# Detecting Anomalous Trajectories and Traffic Services

Mazen Ismael

Faculty of Information Technology, BUT
Božetěchova 1/2, 66 Brno

Mazen.ismael@vut.cz

**Abstract.** Among the traffic studies; the importance of detecting anomalous trajectories of vehicles rises to support many services, starting from securing and safety services to the maps and navigation services. The combination of many methods and concepts could offer interesting advantages, and *iBAT* (Isolation-Based Anomalous Trajectory) is one of the advanced frameworks which detect anomalous traffic trajectories. *iBOAT* (Isolation Based Online Trajectory) came after that as a version of *iBAT* able to process online data. Beside of that, using semantic locations could support the navigations studies, and increase the maps' accuracy. In fact, developing the *iBOAT* framework with use semantic ocation could bring out interesting results. The aim of this paper is to present the progress of detecting anomalous driving patterns from GPS trajectories, which will be achieved by using the concept of semantic locations for improving the scene partitioning.

**Keywords:** Anomalous Trajectory, Semantic Location, Traffic, iBAT/iBOAT

## 1 Introduction

Nowadays, several methods and frameworks have been proposed to achieve the goal of detecting anomalous trajectories by using *GPS* data, which aim to support securing, safety services, the maps and navigation services, and many other services. Some of these frameworks [1] [2] had good results but still need to improve their methods to achieve satisfactory results. Important preprocessing methods have been used by some frameworks [3] [4] to prepare the backgrounds. Moreover, semantic locations concept could play an efficient action in these methods for improving clustering of the trajectories to divide the background into zones. It is useful to mention here that the techniques will be used for design the framwork are connected to the offline data only. In this paper, we will present the work progress in design our framework, and review the current progress. Beside of that, we will show our implementation of the *iBOAT* framework. Finally, conclusion and future work will be presented.

*J. Steinberger, M. Zíma, D. Fiala, M. Dostal, M. Nykl (eds.)*
*Data a znalosti 2017, Plzeň, 5. - 6. října 2017, pp. 192-197.*

## 2 Related work

In this section; related methods in each phase of our framework will be presented. The first phase is a preprocessing which mainly focuses on preparing the data and by partitioning the scene. In this area, *Zheng, Y* [5] reviewed many algorithms for generating and reducing the storage of the data (Uniform sampling, Douglas-Peucker (DP), Top-down time-ratio (TD-TR), Bellman). Besides that, many filters were reviewed by *Zheng, Y* [5](Mean and Median Filters, Kalman Filter, Particle Filter) to filter the trajectories and ignore the outliers and reduce their effect. In the direction of preparing the background, many frameworks processed the background by dividing the scene to grid view cells similar to how iBAT and iBOAT frameworks processed the scene [1][2]. On the other hand, *Brun, L* [3][4] divided the scene into zones using an algorithm is able to cluster a dataset from the trajectories. In details, the algorithm could summarize as follows: consider the entire scene as one zone and then divide the zone into *L* fixed number of zones by using the distribution of training set. After that, each zone will be represented by using statistical properties (mean, the major axis and covariance matrix).

For the secound phase, studies of detecting anomalous behavior which is related to the traffic are widely common, most of these studies have built their frameworks based on Hidden Markov Model (HMM) *Cai, Y* [6]. While other methods used graphs[4] or built other frameworks similar to *Zhang et al.* [2] when he built iBAT framework and after that iBOAT [1] framework. iBAT/iBOAT frameworks aim to discover anomalous driving patterns from taxis trajectories. In fact, iBOAT improved iBAT to be able to work in online environments and maximized the grid cell size in the step of the scene's preprocessing to ensure the accuracy, by using experimental sizes (250m x 250m). Beside of that, iBOAT used a function to solve the problem of low sampling rate and cells' gaps, by this function the algorithm considers the points of trajectories located in the neighbor cells as normal points. Moreover, iBOAT used an *adaptive working window* and hasPath method which is described in the following algorithm [1]:

$\chi \leftarrow \emptyset$ // initialization, $\chi$ is the set of anomalous points
$T_0 \leftarrow T$ // $T_0$ first trajectory, $T$ set of trajectories
$i \leftarrow 0$ // Position in incoming trajectory
$w \leftarrow \emptyset$ // Adaptive window from $t$
$score(0) \leftarrow 0$
**while** the testing trajectory is not completed **do**
$i \leftarrow i + 1$
$g_i = \rho(p_i)$
$w \leftarrow w \cdot g_i$
$support(i) = |hasPath(T_{i-1}, w)| / |T_{i-1}|$
// *hasPath* returns the set of trajectories that contain all of the points in $t$ in the correct order
$T_i \leftarrow hasPath(T_{i-1}, w)$ // working set reduced.
**if** $support(i) < \theta$ **then** // where $\theta$ is threshold.
$\chi \leftarrow \chi \cup p_i$
$T_i \leftarrow T$ // reset the working set
$w \leftarrow g_i$
**end if**
$score(i) = score(i - 1) + \sigma(support(i)) * dist(p_{i-1}, p_i)$ // $\sigma(x) = 1/1 + e^{\lambda(x-\theta)}$
**end while**

# 3 Proposed framework

The structure of the proposed framework contains two main phases, preprocessing and detecting abnormal trajectories. In this framework, besides of preparing the data, we are looking to develop a method for dividing the scene into zones based on a dataset of trajectories. Semantic locations from the third party should affect the weight of the places, there for, GPS points are near the places have a significant weight will consider as visited the places and could cluster better. This improvement should avoid the outliers (which exist because of inaccuracies in sampling the GPS) and increase the robustness of clustering of the trajectories.

The second phase of this framework aims to detect anomalous trajectories, using the *adaptive working window*. Updating maps could be one of the implementations of this framework, by discovering the new official roads based on the number of anomalous trajectories passing through the same path.

Most of the studies of the related art of the works were finished, and some datasets from many sources have been tested if they could be suitable for this study. After that, an interesting dataset has been selected Berlin Moving Object Data (BerlinMOD[1]) for two days as *Figure 1* shows; the data has been preprocessed and imported in PostgreSQL server. Moreover, an interesting daily updated information about Berlin City offered by Geofabrik – German[2] has been imported to the PostgreSQL server and presented by ArcMap system. And by using filters of the building in ArcMap, and by

[1] http://dna.fernuni-hagen.de/secondo/BerlinMOD/BerlinMOD.html
[2] http://download.geofabrik.de/europe/germany/berlin.html

using filters of the building in ArcMap, museums (as examples of important places which should have more weight) could be separated from other buildings.

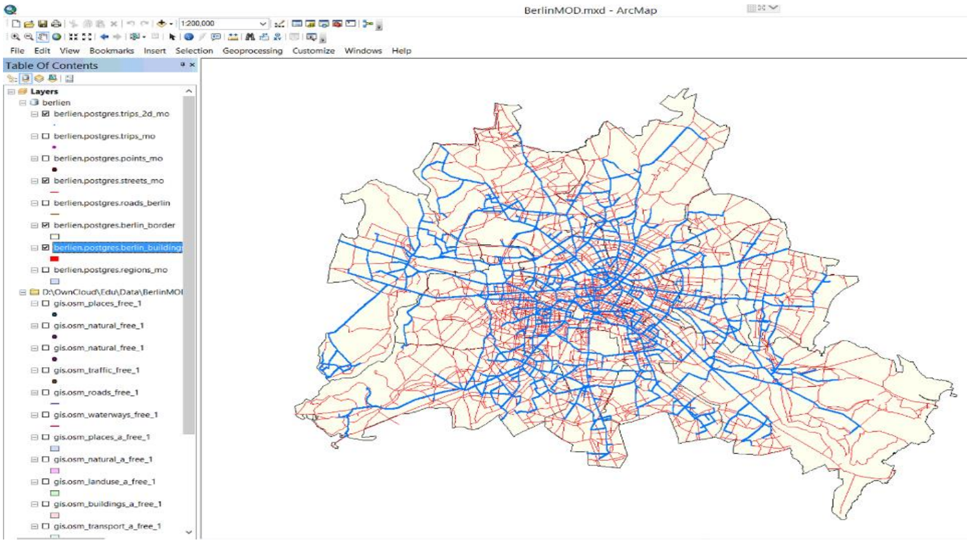


**Figure 1** BerlinMOD with Geofabrik data (Berlin)

In addition, iBOAT has been implemented by Koňárek, P [7], this implementation was developed and special functions were added to load and process BerlinMOD dataset which recorded two days from Berlin traffic. *Figure 2*. For that process the data. And according to the trajectories dimensions, the cell size is (37.08 x 37.08) in the metric system, and *theta* threshold is 0.005. Moreover, when trajectory with an *ID* is selected in the left-hand table, the anomalous trajectory is drawn in bold line on the map, and *iBOAT Score* chart draws where the trajectory act anomaly (fixed score values are normal sub-trajectories).

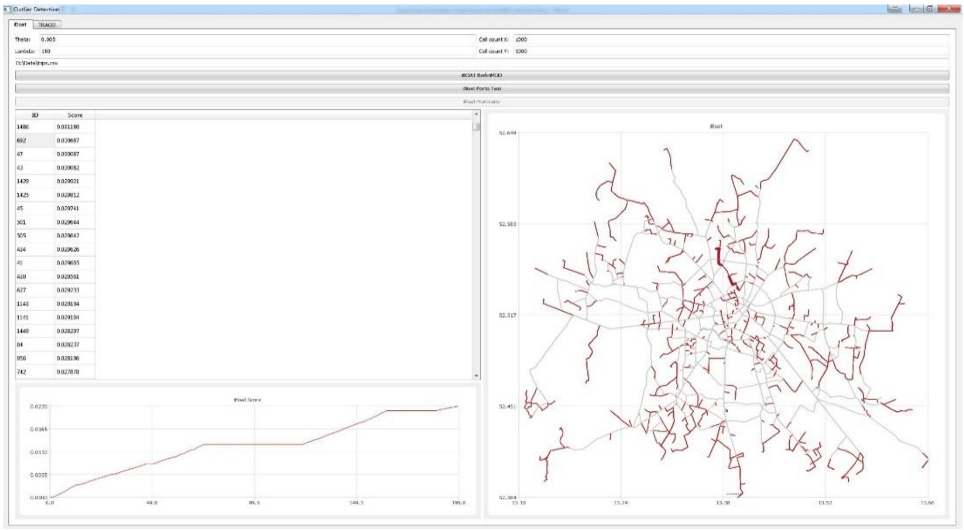

**Figure 2** Implementation of BerlinMOD dataset

The big cells' size will allow many trajectories to act anomalously without discovering that by the algorithm. On the other hand, the small cells' size means the algorithm will consume more computing time (more cells objects), and the Pos function will find that the new cell is not neighbor to any of the trajectory's cells and, as a result, the normal points will be detected as abnormal. For that, the cells size should be chosen manually based on the dataset dimensions. Furthermore, the threshold $\theta$ should not be big (no anomalous points will be selected) or too small (all the points will be recognized as anomalous trajectories).

## 4 Conclusion and future work

This paper described the most important issues related to the presented framework designed to detect traffic anomalous trajectories. In our framework, the development will affect all the levels to solve the problems of the outliers which exist because the low sampling rates and will change the results to be closer to the lifestyle.

Finally, information extracted from the web for rating the attraction places in Berlin will be used in the next step to affect the zoning process. Moreover, update the maps, or detecting the importance of locations based on detecting anomalous trajectories, all of that should be implemented by this framework.

## References:

1. Chen, C., Zhang, D., Castro, P. S., Li, N., Sun, L., Li, S., & Wang, Z. (2013). iBOAT: Isolation-based online anomalous trajectory detection. IEEE Transactions on Intelligent Transportation Systems, 14(2), 806–818. https://doi.org/10.1109/TITS.2013.2238531

2. Zhang, D., Li, N., Zhou, Z.-H., Chen, C., Sun, L., & Li, S. (2011). iBAT: Detecting anomalous taxi trajectories from GPS traces. Proceedings of the 13th International Conference on Ubiquitous Computing, 99–108. https://doi.org/10.1145/2030112.2030127
3. Brun, L., Capellini, B., Saggese, A., & Vento, M. (2014). Detection of anomalous driving behaviors by unsupervised learning of graphs. 11th IEEE International Conference on Advanced Video and Signal-Based Surveillance, AVSS 2014, 405–410. https://doi.org/10.1109/AVSS.2014.6918702
4. Brun, L., Saggese, A., & Vento, M. (2014). Dynamic Scene Understanding for Behavior Analysis Based on String Kernels, 24(10), 1669–1681.
5. Zheng, Y. (2011). Compute with spatial trajectories, ISBN: 9781461416289. Chapter 1.
6. Cai, Y., Wang, H., Chen, X., & Jiang, H. (2015). Trajectory-based anomalous behavior detection for intelligent traffic surveillance. Iet Intelligent Transport Systems, 9(8), 810–816. https://doi.org/10.1049/iet-its.2014.0238
7. Koňárek, P., (2017). Graduation thesis "Mining Anomalous Behavior in Trajectory Data." Faculty of Information Technology, BUT.

**Acknowledgments:** The research is supported by project "ICT tools, methods, and technologies to support smart cities" FIT-S-17-3964. And data used in the research was provided by Secondo BerlinMOD and Geofabrik – German for "Geo factory" projects.