INTERNATIONAL JOURNAL OF RESEARCH IN COMPUTER APPLICATION & MANAGEMENT

IJRCM

A Monthly Double-Blind Peer Reviewed (Refereed/Juried) Open Access International e-Journal - Included in the International Serial Directories

Indexed & Listed at:

Ulrich's Periodicals Directory ©, ProQuest, U.S.A., EBSCO Publishing, U.S.A., Cabell's Directories of Publishing Opportunities, U.S.A., Google Scholar,
Open J-Gage, India [link of the same is duly available at Inflibnet of University Grants Commission (U.G.C.)],
Index Copernicus Publishers Panel, Poland with IC Value of 5.09 & number of libraries all around the world.
Circulated all over the world & Google has verified that scholars of more than **5000 Cities** in **187 countries/territories** are visiting our journal on regular basis.
Ground Floor, Building No. 1041-C-1, Devi Bhawan Bazar, JAGADHRI – 135 003, Yamunanagar, Haryana, INDIA

http://ijrcm.org.in/

CONTENTS

CALL FOR MANUSCRIPTS

We invite unpublished novel, original, empirical and high quality research work pertaining to the recent developments & practices in the areas of Computer Science & Applications; Commerce; Business; Finance; Marketing; Human Resource Management; General Management; Banking; Economics; Tourism Administration & Management; Education; Law; Library & Information Science; Defence & Strategic Studies; Electronic Science; Corporate Governance; Industrial Relations; and emerging paradigms in allied subjects like Accounting; Accounting Information Systems; Accounting Theory & Practice; Auditing; Behavioral Accounting; Behavioral Economics; Corporate Finance; Cost Accounting; Econometrics; Economic Development; Economic History; Financial Institutions & Markets; Financial Services; Fiscal Policy; Government & Non Profit Accounting; Industrial Organization; International Economics & Trade; International Finance; Macro Economics; Micro Economics; Rural Economics; Co-operation; Demography: Development Planning; Development Studies; Applied Economics; Development Economics; Business Economics; Monetary Policy; Public Policy Economics; Real Estate; Regional Economics; Political Science; Continuing Education; Labour Welfare; Philosophy; Psychology; Sociology; Tax Accounting; Advertising & Promotion Management; Management Information Systems (MIS); Business Law; Public Responsibility & Ethics; Communication; Direct Marketing; E-Commerce; Global Business; Health Care Administration; Labour Relations & Human Resource Management; Marketing Research; Marketing Theory & Applications; Non-Profit Organizations; Office Administration/Management; Operations Research/Statistics; Organizational Behavior & Theory; Organizational Development; Production/Operations; International Relations; Human Rights & Duties; Public Administration; Population Studies; Purchasing/Materials Management; Retailing; Sales/Selling; Services; Small Business Entrepreneurship; Strategic Management Policy; Technology/Innovation; Tourism & Hospitality; Transportation Distribution; Algorithms; Artificial Intelligence; Compilers & Translation; Computer Aided Design (CAD); Computer Aided Manufacturing; Computer Graphics; Computer Organization & Architecture; Database Structures & Systems; Discrete Structures; Internet; Management Information Systems; Modeling & Simulation; Neural Systems/Neural Networks; Numerical Analysis/Scientific Computing; Object Oriented Programming; Operating Systems; Programming Languages; Robotics; Symbolic & Formal Logic; Web Design and emerging paradigms in allied subjects.

Anybody can submit the **soft copy** of unpublished novel; original; empirical and high quality **research work/manuscript** ***anytime*** in ***M.S. Word format*** after preparing the same as per our **GUIDELINES FOR SUBMISSION**; at our email address i.e. infoijrcm@gmail.com or online by clicking the link **online submission** as given on our website (***FOR ONLINE SUBMISSION, CLICK HERE***).

GUIDELINES FOR SUBMISSION OF MANUSCRIPT

1. **COVERING LETTER FOR SUBMISSION:**

DATED: _____________

THE EDITOR

IJRCM

Subject: **SUBMISSION OF MANUSCRIPT IN THE AREA OF ______________________.**

(e.g. Finance/Mkt./HRM/General Mgt./Engineering/Economics/Computer/IT/ Education/Psychology/Law/Math/other, please specify)

DEAR SIR/MADAM

Please find my submission of manuscript titled '____' for likely publication in one of your journals.

I hereby affirm that the contents of this manuscript are original. Furthermore, it has neither been published anywhere in any language fully or partly, nor it is under review for publication elsewhere.

I affirm that all the co-authors of this manuscript have seen the submitted version of the manuscript and have agreed to inclusion of their names as co-authors.

Also, if my/our manuscript is accepted, I agree to comply with the formalities as given on the website of the journal. The Journal has discretion to publish our contribution in any of its journals.

NAME OF CORRESPONDING AUTHOR	:
Designation/Post*	:
Institution/College/University with full address & Pin Code	:
Residential address with Pin Code	:
Mobile Number (s) with country ISD code	:
Is WhatsApp or Viber active on your above noted Mobile Number (Yes/No)	:
Landline Number (s) with country ISD code	:
E-mail Address	:
Alternate E-mail Address	:
Nationality	:

* i.e. Alumnus (Male Alumni), Alumna (Female Alumni), Student, Research Scholar (M. Phil), Research Scholar (Ph. D.), JRF, Research Assistant, Assistant Lecturer, Lecturer, Senior Lecturer, Junior Assistant Professor, Assistant Professor, Senior Assistant Professor, Co-ordinator, Reader, Associate Professor, Professor, Head, Vice-Principal, Dy. Director, Principal, Director, Dean, President, Vice Chancellor, Industry Designation **etc**. ***The qualification of author is not acceptable for the purpose***.

NOTES:

a) The whole manuscript has to be in ***ONE MS WORD FILE*** only, which will start from the covering letter, inside the manuscript. ***pdf. version* is liable to be rejected without any consideration**.

b) The sender is required to mention the following in the **SUBJECT COLUMN of the mail**:

New Manuscript for Review in the area of (e.g. Finance/Marketing/HRM/General Mgt./Engineering/Economics/Computer/IT/ Education/Psychology/Law/Math/other, please specify)

c) There is no need to give any text in the body of the mail, except the cases where the author wishes to give any **specific message** w.r.t. to the manuscript.

d) The total size of the file containing the manuscript is expected to be below **1000 KB**.

e) Only the **Abstract will not be considered for review** and the author is required to submit the **complete manuscript** in the first instance.

f) ***The journal gives acknowledgement w.r.t. the receipt of every email within twenty-four hours*** and in case of non-receipt of acknowledgment from the journal, w.r.t. the submission of the manuscript, within two days of its submission, the corresponding author is required to demand for the same by sending a separate mail to the journal.

g) The author (s) name or details should not appear anywhere on the body of the manuscript, except on the covering letter and the cover page of the manuscript, in the manner as mentioned in the guidelines.

2. **MANUSCRIPT TITLE:** The title of the paper should be typed in **bold letters**, **centered** and **fully capitalised**.
3. **AUTHOR NAME (S) & AFFILIATIONS:** Author (s) **name**, **designation**, **affiliation** (s), **address**, **mobile/landline number** (s), and **email/alternate email address** should be given underneath the title.
4. **ACKNOWLEDGMENTS:** Acknowledgements can be given to reviewers, guides, funding institutions, etc., if any.
5. **ABSTRACT:** Abstract should be in **fully Italic printing**, ranging between **150** to **300 words**. The abstract must be informative and elucidating the background, aims, methods, results & conclusion in a **SINGLE PARA**. ***Abbreviations must be mentioned in full***.
6. **KEYWORDS:** Abstract must be followed by a list of keywords, subject to the maximum of **five**. These should be arranged in alphabetic order separated by commas and full stop at the end. All words of the keywords, including the first one should be in small letters, except special words e.g. name of the Countries, abbreviations etc.
7. **JEL CODE:** Provide the appropriate Journal of Economic Literature Classification System code (s). JEL codes are available at www.aeaweb.org/econlit/jelCodes.php. However, mentioning of JEL Code is not mandatory.
8. **MANUSCRIPT:** Manuscript must be in ***BRITISH ENGLISH*** prepared on a standard A4 size ***PORTRAIT SETTING PAPER***. *It should be free from any errors i.e.* ***grammatical, spelling*** *or* ***punctuation.*** *It must be thoroughly edited at your end*.
9. **HEADINGS:** All the headings must be bold-faced, aligned left and fully capitalised. Leave a blank line before each heading.
10. **SUB-HEADINGS:** All the sub-headings must be bold-faced, aligned left and fully capitalised.
11. **MAIN TEXT:**

THE MAIN TEXT SHOULD FOLLOW THE FOLLOWING SEQUENCE:

INTRODUCTION

REVIEW OF LITERATURE

NEED/IMPORTANCE OF THE STUDY

STATEMENT OF THE PROBLEM

OBJECTIVES

HYPOTHESIS (ES)

RESEARCH METHODOLOGY

RESULTS & DISCUSSION

FINDINGS

RECOMMENDATIONS/SUGGESTIONS

CONCLUSIONS

LIMITATIONS

SCOPE FOR FURTHER RESEARCH

REFERENCES

APPENDIX/ANNEXURE

The manuscript should preferably be in ***2000** to **5000 WORDS***, But the limits can vary depending on the nature of the manuscript.

12. **FIGURES & TABLES:** These should be simple, crystal **CLEAR**, **centered**, **separately numbered** & self-explained, and the **titles must be above the table/figure**. **Sources of data should be mentioned below the table/figure**. *It should be ensured that the tables/figures are referred to from the main text*.
13. **EQUATIONS/FORMULAE:** These should be consecutively numbered in parenthesis, left aligned with equation/formulae number placed at the right. The equation editor provided with standard versions of Microsoft Word may be utilised. If any other equation editor is utilised, author must confirm that these equations may be viewed and edited in versions of Microsoft Office that does not have the editor.
14. **ACRONYMS:** These should not be used in the abstract. The use of acronyms is elsewhere is acceptable. Acronyms should be defined on its first use in each section e.g. Reserve Bank of India (RBI). Acronyms should be redefined on first use in subsequent sections.
15. **REFERENCES:** The list of all references should be alphabetically arranged. ***The author (s) should mention only the actually utilised references in the preparation of manuscript*** and they may follow Harvard Style of Referencing. **Also check to ensure that everything that you are including in the reference section is duly cited in the paper**. The author (s) are supposed to follow the references as per the following:

- All works cited in the text (including sources for tables and figures) should be listed alphabetically.
- Use (**ed.**) for one editor, and (**ed.s)** for multiple editors.
- When listing two or more works by one author, use --- (20xx), such as after Kohl (1997), use --- (2001), etc., in chronologically ascending order.
- Indicate (opening and closing) page numbers for articles in journals and for chapters in books.
- The title of books and journals should be in italic printing. Double quotation marks are used for titles of journal articles, book chapters, dissertations, reports, working papers, unpublished material, etc.
- For titles in a language other than English, provide an English translation in parenthesis.
- ***Headers**, **footers**, **endnotes** and **footnotes** should **not be used** in the document.* However, **you can mention short notes to elucidate some specific point**, which may be placed in number orders before the references.

PLEASE USE THE FOLLOWING FOR STYLE AND PUNCTUATION IN REFERENCES:

BOOKS

- Bowersox, Donald J., Closs, David J., (1996), "Logistical Management." Tata McGraw, Hill, New Delhi.
- Hunker, H.L. and A.J. Wright (1963), "Factors of Industrial Location in Ohio" Ohio State University, Nigeria.

CONTRIBUTIONS TO BOOKS

- Sharma T., Kwatra, G. (2008) Effectiveness of Social Advertising: A Study of Selected Campaigns, Corporate Social Responsibility, Edited by David Crowther & Nicholas Capaldi, Ashgate Research Companion to Corporate Social Responsibility, Chapter 15, pp 287-303.

JOURNAL AND OTHER ARTICLES

- Schemenner, R.W., Huber, J.C. and Cook, R.L. (1987), "Geographic Differences and the Location of New Manufacturing Facilities," Journal of Urban Economics, Vol. 21, No. 1, pp. 83-104.

CONFERENCE PAPERS

- Garg, Sambhav (2011): "Business Ethics" Paper presented at the Annual International Conference for the All India Management Association, New Delhi, India, 19–23

UNPUBLISHED DISSERTATIONS

- Kumar S. (2011): "Customer Value: A Comparative Study of Rural and Urban Customers," Thesis, Kurukshetra University, Kurukshetra.

ONLINE RESOURCES

- Always indicate the date that the source was accessed, as online resources are frequently updated or removed.

WEBSITES

- Garg, Bhavet (2011): Towards a New Gas Policy, Political Weekly, Viewed on January 01, 2012 http://epw.in/user/viewabstract.jsp

OVERVIEW OF TRAJECTORY DATA MINING AND THE TECHNIQUES USED

DR. R. KOUSALYA
PROFESSOR & HEAD
DEPARTMENT OF COMPUTER APPLICATIONS
DR. N. G. P. ARTS & SCIENCE COLLEGE
COIMBATORE

S. DHARANI
RESEARCH SCHOLAR
DEPARTMENT OF COMPUTER APPLICATIONS
DR. N. G. P. ARTS & SCIENCE COLLEGE
COIMBATORE

ABSTRACT

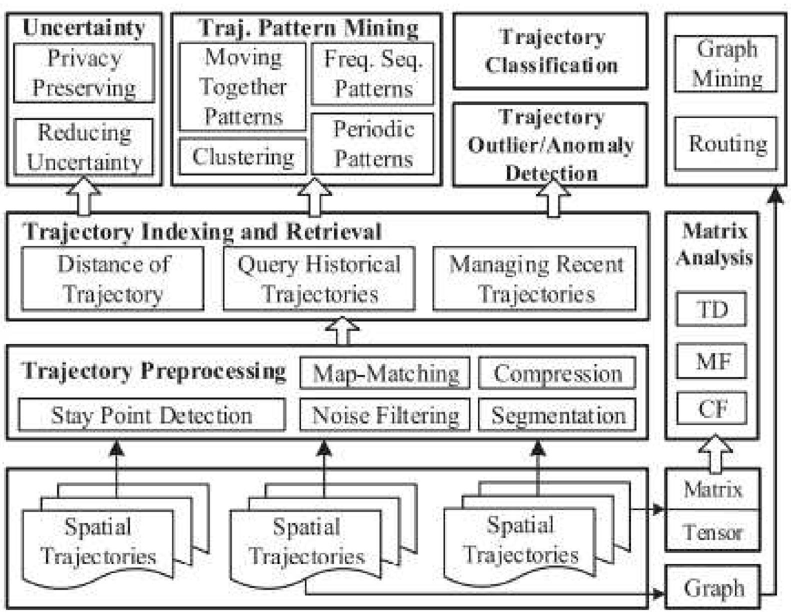
The advances in location-acquisition and mobile computing techniques have generated massive spatial trajectory data, which represent the mobility of a diversity of moving objects, such as people, vehicles, and animals. Many techniques have been proposed for processing, managing, and mining trajectory data in the past decade, fostering a broad range of applications. In this article, we present a systematic approach on the major research into trajectory data mining, providing a panorama of the field as well as the scope of its research topics. This paper also introduces the methods that transform trajectories into other data formats, such as graphs, matrices, and tensors, to which more data mining and machine learning techniques can be applied. Finally, some public trajectory datasets are presented.

KEYWORDS
stay point detection, noisy filtering, segmentation, compression, map matching, tensor, clustering, patterns.

1. INTRODUCTION

A spatial trajectory is a trace generated by a moving object in geographical spaces, usually represented by a series of chronologically ordered points, for example, p1 → p2 → ··· → pn, where each point consists of a geospatial coordinate set and a time stamp such as p = (x, y,t). The advance in location-acquisition technologies has generated a myriad of spatial trajectories representing the mobility of various moving objects, such as people, vehicles, and animals. Such trajectories offer us unprecedented information to understand moving objects and locations, fostering a broad range of applications in location-based social networks [1], intelligent transportation systems, and urban computing [2]. The prevalence of these applications in turn calls for systematic research on new computing technologies for discovering knowledge from trajectory data. Under the circumstances, trajectory data mining has become an increasingly important research theme, attracting the attention from numerous areas, including computer science, sociology, and geography.

FIG. 1: PARADIGM OF TRAJECTORY DATA MINING



Intensive and extensive individual research has been done in the field of trajectory data mining. However, we are lack of a systematic review that can well shape the field and position existing research. Facing a huge volume of publications, the community is still not very clear about the connections, correlations and difference among these existing techniques. To this end, a comprehensive survey that thoroughly explores the field of trajectory data mining, according to the paradigm is shown in Figure 1.

Before using trajectory data, we need to deal with a number of issues, such as noise filtering, segmentation, and map matching. This stage is called trajectory preprocessing, which is a fundamental step of many trajectory data mining tasks. The goal of noise filtering is to remove from a trajectory some noise points that may be caused by the poor signal of location positioning systems (e.g., when traveling in a city canyon). Trajectory compression is to compress the size of a trajectory while maintaining the utility of the trajectory. A stay point detection algorithm identifies the location where a moving object has stayed for a while within a certain distance threshold. A stay point could stand for a restaurant or a shopping mall that a user has been to, carrying more semantic meanings than other points in a trajectory. Trajectory segmentation divides a trajectory into fragments by time interval, spatial shape, or semantic meanings, for a further process like clustering and classification. Map matching aims to project each point of a trajectory onto a corresponding road segment where the point was truly generated.

Many online applications require instantly mining trajectory data (e.g., detecting traffic anomalies), calling for effective data management algorithms that can quickly retrieve particular trajectories satisfying certain criteria (such as spatiotemporal constraints) from a big trajectory corpus. There are usually two major types of queries: the nearest neighbors and range queries. The former is also associated with a distance metric.

2. TRAJECTORY DATA

In this section, we classify the derivation of trajectories into four major categories, briefly introducing a few application scenarios in each category. Trajectory data representing human mobility can help build a better social network [3] and travel recommendation [4].

2.1. Mobility of people: People have been recording their real-world movements in the form of spatial trajectories, passively and actively, for a long time.

- **Active Recording:** Travelers log their travel routes with GPS trajectories for the purpose of memorizing a journey and sharing experiences with friends. Bicyclers and joggers record their trails for sports analysis.
- **Passive Recording:** A user carrying a mobile phone unintentionally generates many spatial trajectories represented by a sequence of cell tower IDs with corresponding transition times.

2.2 Mobility of transportation vehicles: A large number of GPS-equipped vehicles (such as taxis, buses, vessels, and aircrafts) have appeared in our daily life. For instance, many taxis in major cities have been equipped with a GPS sensor, which enables them to report a time-stamped location with a certain frequency. Such reports formulate a large amount of spatial trajectories that can be used for resource allocation [5], traffic analysis [6], and improving transportation networks [7].

2.3. Mobility of animals: Biologists have been collecting the moving trajectories of animals like tigers and birds, for the purpose of studying animals' migratory traces, behavior, and living situations.

2.4. Mobility of natural phenomena: Meteorologists, environmentalists, climatologists, and oceanographers are busy collecting the trajectories of some natural phenomena, such as hurricanes, tornados, and ocean currents. These trajectories capture the change of the environment and climate, helping scientists deal with natural disasters and protect the natural environment we live in.

3. TRAJECTORY DATA PREPROCESSING

This section introduces a fourfold of basic techniques that we need to process a trajectory before starting a mining task, consisting of noise filtering, stay point detection, trajectory compression, and trajectory segmentation.

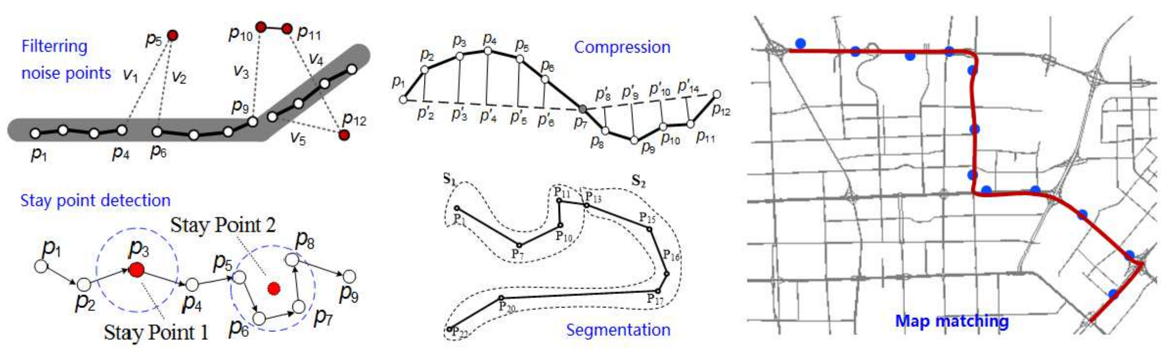
3.1. Noise Filtering: Spatial trajectories are never perfectly accurate, due to sensor noise and other factors, such as receiving poor positioning signals in urban canyons. Sometimes, the error is acceptable (e.g., a few GPS points of a vehicle fall out of the road the vehicle was actually driven), which can be fixed by map-matching algorithms (discussed in Section 3.5). In other situations, as shown in Figure 2, the error of a noise point like p5 is too big (e.g., several hundred meters away from its true location) to derive useful information, such as travel speed. So, we need to filter such noise points from trajectories before starting a mining task.

3.2. Stay Point Detection: Spatial points are not equally important in a trajectory. Some points denote locations where people have stayed for a while, such as shopping malls and tourist attractions, or gas stations where a vehicle was refueled. We call this kind of points "Stay Points." As shown in Figure 2, there are two types of stay points occurring in a trajectory. One is a single point location, for example, Stay Point 1, where a user remains stationary for a while. This situation is very rare, because a user's positioning device usually generates different readings even in the same location. The second type, like Stay Points 2 shown in Figure 2, is more generally observed in trajectories.

3.3 Trajectory Compression: Basically, we can record a time-stamped geographical coordinate every second for a moving object. But, this costs a lot of battery power and the overhead for communication, computing, and data storage. In addition, many applications do not really need such a compression strategies (based on the shape of a trajectory) have been proposed, aiming to reduce the size of a trajectory while not compromising much precision in its new data representation [8]. One is the offline compression (a.k.a. batch mode), which reduces the size of trajectory after the trajectory has been fully generated. The other is online compression, compressing a trajectory instantly as an object travels.

3.4. Trajectory Segmentation: In many scenarios, such as trajectories clustering and classification, we need to divide a trajectory into segments for a further process. The segmentation not only reduces the computational complexity but also enables us to mine richer knowledge, such as sub-trajectory patterns, beyond what we can learn from an entire trajectory.

FIG 2: A) NOISE FILTERING, B) STAY POINT DETECTION, C) TRAJECTORY COMPRESSION, D) TRAJECTORY SEGMENTATION, E) MAP MATCHING



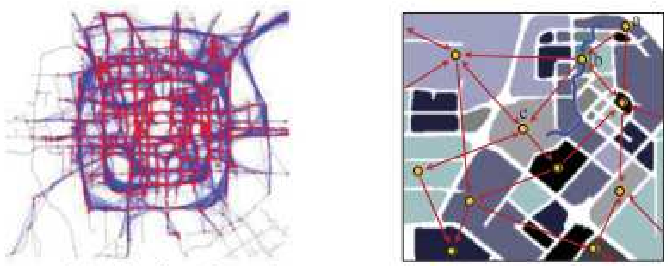
3.5. Map Matching: Map matching is a process to convert a sequence of raw latitude/longitude coordinates to a sequence of road segments. Knowledge of which road a vehicle was/is on is important for assessing traffic flow, guiding the vehicle's navigation, predicting where the vehicle is going, and detecting the most frequent travel path between an origin and a destination, and so forth. Map matching is not an easy problem, given parallel roads, overpasses, and spurs [8].

4. TRANSFER TRAJECTORY TO OTHER REPRESENTATIONS

4.1. From Trajectory to Graph: Trajectories can be transformed into other data structures, besides being processed in its original form. This enriches the methodologies that can be used to discover knowledge from trajectories. Turning trajectories into graphs is one of the representative types of transformation. When conducting such a transformation, the main effort is to define what a node and an edge is in the transformed graph. The methods for transforming trajectories into a graph differentiate between one another, depending on whether a road network is involved in the transformation.

FIG. 3: TRANSFORMING TRAJECTORIES INTO GRAPHS

A) A LAND MARK GRAPH B) A REGION GRAPH

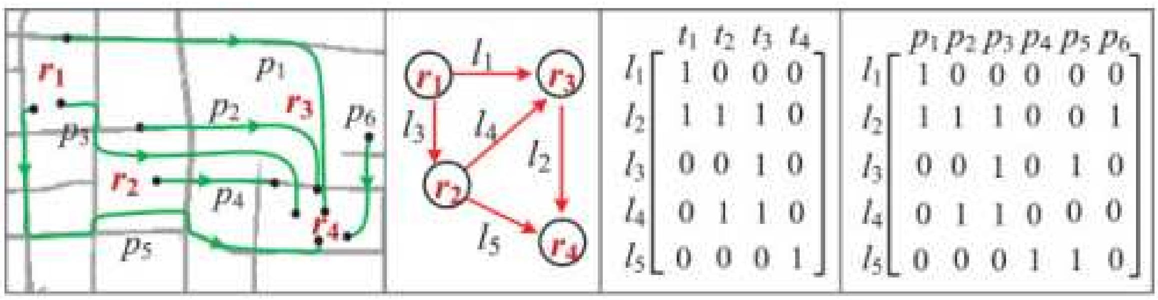


4.2. From Trajectory to Matrix: Another form that we can transform trajectories into is a matrix. Using existing techniques, such as CF and MF, a matrix can help complement missing observations. A matrix can also be used as an input to identify anomalies. The key of the transformation lies in three aspects: (1) what does a row mean, (2) what is a column, and (3) what does an entry denote?

4.3. From Trajectory to Tensor: A nature extension of the matrix-based transformation is turning trajectories into a (3D) tensor, where the third dimension is added to a matrix so as to accommodate additional information.

FIG. 3: TRANSFORMING TRAJECTORIES INTO MATRIX

A) TRAJECTORIES TO PATH B) LINKS C) LINKS TRAFFIC MATRIX D) LINK-PATH MATRIX



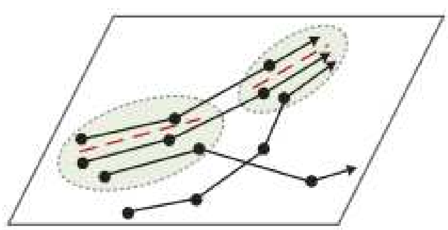
The goal of the transformation is usually to fill the missing entries (in a tensor) or find the correlation between two objects, like two road segments or gas stations. A common approach to solving this problem is to decompose a tensor into the multiplication of a few (low-rank) matrices and a core tensor (or just a few vectors), based on the tensor's nonzero entries. When a tensor is very sparse, in order to achieve a better performance, the tensor is usually decomposed with other (context) matrices in a framework of CF.

5. TRAJECTORY CLUSTERING

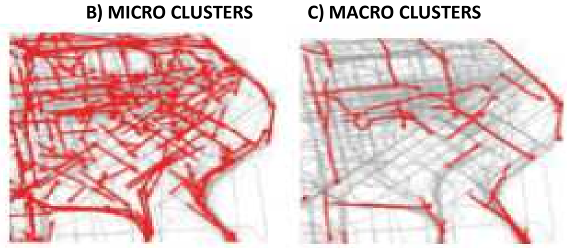
To find representative paths or common trends shared by different moving objects, we usually need to group similar trajectories into clusters. A general clustering approach is to represent a trajectory with a feature vector, denoting the similarity between two trajectories by the distance between their feature vectors. However, it is not easy to generate a feature vector with a uniform length for different trajectories, as different trajectories contain different and complex properties, such as length, shape, sampling rate, number of points, and their orders. In addition, it is difficult to encode the sequential and spatial properties of points in a trajectory into its feature vector.

FIG. 4: TRAJECTORY CLUSTERING BASED ON PARTIAL SEGMENTS [Li et al. 2010b]

A) CLUSTERS OF SEGMENTS



B) MICRO CLUSTERS C) MACRO CLUSTERS



Gaffney and Smyth [1999] [9] and Cadez et al. [2000] [10] proposed to group similar trajectories into clusters by using a regression mixture model and the Expectation Maximization (EM) algorithm. This algorithm clusters trajectories with respect to the overall distance between two entire trajectories. However, moving objects rarely travel together for an entire path in the real world. To this end, Lee et al. [2007] [11] proposed to partition trajectories into line segments and to build groups of close trajectory segments using the Trajectory-Hausdorff Distance, as illustrated in Figure 4(a).

A representative path is later found for each cluster of segments. Since trajectory data are often received incrementally, Li et al. [2010b] further proposed an incremental clustering algorithm, aiming to reduce the computational cost and storage of received trajectories. Both Lee [2007] and Li [2010] adopted a Micro-and-Macroclustering framework, which was proposed by Aggarwal et al. [2003] to cluster data streams. That is, their methods first find mircoclusters of trajectory segments (as demonstrated in Figure 4 (b)), and then group microclusters into macroclusters (as shown in Figure 4 (c)). A major insight of Li's work [Li et al. 2010b] is that new data will only affect the local area where the new data were received rather than the faraway areas.

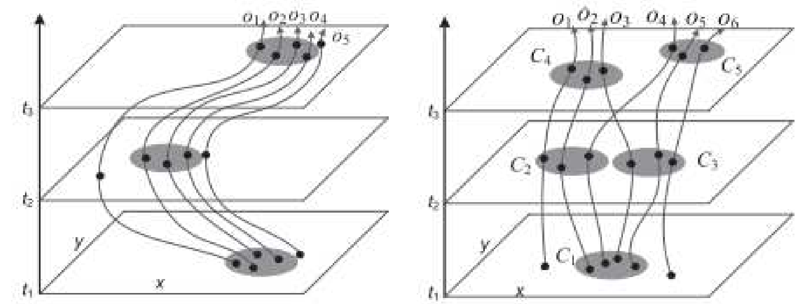
6. TRAJECTORY PATTERN MINING

In this section, we study three major categories of patterns that can be discovered from a single trajectory or a group of trajectories. They are moving together patterns, sequential patterns, and periodic patterns.

6.1. Moving Together Patterns: This branch of research is to discover a group of objects that move together for a certain time period, such as flock [Gudmundsson and Kreveld 2006; Gudmundsson et al. 2004], convoy [Jeung et al. 2008a, 2008b], swarm [Li et al. 2010a], traveling companion [Tang et al. 2012a, 2012b], and gathering [Zheng et al. 2013; Zheng et al. 2014a]. These patterns can help the study of species' migration, military surveillance, and traffic event detection, and so on. These patterns can be differentiated between each other based on the following factors: the shape or density of a group, the number of objects in a group, and the duration of a pattern.

A major concern with flock is the predefined circular shape, which may not well describe the shape of a group in reality, and therefore may result in the so-called lossy-flock problem. To avoid rigid restrictions on the size and shape of a moving group, the convoy is proposed to capture generic trajectory pattern of any shape by employing the density-based clustering. While both flock and convoy have a strict requirement on consecutive time period, Li et al. [2010a] proposed a more general type of trajectory pattern, called swarm, which is a cluster of objects lasting for at least k (possibly nonconsecutive) time stamps.

FIG. 5: EXAMPLES OF MOVING TOGETHER PATTERNS
A) FLOCK, CONVOY AND SWARM B) GATHERING



6.3. Mining Sequential Patterns from Trajectories: A branch of research is to find the sequential patterns from a single trajectory or multiple trajectories. Here, a sequential pattern means a certain number of moving objects traveling a common sequence of locations in a similar time interval. The locations in a travel sequence do not have to be consecutive.

When the occurrence of such a common sequence in a corpus, usually called support, exceeds a threshold, a sequential trajectory pattern is detected. Finding such kind of patterns can benefit travel recommendation [Zheng and Xie 2011b; Giannotti et al. 2007], life pattern understanding [Ye et al. 2009], next location prediction [Monreale et al. 2009], estimating user similarity [Xiao et al. 2014; Li et al. 2008], and trajectory compression [Song et al. 2014]. To detect the sequential patterns from trajectories, we first need to define a (common) location in a sequence. Ideally, in trajectory data, like user check-in sequences from a social net-working service, each location is tagged with a unique identity (such as the name of a restaurant). If two locations share the same identity, they are common. In many GPS trajectories, however, each point is characterized by a pair of GPS coordinates, which do not repeat themselves exactly in every pattern instance. This makes the points from two different trajectories not directly comparable. In addition, a GPS trajectory may consist of thousands of points. Without handled properly, these points will result in a huge computational cost.

6.4. Mining Periodical Patterns from Trajectories: Moving objects usually have periodic activity patterns. For example, people go shopping every month and animals migrate yearly from one place to another. Such periodic behaviors provide an insightful and concise explanation over a long moving history, helping compress trajectory data and predict the future movement of a moving object.

Periodic pattern mining has been studied extensively for time series data. For example, Yang et al. tried to discover asynchronous patterns, surprising periodic patterns [Yang et al. 2001], and patterns with gap penalties, from (categorical) time series. Due to the fuzziness of spatial locations, existing methods designed for time series data are not directly applicable to trajectories. To this end, Cao et al. [2007] proposed an efficient algorithm for retrieving maximal periodic patterns from trajectories. This algorithm follows a paradigm that is similar to frequent pattern mining, where a (global) minimum support threshold is needed. In the real world, however, periodic behaviors could be more complicated, involving multiple interleaving periods, partial time span, and spatiotemporal noises and outliers.

7. TRAJECTORY CLASSIFICATION

Trajectory classification aims to differentiate between trajectories (or its segments) of different status, such as motions, transportation modes, and human activi-ties. Tagging a raw trajectory (or its segment) with a semantic label raises the value of trajectories to the next level, which can facilitate many applications, such as trip recommendation, life experiences sharing, and context-aware computing.

In general, trajectory classification is comprised of three major steps: (1) Divide a trajectory into segments using segmentation methods. Sometimes, each single point is regarded as a minimum inference unit. (2) Extract features from each segment (or point). (3) Build a model to classify each segment (or point). As a trajectory is essentially a sequence, we can leverage existing sequence inference models, such as Dynamic Bayesian Network (DBN), HMM, and Conditional Ran-dom Field (CRF), which incorporate the information from local points (or segments) and the sequential patterns between adjacent points (or segments).

8. PUBLIC TRAJECTORY DATASETS

Collecting data is always the first priority of trajectory data mining. Thanks to researchers in this field, there are quite a few real trajectory datasets that are publicly available:

GeoLife Trajectory Dataset [GeoLife Data]: a GPS trajectory dataset from Microsoft Research GeoLife project [Zheng et al. 2010d], collected by 182 users from April 2007 to August 2012.

T-Drive Taxi Trajectories [T-Drive Data]: A sample of trajectories from Microsoft Research T-Drive project [Yuan et al. 2010a], generated by over 10,000 taxicabs in a week of 2008 in Beijing.

GPS Trajectory with Transportation Labels [Trajectory with transportation modes]: Each trajectory has a set of transportation mode labels, such as driving, taking a bus, riding a bike, and walking. The dataset can be used to evaluate trajectory classification and activity recognition [Zheng et al. 2008a, 2008b].

Check-in Data from Location-based Social Networks [User check-in data]: The dataset consists of the check-in data generated by over 49,000 users in New York City and 31,000 users in Los Angeles as well as the social structure of the users. Each check-in includes a venue ID, the category of the venue, a time stamp, and a user ID.

Hurricane Trajectories [Hurricane trajectory (HURDAT)]: This dataset is provided by the National Hurricane Service (NHS), containing 1,740 trajectories of Atlantic Hurricanes (formally defined as tropical cyclone) from 1851 to 2012. NHS also provides annotations of typical hurricane tracks for each month throughout the annual hurricane season that spans from June to November. The dataset can be used to test trajectory clustering and uncertainty.

9. CONCLUSION

The wide availability of trajectory data has fostered a diversity of applications, calling for algorithms that can discover knowledge from the data effectively and efficiently. This article shows the techniques concerned with different stages of trajectory data mining, recapping them by categories and exploring the differences between one another. This article also suggests the approaches of transforming raw trajectories into other data structures, to which more existing data mining techniques can be applied. And provides an overview on how to unlock the power of knowledge from trajectories, for researchers and professionals from not only computer sciences but also a broader range of communities dealing with trajectories. At the end of this article, a list of public trajectory datasets has been given and a few future directions have been suggested.

REFERENCES

1. Y. Zheng and X. Zhou. 2011. *Computing with Spatial Trajectories*. Springer.
2. Y. Zheng, T. Liu, Y. Wang, Y. Liu, Y. Zhu, and E. Chang. 2014c. Diagnosing New York City's noises with ubiquitous data. In *Proceedings of the 16th ACM International Joint Conference on Pervasive and Ubiquitous Computing.* ACM, 715–725.
3. J. Bao, Y. Zheng, D. Wilkie, and M. F. Mokbel. 2015. A survey on recommendations in location-based social
4. networks. *GeoInformatica*, 19, 3, 525–565.
5. Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W.-Y. Ma. 2011c. Recommending friends and locations based on individual location history. *ACM Transaction on the Web* 5, 1 (2011), 5–44.
6. Y. Zheng, Y. Liu, J. Yuan, and X. Xie. 2011a. Urban computing with taxicabs. In *Proceedings of the 13th International Conference on Ubiquitous Computing.* ACM, 89–98.
7. Y. Wang, Y. Zheng, and Y. Xue. 2014. Travel time estimation of a path using sparse trajectories. In *Proceedings of the 20th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, 25–34.
8. Y. Zhu, Y. Zheng, L. Zhang, D. Santani, X. Xie, and Q. Yang. 2011. *Inferring Taxi Status using GPS Trajectories.* Technical Report MSR-TR-2011-144.
9. W.-C. Lee and J. Krumm. 2011. Trajectory preprocessing. *Computing with Spatial Trajectories*, Y. Zheng and X. Zhou (Eds.). Springer, 1–31.
10. S. Gaffney and P. Smyth. 1999. Trajectory clustering with mixtures of regression models. In *Proceedings of the 5th ACMSIGKDD International Conference on Knowledge Discovery and Data Mining.* ACM, 63–67.
11. D. Pfoser, C. S. Jensen, and Y. Theodoridis. 2000. Novel approaches to the indexing of moving object trajectories. In *Proceedings of the 26th International Conference on Very Large Data Bases.* VLDB Endowment, 395–406.

REQUEST FOR FEEDBACK

Dear Readers

At the very outset, International Journal of Research in Computer Application & Management (IJRCM) acknowledges & appreciates your efforts in showing interest in our present issue under your kind perusal.

I would like to request you to supply your critical comments and suggestions about the material published in this issue, as well as on the journal as a whole, on our e-mail **infoijrcm@gmail.com** for further improvements in the interest of research.

If you have any queries, please feel free to contact us on our e-mail **infoijrcm@gmail.com**.

I am sure that your feedback and deliberations would make future issues better – a result of our joint effort.

Looking forward to an appropriate consideration.

With sincere regards

Thanking you profoundly

Academically yours

Sd/-
Co-ordinator

DISCLAIMER

ABOUT THE JOURNAL

In this age of Commerce, Economics, Computer, I.T. & Management and cut throat competition, a group of intellectuals felt the need to have some platform, where young and budding managers and academicians could express their views and discuss the problems among their peers. This journal was conceived with this noble intention in view. This journal has been introduced to give an opportunity for expressing refined and innovative ideas in this field. It is our humble endeavour to provide a springboard to the upcoming specialists and give a chance to know about the latest in the sphere of research and knowledge. We have taken a small step and we hope that with the active co-operation of like-minded scholars, we shall be able to serve the society with our humble efforts.

Our Other Journals