GeT_Move: An Efficient and Unifying Spatio-Temporal Pattern Mining System for Moving Objects

Phan Nhat Hai, Dino Ienco, Pascal Poncelet, and Maguelonne Teisseire

LIRMM - TETIS Laboratory, University Montpellier 2, France
{nhat-hai.phan,dino.ienco,maguelonne.teisseire}@teledetection.fr
{pascal.poncelet}@lirmm.fr
http://www2.lirmm.fr/TATOO/spip.php

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 data mining tools are available for flexible and scalable analysis of massive scale moving object data. The main reason is that there is no a unifying approach to manage the patterns while many different kinds of spatio-temporal patterns have been proposed in recent years. Each approach only focuses on mining a specific kind of pattern. Our system, *GeT_Move*, is designed to extract and manage different 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 at the same time.

Keywords: Spatio-temporal patterns, trajectories, visualization

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 data and has led to an increasingly large amount of data that contain moving objects and numerical data. Therefore, analysis on such data to find interesting 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 available for mining massive and sophisticated moving object data. Even if some tools are available for extracting different kinds of patterns, they suffer the problem that they only extract a specific kind of pattern at a moment.

When considering a dataset, it is quite difficult, for the decision maker, to know in advance the kind of pattern embedded in the data. Additionally, it is difficult to recognize the differences between patterns. Therefore proposing an system able to automatically extract all these different kinds of patterns can be very useful. To tackle the issue, our system, GeT_Move is designed to reveal collective movement patterns like convoys [1], group patterns [5], closed swarms [3], moving clusters [2] and also periodic patterns [7]. GeT_Move has been applied on many different datasets published by the Movebank[1] project.

2 The GeT_Move System Architecture

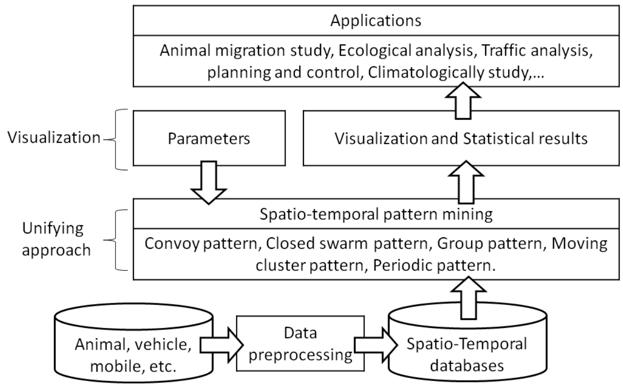
GeT_Move general architecture[2], described in Figure 1-a, has three main layers: (i) collection and cleaning, (ii) mining, and (iii) visualization. The bottom layer is responsible for collecting and preprocessing of moving objects. Object movements are quite complex and therefore data can be noisy. So preprocessing is needed to integrate and clean the raw data and to interpolate missing points.

Unifying approach is then applied on the preprocessed data on the middle layer. Then GeT_Move can automatically extract different kinds of patterns such as convoys, closed swarms, group patterns, moving clusters and periodic patterns. The brief definitions of such patterns and an overview of GeT_Move are presented in the following.

The top layer shows the visualized results with some statistics. They can be plotted on 2D plane or embedded into other visualization tools (e.g., Google Map[3] and Google Earth[4]). Our system enable users to visualize all the patterns with some statistics. Moreover, different patterns can be visualized in the same time on the same map. Consequently, users can be more insights into these results.

Informally, a *swarm* is a group of moving objects O containing at least ε 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,

[1] http://www.movebank.org/
[2] http://www.lirmm.fr/∼phan/index.jsp
[3] http://code.google.com/apis/maps/
[4] http://earth.google.com/



(a) General system

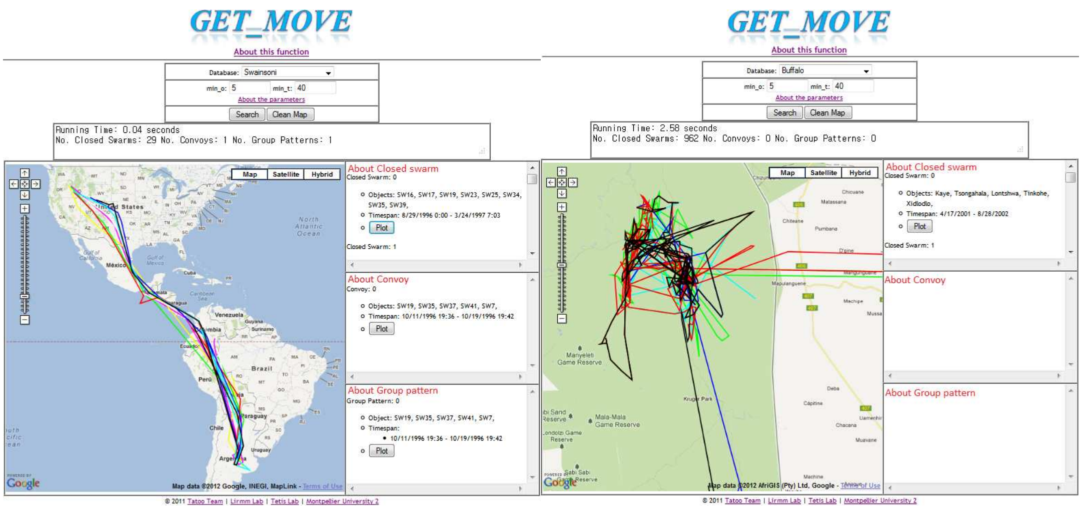
Fig. 1. General System.

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.

By taking into account the similar characteristics of the patterns, our system is able to extract all the patterns. Additionally, 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 Map and Google Earth format to help users better explore the results. Furthermore, users can compare the results in different datasets to understand the differences between object movement behaviors. Moreover, thanks to the system, one can explore that which kinds of patterns are embedded in each dataset. Moreover, users can plot as many patterns as users would like at the same time or clean and redraw the map with the other patterns. Additionally, to due with the large scale moving object databases, GeT_Move is implemented by combining different techniques such as clustering technique and pattern mining technique so that GeT_Move is very efficient. Therefore, the system can performed online and therefore it is easier for interested one to use and discover interesting patterns.

Let us examine a real example of Swainsoni and Buffalo movements. The raw trajectories of 43 Swainsonies are gathered from July 1995 to June 1998. Buffalo dataset concerns 165 buffalos and the tracking time is from year 2000 to year 2006. We further select the parameters ε and min_t and then the system automatically discover different patterns. See Figure 2, there are two closed swarms discovered from Swainsoni and Buffalo. Additionally, with the same parameters there are closed swarms, convoys and group patterns in Swainsoni while there are only closed swarms in Buffalo. Therefore, it is interesting for users to recognize the differences between Swainsonies and Buffalos.

The GeT_Move effectiveness and efficiency have been evaluated by using real and synthetic datasets. All the experimental results are also integrated on our demo website.



(a) One of discovered closed swarms in Swainsoni (b) One of discovered closed swarms in Buffalo

Fig. 2. Screen shots of GeT_Move.

3 Outlook

In this paper, we propose a system, GeT_Move, which is designed to automatically extract different kinds of spatio-temporal patterns by applying a unifying system. By experimenting with real data sets, one can efficiently observe interesting patterns, and also find the limitations of the current kinds of patterns. This is useful to promote further research into the new challenge issues in moving object mining.

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*. VLDB2010, 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. SIGKDD'04, pp.236-245.