensional Space Using Direction of Arrival Measurements, IFAC, vol. 50, 2017, issue 1:8019-8024.
5. Singh, S., Sujit, P.B.: Landmarks based path planning for UAVs in GPS-denied areas, IFAC, vol 49, 2016, issue 1:396-400.
6. Kim, B., Kim, D., Park, S., Jung, Y., Yi, K.: Automated Complex Urban Driving based on Enhanced Environment Representation with GPS/map, Radar, Lidar and Vision, IFAC, vol. 49, 2016, issue 11: 190-195.
7. Kheiri, A.A.: Intra-urban movement flow estimation using location based social networking data, The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XL-1/W5, 2015 International Conference on Sensors & Models in Remote Sensing & Photogrammetry, Kish Island, Iran.
8. Zhang, Z., Wang, Z.: The data-driven null models for information dissemination tree in social networks, Physica A: Statistical Mechanics and its Applications, 2017, 484:394-411.
9. Yu, C., Xiao, B., Yao, D., Ding, X., Jin, H.: Using check-in features to partition locations for individual users in location based social network, Information Fusion, 2017, 37:86-97.
10. Jiao, Z., Ran, L., Chen, J., Meng, H., Li, C.: Data-Driven Approach to Operation and Location Considering Range Anxiety of One-Way Electric Vehicles Sharing System, Energy Procedia, 2017, 105: 2287-2294.
11. Friso, K.: Enriching the transport model of the Rotterdam region by cell phone data, Models and Technologies for Intelligent Transportation Systems (MT-ITS), 2015.
12. Simini, F.: A universal model for mobility and migration patterns, Macmillan Publishers Limited, 2012.
13. Reilly, W.J.: The law of retail gravitation, New York, 1931.
14. Wiedemann, R.: Simulation des Straßenverkehrsflusses, PhD-thesis, University of Karlsruhe, Germany, 1974.
15. Beckmann, M., McGuire, C.B., Winsten, C.B.: Studies in the economics of transportation, RM 1488. Santa Monica: RAND Corporation, 1955.
16. Fang, S.C., Rajasekera, J.R., Tsao, H.-S.J.: Entropy optimization and mathematical programming, Kluwer Academic Publisher, 1997.
17. Wilson, A.G.: Entropy in urban and regional modeling, London: Pion, 1970.
18. Balmer, M.: Generating Day Plans Based on Origin-Destination, 2005.
19. Roughgarden, T., Tardos, E.: How bad is selfish routing? Journal of the ACM, 2002.
20. Krajzewicz, D., Erdmann, J., Behrisch, M., Bieker, L.: Recent Development and Applications of SUMO – Simulation of Urban Mobility, International Journal On Advances in Systems and Measurements, 2012.
21. Hofer, C., Jäger, G., Füllsack, M.: Generating Realistic Road Usage Information and Origin-Destination Data for Traffic Simulations: Augmenting Agent-Based Models with Network Techniques, In: Cherifi C., Cherifi H., Karsai M., Musolesi M. (eds) Complex Networks & Their Applications VI. COMPLEX NETWORKS 2017. Studies in Computational Intelligence, 689, pages 1223-1233. Springer, Cham, 2018.

22. Saprykin, O., Saprykina, O.: Multilevel Modelling of Urban Transport Infrastructure, In Proceedings of the 1st International Conference on Vehicle Technology and Intelligent Transport Systems (VEHITS-2015). Portugal, Lisbon: SCITEPRESS, 2015: 78-82.
23. Chen, M., Bernstein, D.H., Chien, S.I., Mouskos, K.: Simplified formulation of toll design problem, Transportation Reseach, 1999.
24. Martinez, L.M., Viegas, J.M., Silva, E.A.: A traffic analysis zone definition: a new methodology and algorithm, Springer Science+Business Media, LLC, 2009.
25. Gorev, A.E., Buttger, K., Prokhorov, A.V., Gizatullin, R.R.: Transport modeling fundamentals, St. Petersburg, Kosta, 2015, 168 p.
26. Saprykin, O., Saprykina, O.: Validation of Transport Infrastructure Changes via Microscopic Simulation: A Case Study for the City of Samara, Russia, In Proceedings of the 5th IEEE International Conference on Models and Technologies for Intelligent Transportation Systems (MT-ITS 2017) in Naples, Italy, 26-28 June, 2017: 788-793.