

Extracting Trajectories through an Efficient and Unifying Spatio-Temporal Pattern Mining System

Phan Nhat Hai^{1,2}, Dino Ienco^{1,2}, Pascal Poncelet^{1,2}, and Maguelonne Teisseire^{1,2}

¹ IRSTEA Montpellier, UMR TETIS - 34093 Montpellier, France
{nhat-hai.phan,dino.ienco,maguelonne.teisseire}@teledetection.fr

² LIRMM CNRS Montpellier - 34090 Montpellier, France
pascal.poncelet@lirmm.fr

Abstract. Recent improvements in positioning technology has led to a much wider availability of massive moving object data. A crucial task is to find the moving objects that travel together. Usually, these object sets are called spatio-temporal patterns. Analyzing such data has been applied in many real world applications, e.g., in ecological study, vehicle control, mobile communication management, etc. However, few tools are available for flexible and scalable analysis of massive scale moving objects. Additionally, there is no framework devoted to efficiently manage multiple kinds of patterns at the same time. Motivated by this issue, we propose a framework, named GET_MOVE, which is designed to extract and manage different kinds of spatio-temporal patterns concurrently. A user-friendly interface is provided to facilitate interactive exploration of mining results. Since GET_MOVE is tested on many kinds of real data sets, it will benefit users to carry out versatile analysis on these kinds of data by exhibiting different kinds of patterns efficiently.

1 Introduction

Nowadays, many electronic devices are used for real world applications. Telemetry attached on wildlife, GPS installed in cars, sensor networks, and mobile phones have enabled the tracking of almost any kind of moving object data and has led to an increasingly large amount of data. Therefore, analysis on such data to find interesting patterns, called spatio-temporal patterns, is attracting increasing attention for applications such as movement pattern analysis, animal behavior study, route planning and vehicle control.

Despite the growing demands for diverse applications, there have been few scalable tools for mining massive and sophisticated moving object data. Even if some tools are available for extracting patterns (e.g. [4]), they mainly focus on specific kinds of patterns at a time. Obviously, when considering a dataset, it is quite difficult, for the decision maker, to know in advance which kinds of patterns are embedded in the data. To cope with this issue, we propose the GET_MOVE

system to reveal, automatically and in a very efficient way, collective movement patterns like convoys [1], group patterns [5], closed swarms [3], moving clusters [2] and also periodic patterns [7]. Starting from the results of GET_MOVE, the user can then visualize, browse and compare the different extracted patterns through a user friendly interface.

2 Spatio-Temporal Patterns

For clarity sake, we briefly remind the spatio-temporal pattern definitions. Informally, a *swarm* is a group of moving objects O containing at least min_o individuals which are closed each other for at least min_t timestamps T . To avoid redundant swarms, Zhenhui Li et al. [3] propose the notion of *closed swarm* for grouping together both objects and time. A swarm (O, T) is a closed swarm if it cannot be enlarged in terms of timestamps T and objects O . Another pattern is *convoy* which is also a group of objects such that these objects are closed each other during at least min_t time points. The main difference between convoy and swarm (or closed swarm) is that convoy lifetimes must be consecutive. Furthermore, moving clusters can be seen as special cases of convoys with the additional condition that they need to share some objects between two consecutive timestamps [6]. We can consider that the main difference between convoys and swarms is about the consecutiveness and non-consecutiveness of clusters during a time interval. In [5], Hwang et al. propose a general pattern, called a *group pattern*, which essentially is a combination of both convoys and closed swarms. Basically, group pattern is a set of disjointed convoys which are generated by the same group of objects in different time intervals. By considering a convoy as a time point, a group pattern can be seen as a closed swarm of disjointed convoys.

3 The GET_MOVE System Architecture

The GET_MOVE general architecture, described in Figure 1, has three main layers: (i) collection and cleaning, (ii) mining, and (iii) visualization. The bottom layer is responsible for collecting and preprocessing moving objects. During this step, some cleaning and interpolations techniques are used to integrate and clean the raw data as well as to interpolate missing points. The interpolation techniques used are similar to the ones provided by most of spatio-temporal pattern mining algorithms.

GET_MOVE uses a new mining algorithm able to exploit the similar characteristics among different kinds of patterns. This new mining method combines both clustering and pattern mining to extract the final results. So, by applying it to the preprocessed data, GET_MOVE can automatically extract different kinds of patterns such as convoys, closed swarms, group patterns, moving clusters and periodic patterns.

The top layer is devoted to the visualization. GET_MOVE provides a platform for users to flexibly tune parameters and supports visualization of the results in

different formats. The output can be written in Google Map³ and Google Earth⁴ formats to help users better explore the results. Furthermore, the system enables users to explore, plot and navigate over the different patterns in order to compare the differences among the various moving objects behaviors.

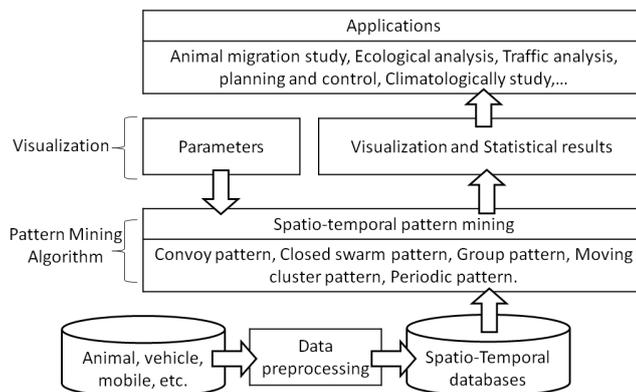


Fig. 1. The GET_MOVE System Architecture.

Figure 2 shows a screenshot of GET_MOVE⁵. It allows to choose a dataset and set the values of the parameter min_o and min_t . In this example we have chosen the *Buffalo* dataset from the Movebank⁶ project with $min_o = 10$ and $min_t = 10$ (combo box in the top of the figure). This dataset concerns the raw trajectories of 165 Buffaloes gathered from year 2000 to year 2006. In the second combo box in the figure, we can observe that, with these parameters, there are 908 closed swarms, 10 convoys and 10 group patterns. In this example, we observe that the execution time for extracting all these patterns is 0.52 seconds. From the combo box on right, the user can thus select a specific pattern and plot it on the main window. For instance, in our example, the user has decided to plot the 811th Closed Swarm. Finally the user can have more information such as pattern identifier, animal name and timespan when clicking on plotted trajectories as illustrated in the main window.

4 Conclusion

In this paper, we propose a system, GET_MOVE, which is designed to automatically and efficiently extract different kinds of spatio-temporal patterns at the same time. Starting from these results, the user can browse, navigate and

³ <http://code.google.com/apis/maps/>

⁴ <http://earth.google.com/>

⁵ <http://www.lirmm.fr/~phan/index.jsp>

⁶ <http://www.movebank.org/>

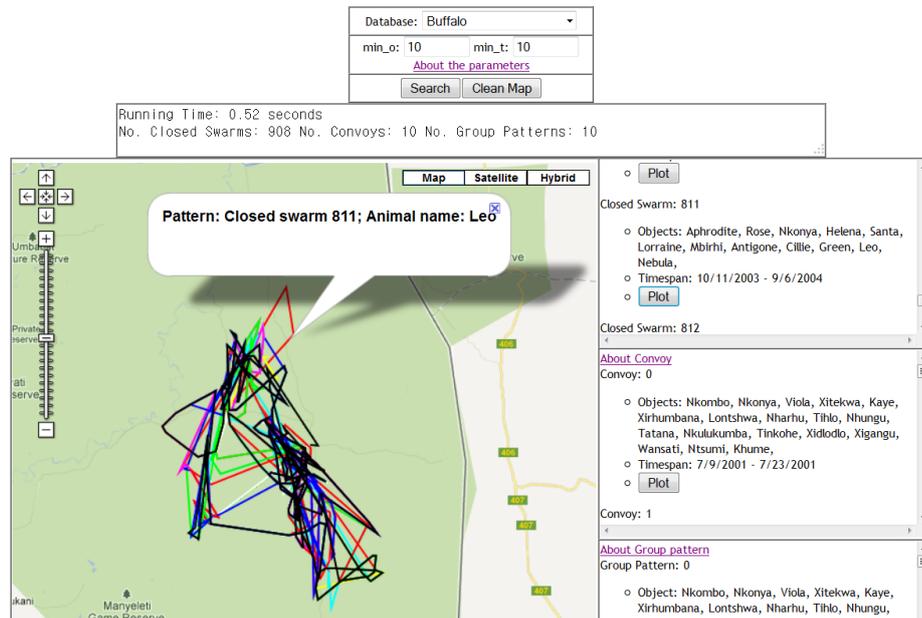


Fig. 2. The Graphical Visualization Interface.

compare different patterns in an easy way. The comparative analysis allows the user to understand and discover which results can fit better her researches.

References

1. Jeung H, Yiu ML, Zhou X, Jensen CS, Shen HT. *Discovery of Convoys in Trajectory Databases*. PVLDB 2008, 1(1):1068-1080.
2. P. Kalnis, N. Mamoulis, S. Bakiras. *On Discovering Moving Clusters in Spatio-temporal Data*. In SSTD 2005, Angra dos Reis, Brazil, pages 364-381.
3. Z. Li, B. Ding, J. Han, R. Kays. *Swarm: Mining Relaxed Temporal Moving Object Clusters*. In VLDB 2010, Singapore, pp. 723-734.
4. Z. Li, M. Ji, J.-G. Lee, L. Tang, Y. Yu, J. Han, and R. Kays. *Movemine: Mining moving object databases*. In SIGMOD 2010, Indianapolis, Indiana, pp.1203-1206.
5. Y. Wang, E.-P. Lim, and S.-Y. Hwang. *Efficient Mining of Group Patterns from User Movement Data*. In DKE(2006), pp. 240-282.
6. J. Han, Z. Li, and L. A. Tang. *Mining Moving Object, Trajectory and Traffic Data*. In DASFAA'10, Tsukuba, Japan.
7. N. Mamoulis, H. Cao, G. Kollios, M. Hadjieleftheriou, Y. Tao, D. W. Cheung. *Mining, Indexing, and Querying Historical Spatiotemporal Data*. In SIGKDD'04, pp.236-245.