ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

A Survey on a Uniting Framework of Mining Trajectory Forms of various Secular Stiffness

S.Devanathan[1], M.Mohan Kumar[2]

PG Scholar, Dept. of CSE., Sri Vidya College of Engineering & Technology, Virudhunagar, Tamil Nadu, India [1]

Assistant Professor, Dept. of CSE., Sri Vidya College of Engineering & Technology, Virudhunagar, Tamil Nadu, India [2]

ABSTRACT: Trajectory patterns discovery is more useful in learning interactions between moving objects. The trajectory patterns are arranged in the order of temporal tightness. A set of trajectory patterns are discovered and their granularity are adjusted by split and merge to detect other types. Trajectory classification is the model construction for predicting the class labels of moving objects based on their trajectories. Trajectory outlier detection detects outlying line segments for trajectory outliers and partitions a trajectory into a set of line segments. The classification process is based upon the fuzzy logic by using Naives byes classification. Groups of moving objects that have followed similar movements in different times are proposed by the Time Relaxed Spatiotemporal Trajectory Join. Unifying trajectory patterns (UT-patterns) of various temporal tightness is proposed in this paper. To show various patterns a pattern forest is constructed .For example, while a group of animals are migrating, some animals may leave the group or new animals may enter it so that a formal definition for moving clusters is provided.

KEYWORDS: Trajectory, fuzzy, temporal tightness

I. INTRODUCTION

Data mining is the process of extracting data from the hidden knowledge or data. Data mining is the process of finding correlations or patterns among dozens of fields in large relational databases. Data mining functionalities are characterization, discrimination, association, classification, clustering, outlier and trend analysis, etc. Knowledge discovery in data mining is the technology to extract interesting and unknown patterns from previous data. A KDD process includes data cleaning, data integration, data selection, transformation, data mining, pattern evaluation, and knowledge presentation. The advancement in location acquisition has generated massive spatial trajectory data that represent the mobility of a diversity of moving objects, such as people, vehicles, and animals.

A set of moving objects that are closely related in terms of location, time, or both is defined as UT patterns. A set of moving objects follow the same path every year is said to asynchronous in moving patterns. UT patterns mining are very useful in learning interactions between objects and group dynamics. Trajectory classification is the model construction for predicting the class labels of moving objects based on their trajectories. The UT patterns proposed is classified into three types based on the temporal tightness a) time controlled patterns b) time relaxed patterns c) time independent patterns. Through the discovery of patterns moving point object data can be analyzed. The main purpose of spatio-temporal data mining is to analyze data sets for interesting patterns. Some examples are moose in Sweden (25 animals reported every 30 minutes), leopards in South Africa (32 animals reported daily), and mountain goats in USA (32 animals reported every 3 hours). The classification presentation is given by decision-tree, classification rule and neural network. Prediction is the process of some unknown or missing numerical values .Supervised learning is the basic process for classification. In *supervised* learning (classification)

- The labels indicating the class of the observations are accompanied by the training data (observations, measurements, etc.).
- Based on the training set the new data is classified.

In *unsupervised* learning (clustering)

- The training data class label is unknown

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

- The training data set establishes the existence of classes or clusters in the data.

The pattern classification and clustering process is done by trajectory forms and the hierarchical diagram is given in Fig 1

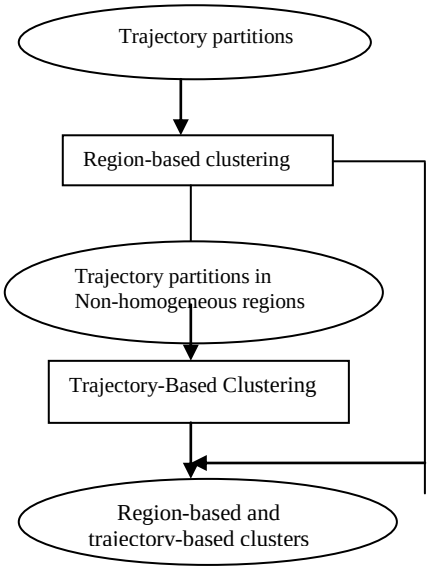

Fig 1.1: The procedure of hierarchical region-based and trajectory-based clustering

Trajectory partitioning and trajectory-based clustering are based on the partition-and-group framework. It consists of two phases a) the partitioning phase b) the grouping phase. Also the flock pattern is derived which is a group of objects that move together within a disc of some user-specified size. The description of a trajectory outlier should be unbiased to applications. The trajectory outlier is detected by two types

- **Positional outlier:** The *position* of a trajectory is different from those of neighboring trajectories.
- **Angular outlier :** The *path* of a trajectory is different from those of neighboring trajectories.

Next the trajectory clustering process is handled. A set of trajectory partitions is called as cluster. A line segment is the trajectory partitioning. According to the distance measure line segments that belong to the same cluster are close to each other. A trajectory can belong to several clusters because a trajectory is partitioned into multiple line segments, and clustering is performed over these line segments. A delegate trajectory is a series of points just like a common trajectory. It is an invented trajectory that indicates the major performance of the trajectory partitions (*i.e.* line segments) that belong to the cluster. In cluster analysis class label is unknown. There is a of group data to form new classes, e.g., cluster houses to find distribution patterns. The principle in clustering is based on maximizing the intra-class similarity and minimizing the interclass similarity. In outlier analysis, the outlier is a data object that does not observe with the general behavior of the data. It can be considered as noise or exception but it is useful in fraud detection, rare events analysis. The following Figure 2 shows the clusters and outliers analysis.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

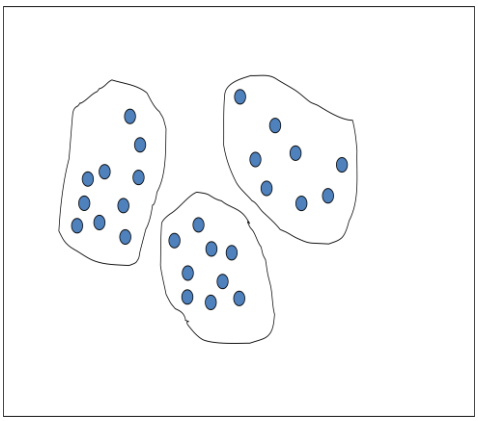

Fig 1.2: Clusters and Outliers

The trajectories are classified into pattern by Association rule mining that includes the process of

- Finding common patterns, relations, correlations, or fundamental structures among sets of items or objects in transaction databases, relational databases, and other information repositories.
- Frequent pattern is a set of items, sequence, etc. that occurs frequently in a database

The different classification methods are

- Decision tree based methods
- Rule-based methods
- Memory based reasoning
- Neural networks
- Naives Bayes and Bayesian belief networks
- Support vector machines

In this survey the classification is based upon the Naives Bayes classification using fuzzy logic. The clustering process is based upon two types

- Partition clustering
- Hierarchical clustering

II. RELATED WORK

Jae Gil Lee et.al, [1] proposed the classification of trajectories that is based on the model construction for predicting the class labels of moving objects and other features. This survey is based upon the shape trajectory which appears to the part of the trajectory or not the relevant to that shape of the trajectory. It is based upon the geographical areas where trajectories are spread along. A *traclass* framework has been proposed which generates the hierarchy for partitioning the trajectories. There are two types of clustering process being proposed 1) region-based 2) trajectory-based. Also they have proposed two real world applications 1) Vessel classification from satellite images 2) Classification of trace gas measurements. The clustering process in this paper has two types 1) region-based clustering 2) trajectory based. The algorithm being proposed here is based upon the region based and trajectory based clustering. Also the Recursive Quantization and Merging problem is being handled in this survey. The animal movement data is generated and the region-based and trajectory-based. In the training set the region based clustering is efficient than the trajectory based which is more expensive.

Kyu-Young Whang et.al, [2] proposed the clustering process algorithm that overcomes the existing systems which group existing trajectory as a whole. Also a partition-and-group framework for trajectories has been proposed that

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

makes the trajectories into line segments. The main benefit of this survey is to build a framework for sub-trajectories for trajectory database. Based on this framework the clustering algorithm called the *TRACLUS* has been proposed. There are two phases in the algorithm 1) partitioning 2) grouping.
The partition algorithm uses the minimum description length [MDL]. The real time examples include the vehicle position data, hurricane track data etc. The key observation of the clustering trajectories is not to detect the similar portions of trajectories. The proposed work presents an efficient density-based clustering algorithm for line segments. In this survey the heuristics for parameter value has been experimented to show the optimal value for the accurate trajectories. The MDL algorithm is extended to support the trajectories with weights. The MDL algorithm proposed here is used for many applications like graph partitioning and distance function design for strings. The various extensions of the trajectory clustering proposed here are

- Extensibility
- Parameter Insensitivity
- Efficiency
- Movement Patterns
- Temporal Information

Jiawei Han et.al [3] proposed the outlier detection by two ways of detection 1) positional outlier 2) angular outlier. In this survey a partition-and-detect framework for trajectory outlier detection is proposed. This partition-and-detect framework partitions a trajectory into a set of line segments, and then, detects outlying line segments for trajectory outliers. A trajectory outlier detection algorithm TRAOD is proposed in this survey that contains two phases 1) partitioning 2) detection. In the partitioning a two-level trajectory partitioning
Strategy has been proposed. It ensures both high quality and high efficiency. Then in the detection phase a hybrid of the distance-based and density-based approaches has been proposed. A data object that is grossly different from or inconsistent with the remaining set of data is an outlier. There are different outlier algorithms proposed in the survey and they are 1) distribution-based, 2) distance-based, 3) density based 4) and deviation based algorithms. A decision-tree-based symbolic
rule induction system has been proposed. A fast decision tree induction algorithm has been proposed that is especially useful for detecting outliers and trajectories. The key observation of this survey is with the help from the summary information; the trajectories are compared as a whole that might not be able to detect outlying portions of the trajectories.

Panos Kalnis et.al [4] calculated the moving clusters with help of spatio-temporal data. A set of objects that move close to each other for a long time interval is defined by a moving cluster. The availability of accurate location information from embedded GPS devices will enable in the near future numerous novel applications requiring the extraction of moving clusters. The survey proposes three algorithms 1) MC1 the straight forward approach 2) MC2 minimizing redundant checks 3) MC3 approximate moving clusters. A generator which generates synthetic datasets with various distributions is developed in this survey. The output of the generator is a series of time slices. MC1 combines the results into a set of moving clusters by performing spatial clustering at each snapshot. In MC2 the process is speed-up by pruning the object combinations. In MC3, to predict the set of clusters at the current snapshot uses information from the past is used. In clustering the static spatial data the partitioning is done by k-mediods that divides the objects into k groups and iteratively exchange objects between them until the quality of the clusters does not further improve. The k mediods are chosen randomly from the dataset. Hierarchical clustering techniques define the clusters in a bottom-up fashion. Density-based methods find out dense regions in space, where objects are secure to each other and divide them from regions of low density.
Dimitris Sacharidis et.al [5] considered frequent of moving objects prepared with location-sensing devices and accomplished of communicating with a central coordinator. The hot motion paths are the routes often followed by numerous objects over the recent past. The applications requiring classification/profiling discovers the hot motion path. The classification is based on monitored movement patterns, such as targeted advertising, resource allocation, etc. In this survey the problem of overcoming the maintenance of hot motion paths is proposed. The location-aware device (GPS, RFID, etc) monitors the moving objects and motion patterns. The algorithm proposed here is the *Ray trace algorithm* that acts as a filter and maintains the spatiotemporal extent. It is termed as Spatial Safe Area (SSA), which is

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

around the object's trajectory. To maintain the count for each motion path, hash table and an event queue is used. The proposed frame work has a coordinator point that maintains the hotness and geometries for the spatiotemporal index.

Marc Benkert et.al [6] represents the data in moving objects especially in the area of wildlife GPS tracking. The common spatio-temporal pattern to detect this is *flocks*. For a certain pre-defined time the objects moving along paths close to each other in a large enough subset is called as flock. In the propose work, the REMO framework (Relative Motion) has been proposed. This framework defines a collection of spatio-temporal patterns based on similar direction or motion path. The algorithm proposed in the work has been performed by a set of range counting queries with higher dimensional space. There are several data structures supporting these queries. For the proposed work skip-quad tree or BBD trees has been implemented.

Hoyoung Jeung et.al [7] formalizes the concept of convoys using density-based notions. This survey is done so to capture group of arbitrary extents and shapes. In the proposed work, there are three algorithms proposed for convoy discovery that adopt the well known filter-refinement framework. The line simplification techniques on the trajectories are applied in the filter step. This is done to establish the distance bound between the simplified trajectories. The candidate convoys are further processed to obtain the actual convoys in the refinement step. The lossy-flock problem has been overcome in this survey. The first process in discovering the convoys is to discover the cluster. The second step defines the CuTS (Convoy Discovery Using Trajectory Classification). This is done in the filter-refinement framework which is a set of candidate convoys are retrieved in the filter step. The next step is CuTS+ which accelerates the process of trajectory simplification. The last step is the CuTS* that enhances the effectiveness of the filter step by introducing the tighter distance bound of the simplified trajectory.

Joachim Gudmundsson et.al [8] computes the computational efficiency of two most basic spatio-temporal patterns in trajectories, namely flocks and meetings. For the longest duration meeting flocks the polynomial time algorithm has been proposed. The longest duration fixed meeting algorithm determines the edges and boundaries of the trajectories. The next solved algorithm in this work is the longest duration varying meeting. The main difference is maintaining all connected components of trajectories inside the column or half-column.

III. PROPOSED WORK

A uniting setting work of mining trajectory forms of various secular-stiffness called uniting trajectory forms (UT-forms) is proposed in this survey. The setting work consists of two phases: original forms detection and granularity alteration. In the first phase a set of original forms are exposed. In the second phase their granularity levels (i.e. level of details) are adjusted. The classification process in this phase is proposed with Naives-Bayes classification algorithm using fuzzy logic. The pattern forest is constructed with the results that are obtained. The trajectory patterns of various temporal tightness: time-constrained, time-relaxed, and time-independent has been evolved.

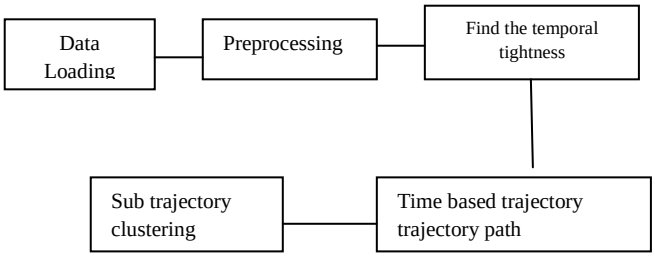

Fig 3.1: The flow diagram

The flow diagram explains about the extension of the proposed work. The different work to be undertaken are collecting of set of trajectories, discovering initial pattern Generating initial pattern granularity adjustment and discovering UT patterns. Merging border work of work of mining trajectory patterns of various secular tightness is the important in this proposed work.

ISSN(Online): 2320-9801
ISSN (Print): 2320-9798

International Journal of Innovative Research in Computer and Communication Engineering

(An ISO 3297: 2007 Certified Organization)

Vol. 3, Issue 11, November 2015

IV. CONCLUSION AND FUTURE WORK

In this paper a unifying framework for mining UT-patterns has been proposed. A pattern mining algorithm UT-Pattern Mine has been developed. The detection of the trajectory patterns of various temporal tightness like time-constrained, time-relaxed, and time-independent is the main advantage. A pattern forest by drill-down and roll-up is constructed. The algorithm classifies the pattern based on the Naives-Bayes classification. It is implemented in fuzzy where it takes the Type 2 fuzzy set. The accuracy is shown to be very high (80~90) for the moving objects. A Framework for trajectories has been proposed in this paper that performs hierarchical region-based and trajectory-based clustering after trajectory partitioning. The high classification accuracy owing to the collaboration between the two types of clustering is the key advantage.

V. ACKNOWLEDGEMENT

This research was supported by **Mr. M. Mohan Kumar (Assistant Professor of CSE) and Dr. P. Murugeswari., Ph.D.,** who provided insight and expertise that greatly, assisted the survey and the comments that greatly improved the manuscript. I also like to show my gratitude to them for sharing their pearls of wisdom with me during the survey.

REFERENCES

[1] JaeGil Lee, Jiawei Han, Xiao lei Li, and Hector Gonzalez "Traclass: Trajectory Classification Using Hierarchical Region-Based and Trajectory-Based Clustering" DEC 2008.
[2] Jae-Gil Lee, Jiawei Han Kyu-Young Whang "Trajectory Clustering: A Partition-and-Group Framework" 2008.
[3] Jae-Gil Lee, Jiawei Han, and Xiao lei Li: "Trajectory Outlier Detection: A Partition-and-Detect Framework" 2009.
[4] Panos Kalnis, Nikos Mamoulis, and Spiridon Bakiras "On Discovering Moving Clusters in Spatio-temporal Data" 2010
[5] Dimitris Sacharidis, Kostas Patroumpas, Manolis Terrovitis: "On-Line Discovery of Hot Motion Paths "at 2010.
[6] Marc Benkert, Joachim Gudmundsson, Florian Hubner, and Thomas Wolle "Reporting Flock Patterns" at 2010
[7] Hoyoung Jeungy Man Lung Yiuz Xiao fang Zhouy Christian S. Jensenz Heng Tao Shen "Discovery of Convoys in Trajectory Databases"
[8] Joachim Gudmundsson, Marc van Kreveld "Computing Longest Duration Flocks in Trajectory Data" 2011
[9] P . B a k a l ov , M . Ha d j i e le f t he r i o u , a n d V. J . T s ot r as , "T i m e r e la x e d sp a t io t e m p o r a l t r a j e c t o r y j o in s ," in Pro c . 13th A CM Int. Symp. G eo graph. Inf. Sys t . , B remen, Germany , Nov . 2005, pp. 182–191.
[10] F. G ia n no t t i, M. Na n ni, F. P i n el l i, a and D. P e d r e s c hi, "T r a j e c -tor y pattern mining ,, " in Pro c. 13th AC M SIGKDD I nt. Conf. Knowledge Discovery Data Mining, San Jose, CA, USA, Aug. 2007, pp. 330–339
[11] P. Laube, S. Imfeld , and R. Weibel , "Discovering relative motion patterns in groups of moving point objects,"Int. J. Geograph . Inf. Sci.,vol.19,no.6,pp.639–668,Jul.2005.
[12] J.-G. Lee, J. Han, and X. Li, "Trajectory outlier detection: A partition and detect framework," in Proc. 24th Int. Conf. Data Eng. ,Cancun, Mexico, Apr. 2008, pp. 140–149.