# Profile-driven synthetic trajectories generation to enhance smart system solutions

Radosław Klimek[1][0000-0002-9061-561X] and Arkadiusz Olesek

[1] AGH University of Science and Technology
al. Mickiewicza 30, 30-059 Krakow, Poland
rklimek@agh.edu.pl

**Abstract.** The knowledge of the individual trajectories of citizens' mobility in the urban space is critical for smart cities. The data concerning trajectories from the providers of mobile phone services are still difficult to be obtained in practice and one of the considerable obstacles here are legal aspects. We have designed and implemented the tourist trajectories generator for objects located in a selected but arbitrary urban area. A generation process is based on the random selection of the pre-defined profiles of tourist activeness, including mobility patterns. It is possible to generate a practically unlimited number of trajectories, if needed, and they may also be directed at the certain specific types of behaviours. Thus obtained large sets of data may be used for understanding urban behaviours, calibrating urban models, recommending systems under construction, as well as anticipating the smart city further software testing.

**Keywords:** smart system · data management · synthetic trajectories generation · urban ecosystem calibration

## 1 Introduction

There is a large number of the possibilities for obtaining data on mobile phone individual trajectories, e.g. tracking applications in smartphones. It appears that the localisation data of mobile phone operators are most interesting. This data is not conditioned on the possession of any advanced smartphone or intentional enabling of a proper application. Nowadays, almost everyone has a mobile phone, even the simplest one; therefore, localisation data from the providers of mobile telephony is the most common in this regard, democratic and reliable. Localisation takes place based on triangulation and trilateration methods in BTSs (Base Transceiver Stations), which guarantee high accuracy in the urban area [2,5]. The spread of mobile devices in the recent years has contributed to the performance of many original researches concerning human mobility, and the mobile phone data collected by network operators have become the invaluable source of information about individual mobility patterns on a large scale. Whereas, mobility and individual trajectory data are still poorly available although it is significant for a *smart city*, which provides supporting services and solving city problems. A main hindrance here are legal aspects referring to privacy, see for example [10,9,3].

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77970-2_47

Our goal and contribution is to design and implement a generator of individual mobility trajectories for mobile phone users. Trajectories mirror the behaviours of tourists staying in a given urban area. Trajectories are generated at random based on the prepared profiles of tourism behaviours, considering the various points of interests (POIs) located within the urban area. The respective elements of behaviour profiles are selected randomly for each object. The following parameters are considered: the prospective interests of tourists, their variability and possible habits connected with tourists mobility, intensity in respect of the implementation of a tourism plan, changing the means of transport and other aspects.

We believe that there are numerous prospective applications for thus obtained mass data concerning tourist mobility. These are: the analysis and understanding of the tourist traffic needs, the calibration of urban models and operational procedures, recommendation systems, increasing tourists' safety, in particular in dangerous regions or districts, solutions supporting the designing of the urban infrastructure, applications supporting the control of epidemic spread, solutions helping in emergency management and many others. Furthermore, possessing such data constitutes a perfect base for testing systems concerning the smart city which are being created or which will be created in the future.

## 2 Related works

The legal regulations concerning mobile devices localisation are not too restrictive in the Asian countries, they are moderately restrictive in the USA and extremely restrictive in Europe, cf. Uhlirz et al. [12]. This last case is the main reason for the low availability of such data, despite their immense potential. A certain example here is Estonia, see Ahas et al. [1], where anonymised data are analysed, which refer to a certain group of users but at the same time they demonstrate huge advantages from possessing and analysing such data. The possibilities of the tourist trajectories analysis are also discussed in the paper [6]. Kwan et al. [7] describe an experiment of the limited data disclosure by Deutsche Telekom for a certain selected period from the past. Such data allow to analyse various hypotheses, verify them in terms of various initiatives, including start-ups, the assumed methodologies of data analysis, etc. Therefore, the objects mobility trajectories are crucial here. In the work by Gonzalez et al. [4], the following was analysed: the use of trajectory in searching for similarities, finding general patterns, the distribution of spatial probability, what is significant for the urban models calibration or understanding the spreading of e.g. an epidemic.

The aspect of trajectory generation has already been considered. Pelekis et al. [10] suggest the generator of objects which follow a specific main objective of such a generator. This differs considerably from our solution which involves object profiles; moreover, our work includes some sets of targets which are inter-related unambiguously with the pre-defined profiles. Pappalardo and Simini [9] generate trajectories taking into account some tendencies, also considering a tendency for breaking the routine. The work does not imply any basic source

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77970-2_47

of obtaining mobile data. In the work by Giurlanda et al. [3], there is the trajectory simulator considered, concerning people's habits based on the behaviour model. This work affects our approach; yet, we consider behaviour profiles for tourists in the urban area. Zhou et al. [13] suggest the system of recommending points connected with travelling on the basis of preferences and journey time estimation. The aim is to match the trajectory with the users based on the objects mobility patterns. Networks are trained and trajectories are corrected as a result of the min-max strategy. In our paper, trajectories are generated without their corrections based on random behaviour profiles. Renso et al. [11] discuss various trajectory generators, comparing real data and synthetic data. The work discussed implies that our generator may have numerous applications. According to our best knowledge, it differs from the recently known generators, and a distinguishing factor is the random generation of trajectory based on various profiles which are weighted additionally in order to obtain better realism. Our paper is based on the unpublished work [8]. Nevertheless, a significant summary was performed and new results were generated.

## 3 User profiles and POIs planner

The main task of the generator is to produce trajectories which represent human behaviours as naturally as possible. The trajectory quality is understood here as the capability of imitating human behaviours in a natural manner. In order to ensure the conformity of the visited POI with the preferences of a given tourist, a *user profile* was designed and formed as a structure allowing for the description of features and the inclinations of a specific tourist in an adequate and implementable manner generator algorithms.

$$Profile = (I, f(i_n) \rightarrow w_n, S, d, M, T, f(t_n) \rightarrow v_n, a, p)$$

$I$ is the set of profile interests. Function $f(z_n) \rightarrow w_n$ allocates weights to a category from set $I$. A weight means here a degree of being interested in each of the elements, expressed as a number from the interval $(0, 1)$. The sum of all the weights for the profile amounts to 1. $S$ – the geographical coordinates of the journey starting point for a given profile. $d$ – a suggested maximum distance in meters between each POI matched with the profile, and a starting point $S$. $M$ – a material status allocated to the profile, it is determined in a five degree scale. $T$ is the set of the preferred means of transport, and function $f(t_n) \rightarrow v_n$ allocates a weight to each means of transport. The preferred activeness time is marked as $a$, it is used when determining time for trajectory. Value $p$ is another structure containing additional tourist's preferences used in the process of allocating the points.

The *Planner of POIs* task is to create a trajectory plan based on the preselected POIs. Each POI set for an urban area is obtained automatically by publicly available services, so it does not require any (manual) preparations. Firstly, the planner establishes the place of a trajectory starting point and the time of its starting. Then, searching for the best POI to be visited next takes

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77970-2_47

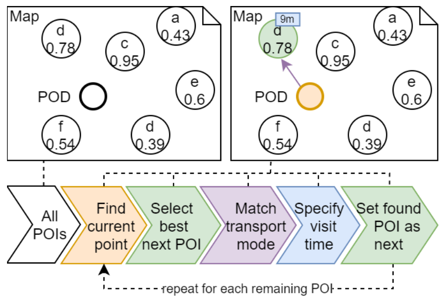


**Fig. 1.** The operating rule of the trajectory planner is presented as a simplified outline. (POD – Place of Departure, POI – Point of Interest). The maps depict the respective stages of trajectory generation. The first one presents an initial state, that is POIs and their assessments for the profile. A middle map is a stage repeated until all the points are planned. The last map is a ready-made trajectory plan; and there is also the outline of the ready-made trajectory plan below the last map. The outline provided below presents the steps taken at each stage; a dashed line joins given steps with the corresponding fragment of the map. Furthermore, the respective corresponding elements of the outline are marked with the same colours

place. It is searched in the supplied points, for which opening hours are known and for which it is possible to reach a given point timely or points for which we have no information about opening hours. In an extraordinary case, when it is not possible to create a plan consistent with opening hours, information about such opening hours may be ignored for given points. After determining another POI, the planner selects a planned transport mode matching the preferences for the profile, it estimates the travelling time to the next point and then it defines the planned visiting time in this point. Defining the visiting time takes place at random but the following factors are taken into consideration: an appropriate profile attribute and average visiting time in a given point. All the pre-calculated parameters are collected to the structure describing one fragment of the trajectory plan. At the end of each step, a selected point is marked as a current one and the process is re-implemented. While modifying the planner configuration, we can decide whether the trajectories are to be terminated in the last POI or the tourists are to return to the starting point after visiting all the points.

After implementing the steps described above for all the supplied POIs, we obtain a daily trajectory plan which consists of parts ordered depending on their planned visiting sequence. Each part contains information enabling the creation of a corresponding fragment of a trajectory. The operating rule of the planner is also presented in Figure 1.

## 4 Generating results

In order to validate our system, we have carried out several generating processes. Table 1 contains details concerning the generated sets. Figure 2 presents a frag-

**Table 1.** The generated data sets ($A$, $B$, $C$)

| Set | Planned number of trajectories | Number of obtained trajectories | Number of unique POIs |
|---|---|---|---|
| $A$ | 100 | 92 | 306 |
| $B$ | 1000 | 976 | 713 |
| $C$ | 5000 | 4853 | 865 |

ment of the visualisation made for a generated large trajectory set (set $C$, 4853 trajectories, 865 unique POIs).

Table 2 presents the statistics generated based on the analysed sets. The statistics refer to transport methods and parameters connected with travelling, visiting POIs and the time in which the respective trajectories were placed.

Table 3 presents the distribution of 10 most popular categories of POIs taken from 976-element set $B$ and 4853-element set $C$.

We believe we have built an important and universal tool. Our system is characteristic in this regard as compared to the existing ones [10,9,3,13,11]. The results obtained prove the credibility and realism in reference to the trajectories which thus are reliable. The trajectories generated are synthetic but much effort was made so that they could look realistic and so that they could consider both various profiles of human behaviours and a randomness factor.

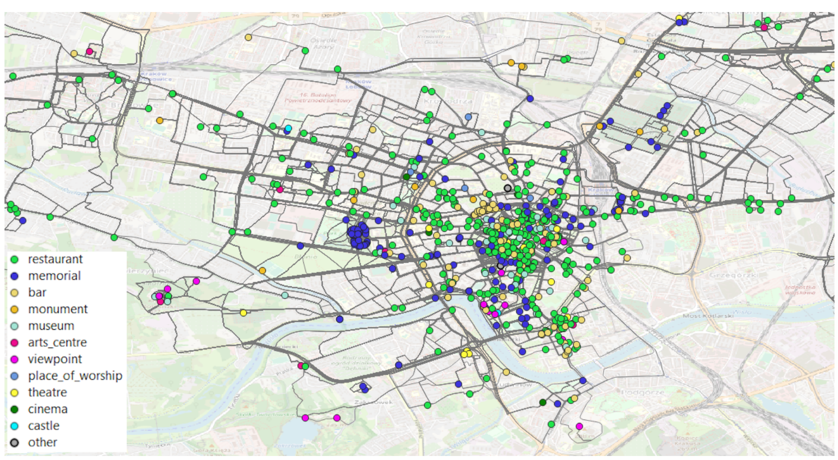


**Fig. 2.** The fragments of the visualisation made for Kraków City for 4853 generated trajectories (set $C$). On account of a large volume of data, trajectories are represented by unicolour lines. POIs are represented by dots whose colours denote different object categories

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77970-2_47

**Table 2.** Table presenting statistical data obtained on the basis of the analysis of the generated sets

| | **Set A** | **Set B** | **Set C** |
|---|---|---|---|
| The earliest time stamp | 2021-01-03 06:00 | 2021-01-03 06:00 | 2021-01-03 06:00 |
| The latest time stamp | 2021-01-04 07:03 | 2021-01-04 11:08 | 2021-01-04 14:24 |
| Average duration time of a trajectory [h] | 7.67 | 7.35 | 7.46 |
| Average number of POIs visited within one hour | 1.18 | 1.25 | 1.27 |
| Average number of all the visits during one day (0-24) | 434.5 | 4393.5 | 22434 |
| Average distance covered in one trajectory [km] | 12.70 | 12.30 | 12.37 |
| Average number of visited POIs within one trajectory | 8.69 | 8.78 | 8.97 |
| Average number of visits for each POI | 2.83 | 12.32 | 51.87 |
| Number of all the POIs | 306 | 713 | 865 |
| Average number of means of transport within one trajectory | 1.96 | 2.02 | 1.97 |
| Available transport mode | driving-car, cycling-regular, cycling-road, foot-walking | | |

**Table 3.** 10 most popular categories of POIs on account of the number of visits for set $B$ and set $C$

| **Set B** | | | **Set C** | | |
|---|---|---|---|---|---|
| **Category** | **Visits** | **Percentage** | **Category** | **Visits** | **Percentage** |
| *Total* | 8787 | 100.0 % | *Total* | 44868 | 100.0 % |
| restaurant | 2572 | 29.27 % | restaurant | 13728 | 30.60 % |
| bar | 1298 | 14.77 % | memorial | 6642 | 14.80 % |
| memorial | 1254 | 14.27 % | bar | 6410 | 14.29 % |
| monument | 682 | 7.76 % | monument | 3542 | 7.89 % |
| museum | 547 | 6.23 % | museum | 2603 | 5.80 % |
| arts_centre | 459 | 5.22 % | arts_centre | 2268 | 5.05 % |
| viewpoint | 423 | 4.81 % | viewpoint | 2029 | 4.52 % |
| place_of_worship | 398 | 4.53 % | place_of_worship | 1946 | 4.34 % |
| theatre | 368 | 4.19 % | theatre | 1943 | 4.33 % |
| cinema | 356 | 4.05 % | cinema | 1744 | 3.89 % |

## 5 Conclusions

By using our system, it is possible to generate practically any number of tourist trajectories, depending on demand and in the real time. Further works may comprise other groups of people instead of tourists, e.g. regular citizens.

ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77970-2_47

## References


1. Ahas, R., Aasa, A., Roose, A., Ülar Mark, Silm, S.: Evaluating passive mobile positioning data for tourism surveys: An Estonian case study. Tourism Management **29**(3), 469–486 (2008). https://doi.org/https://doi.org/10.1016/j.tourman.2007.05.014
2. Calabrese, F., Colonna, M., Lovisolo, P., Parata, D., Ratti, C.: Real-time urban monitoring using cell phones: A case study in Rome. IEEE Transactions on Intelligent Transportation Systems **12**(1), 141–151 (2011). https://doi.org/10.1109/TITS.2010.2074196
3. Giurlanda, F., Perazzo, P., Dini, G.: Humsim: A privacy-oriented human mobility simulator. In: Kanjo, E., Trossen, D. (eds.) Sensor Systems and Software - 5th International Conference, S-CUBE 2014, Coventry, UK, October 6-7, 2014, Revised Selected Papers. Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering, vol. 143, pp. 61–70. Springer (2014). https://doi.org/10.1007/978-3-319-17136-4_7
4. Gonzalez, M.C., Hidalgo, C.A., Barabasi, A.L.: Understanding individual human mobility patterns. Nature **453**(7196), 779–782 (June 2008). https://doi.org/10.1038/nature06958
5. Klimek, R.: Exploration of human activities using message streaming brokers and automated logical reasoning for ambient-assisted services. IEEE Access **6**, 27127–27155 (2018). https://doi.org/10.1109/ACCESS.2018.2834532
6. Klimek, R.: Towards recognising individual behaviours from pervasive mobile datasets in urban spaces. Sustainability **11**(6), 1–25 (2019). https://doi.org/10.3390/su11061563
7. Kwan, M., Cartwright, W., Arrowsmith, C.: Tracking Movements with Mobile Phone Billing Data: A Case Study with Publicly-Available Data, pp. 109–117. Springer Berlin Heidelberg, Berlin, Heidelberg (2012). https://doi.org/10.1007/978-3-642-24198-7_7
8. Olesek, A.: Mobile phone trajectory generator using user profiles [in Polish], Master thesis, supervisor: Radosław Klimek, AGH University of Science and Technology (2020)
9. Pappalardo, L., Simini, F.: Data-driven generation of spatio-temporal routines in human mobility. Data Min. Knowl. Discov. **32**(3), 787–829 (2018). https://doi.org/10.1007/s10618-017-0548-4
10. Pelekis, N., Ntrigkogias, C., Tampakis, P., Sideridis, S., Theodoridis, Y.: Hermoupolis: A trajectory generator for simulating generalized mobility patterns. In: Blockeel, H., Kersting, K., Nijssen, S., Železný, F. (eds.) Machine Learning and Knowledge Discovery in Databases. pp. 659–662. Springer Berlin Heidelberg, Berlin, Heidelberg (2013)
11. Renso, C., Puntoni, S., Frentzos, E., Mazzoni, A., Moelans, B., Pelekis, N., Pini, F.: Wireless Network Data Sources: Tracking and Synthesizing Trajectories, pp. 73–100. Springer Berlin Heidelberg, Berlin, Heidelberg (2008). https://doi.org/10.1007/978-3-540-75177-9_4
12. Uhlirz, M.: A Market and User View on LBS, pp. 47–58. Springer Berlin Heidelberg (2007). https://doi.org/10.1007/978-3-540-36728-4_4
13. Zhou, F., Yin, R., Trajcevski, G., Zhang, K., Wu, J., Khokhar, A.: Improving human mobility identification with trajectory augmentation. GeoInformatica pp. 1–31 (2019)


ICCS Camera Ready Version 2021
To cite this paper please use the final published version:
DOI: 10.1007/978-3-030-77970-2_47