

Modelling of Trajectory Moving Data with use of Social Media: A Review

Prachi Junwale¹

¹Assistant Professor in Dept. of Computer Engineering, Ramrao Adik Institute of technology, Navi Mumbai, India. prachi.junwale@gmail.com

*Correspondence: Prachi Junwale; Prachi.junwale@gmail.com

ABSTRACT- From last decades, there has been more research on mobile phone network data. Usually, mobile phone user mobility can be track from GIS of device or online social media information with up gradation of location enabled services. Tracking using GIS system is much costlier, time consuming and requires more man power while social media used by smart phones and location enabled devices can easily track the location and time, which can be further utilized by researcher in spatiotemporal domain along with other application related attribute or information. User mobility pattern and prediction are recent interesting area in data and pattern mining. There are many applications for trajectory data modelling such as transportation, tourist place recommendation, land use application, spread of diseases tracking etc. As information in social media data is not dense and difficult to analysis, this article proposes group methodology of user textual data with respect to location and time of user post. In this, user clustering and moving pattern design are made alternatively with each other so that efficiency of method will increases in each stage of method. User clustering is based on Hidden Markov Models (HMMs) which uses definite number of unrecognizable latent states where each record keep up a correspondence to each latent state which has a probabilistic distribution that governs the generation of records as well as text augmentation for proper distribution of textual data with respect to location and time, which is also used to reduce text complexity and not dense data [2]. This method is used to increase prediction accuracy of trajectory moving data.

General Terms: Spatiotemporal data mining.

Keywords: Mobility pattern, Trajectory moving data, Hidden Markov Models.

ARTICLE INFORMATION

Author(s): Prachi Junwale;

Received: 11/07/2023; Accepted: 20/09/2023; Published: 30/09/2023;

e-ISSN: 2347-470X;

Paper Id: IJCSR-030104;

Citation: 10.37391/IJCSR.030104



Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Now a days, tremendous amount of spatio-temporal information are collected, logically gathered, studied and processed in multiple domain such as neurological domain, transport service, agriculture domain, health care, environmental science, crime investigation department and online social network. Spatio-temporal data is different with compare to traditional relational data for which developer and researcher developed many methods and techniques from many years. Traditional relational data contains information related to particular application while spatio-temporal data contains information related to location in the form of latitude and longitude and time along with other normal information. The existence of this additional information brings lot of challenges while processing it. Spatio-temporal mining techniques have been studied from many years in data mining domain. In this article, we have used this technology to analyse trajectory data such as human mobility, animal migration and many more. Trajectory information is a data from moving bodies in terms of location and time for examples animals or birds are moving

from one location to other over time for searching food and to survive better. Another example is taxi transportation application where they use sensors to trace the route of taxi. Usually, information is gathered by using sensors or GIS system periodically from moving bodies about the location over time [1]. Figure 1 shows trajectories of human mobility.

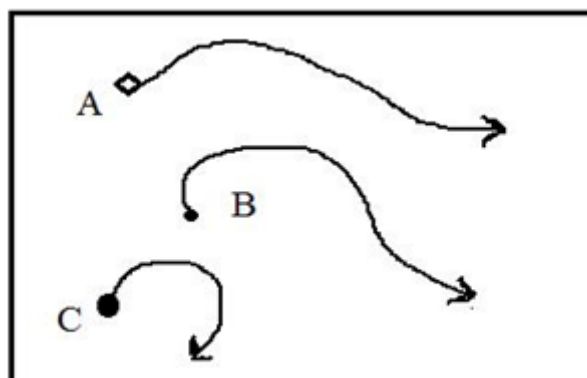


Figure 1: Trajectories of human mobility

There are several recent technologies on Human mobility with a help of social media data in data mining domain for example location prediction, travelling services, urban planning and activity recommendation [2]. GPS information has been used for tracking data about human mobility in many earlier research

technologies while social media enhances these technologies in recent decades. Social media used by smart phones and location enabled devices can easily track the location and time of which can be further utilized by researcher in spatiotemporal domain along with other application related attribute or information. There are two benefits of using online social data. First one is it uses location in terms of latitude and longitude, timestamp along with textual data so one can track who is doing what activity at what time. Another benefit is this data is available in huge amount in terms of size and volume which will cover large information in less cost as compare to cost and time required collecting tracking data by volunteers. Such social information can be easily collected from social media application programming interfaces but it quite complicated to compose high-resolution mobility logs from such data because it has more sparse data [3]. Also only a low percentage of online social network documents are actually geo-tagged, in fact, only 1.6 percent of Twitter users actually have the automatic geo-tagging functionality turned on [4]. There are several challenges of modelling human mobility with spatiotemporal social media data [2]:

- (a) In this application domain, three type of information such as text, location and timestamp are combined. All are having dissimilar syntax, representation and relation with each other so it very difficult to combine them and use efficiently in application.
- (b) As data is not dense and complex, creating reliable application models is difficult [2]. As very less user post their activity at every seen place so it is difficult to train application model for each user. If we train the model for combination of multiple user information so there is possibility of data inconsistency as different people show different mobility pattern.
- (c) As social media message are not so lengthy so it is difficult to take out some knowledge pattern from them.

In this article, social medial information having location of user and timestamp are used in many applications most importantly is to predict next place a user can visit using his text data with spatiotemporal information [6]. Another example is to track mobility prediction of worker and other equipment on construction site area to avoid accidental death or other tragedy [5]. This application gather GPS information of worker and other equipment's along with materials required on construction site and reconstructs workers trajectories using data mining techniques which further pre-processed and cleaned to semantically enhance the information for this application of safety management [5].

Social media information includes location, time and textual data which are used to find users mobility modelling. This method uses grouping of similar moving bodies for mobility modelling which basically use HMM model in which user grouping and mobility modeling both mutually enhance each other [2], to reduce data inconsistency and to reduce the data which is not dense. In addition to this, text augmentation is also used to get efficient latent states.

2. RELATED WORK

The present research on modeling of trajectory moving data can be classified into pattern based, law based and model based [2]. Pattern based method is used to find out trajectory moving prediction information such as T- pattern to find out area or region of interest which is appearing many times in given pattern [7]. There are other methods for pattern mining like flock mining [8], swarm mining [9] and gathering patterns from trajectory data [10] along with periodic mobility pattern mining [2]. Also, so many techniques are available for moving pattern mining from semantic trajectory data [11]. In this article, some different consideration is taken than earlier methods like they are using unstructured textual data, also for getting group of users mobility pattern they are using efficient techniques using spatial and temporal data mining. Law based methods uses physical laws to predict human mobility [2]. Another type is model based method predicted from trajectory movement data [2], in which some fixed location spots is taken where user is going usually at fixed timeslot to find out human mobility as a mixture of Gaussians. This technique is get enhanced when consider social media information. Other technique uses hybrid mobility model that uses heterogeneous mobility data for better location prediction [12]. In this article it uses latent states which show human movement as well as sequential transition among them [2]. Similar research work to this technique uses HMM-based methods in which it proposes the probabilistic latent semantic model, which essentially uses HMM to extract latent semantic locations from cell tower and Bluetooth data [13]. Some techniques use geo-enabled social media as advanced method. In this article they use GMove method which also depends on HMM method, but it differs as it uses textual data, location data and timestamp to predict latent states, as well as they use group of users for mobility pattern prediction [2].

Social media mining study in mobility mining is divided on the basis of its use such as event detection, land-property utilization, mobility pattern and prediction etc. [3]. Social media information usually taken from Twitter and Flickr gives time, location and textual information which is use in many more methods and applications now a days.

For human mobility pattern using model-based method, many research algorithms are used like EM- algorithm, Latent Dirichlet Allocation (LDA) [14], Bayesian networks [16] or Origin Destination (OD) matrices [17]. A density-based clustering for trajectories such as OPTICS algorithm is used to find mobility patterns using spatio-temporal data and textual data and KullbackLeibler (KL) divergence is used as the similarity measurement to mine the evolution of these patterns through time [15]. Fuzzy C-Means (FCM) clustering algorithm is also used on social media like Twitter data for finding mobility pattern in spherical-shape clusters while GK algorithm is used for generating clusters with different geometrical shapes [18]. This both algorithms use only spatial and temporal information. For location prediction, location and time feature is used in ad hoc DBSCAN, Markov model [19], in contrast, Fuzzy Modelling for Human Dynamics [2] uses a technique which consider text feature also. Markov method consider real time observation, in contrast, Fuzzy Modelling uses

combination of offline observation to create clusters and moving pattern and online observation for prediction using this pattern in real time [2].

3. PROPOSED SYSTEM

Social media information includes location, time and textual data which are used to find users mobility modelling. This method uses grouping of similar moving bodies for mobility modelling which basically use HMM model in which user grouping and mobility modelling both mutually enhance each other [2], to reduce data inconsistency and to reduce the data which is not dense. In addition to this, text augmentation is also used to get efficient latent states [2].

Social media data contain a tuple which consist of user identification code, time recorded from message, location in terms of latitude and longitude and group of keywords taken from textual message. This article uses group of user's trajectory records which are dense in nature for better understanding of user's mobility pattern. But if we aggregate all users record to know the mobility pattern then there is possibility of data inconsistency as different people show different mobility pattern. So, the best solution is to use the grouping of people having similar behaviour activity for example consider the tourist visiting to Diu place are usually visited to Daman place in Gujarat state of India. Another example is people usually go to restaurant after watching movie in theatre. Hence one can define this problem by considering human record information taken from social media message and a lot of users profile, make a cluster of similar behaviour activity people and then for each cluster find out the latent states *i.e.* distinct activity people are doing and transition among this states by users.

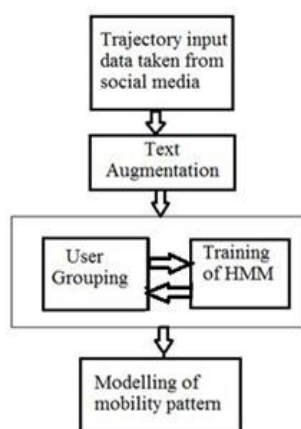


Figure 2: Architecture of System

Figure 2 shows architecture of grouping concept of spatial-temporal social messages to find mobility pattern. In this architecture, trajectory input data is taken from social media messages which include textual, location and timestamp information. Initially text distribution is done on the basis of relation between location and time period of users. Then using this information, grouping technique of online user is done along with training of HMM model alternatively to obtain moving prediction and pattern analysis. In textual data

distribution, it is necessary to find number of time the keyword is appear in text with respect to location in terms of latitude and longitude and time period. After this, semantic relations between words are calculated in each spatiotemporal distribution to get proper grouping of users.

4. STAGES OF MODEL

Proposed system uses grouping of similar moving bodies for mobility modelling which basically use HMM model in which user grouping and mobility modeling both mutually enhance each other [2], to reduce data inconsistency and to reduce the data which is not dense. GMove is based on the Hidden Markov Model (HMM), an effective statistical model for sequential data [2]. HMM uses definite number of unrecognizable latent states where each record keep up a correspondence to each latent state which has a probabilistic distribution that governs the generation of records also with this, the latent state sequence should follows the Markov process *i.e.* a current state only depends on the previous state [2]. This model is based on incremental concept where grouping of people and mobility modelling is done in alternate way to each other. Because of this strategy, in later stages of this model it increments the accuracy also. Other stage of this method is text augementer in which it is difficult to find excellent latent states because textual messages are small in length and includes much text impurity [2]. To avoid this problem, this method find group of textual words depending on relation between location and time period of users which ultimately used to find knowledge from user message. From this information one can perform weighted distribution process of textual data which is used to reduce rare availability of textual data of user and to reduce text impurity. Consider a message, Fun night! at diu beach in Gujarat. This message contain four keywords with supposed weight *i.e.* night (23), diu(45), Gujarat(53), beach (20). The relationship of these words are to be calculated with reference to location and time-stamp, making it more semantically meaningful to form latent states. This is used to reduce rare availability of textual data of user and to reduce text impurity.

5. CONCLUSION

In this article, we studied many research technology used on human trajectory moving data using social media messages. To obtain efficient mobility pattern method, one is using grouping method of similar behavior pattern from user message and does mining and prediction about moving behavior of user. This method uses three information from messages that are text, location in terms of latitude and longitude as well as time period of message when it actually posts on social media. It performs users grouping and moving pattern modelling alternatively with each other to enhance the result. Also this method find group of textual words depending on relation between location and time period of users which ultimately used to find knowledge from user message. This technique has many applications such as location prediction using social media messages, tourist place of interest and next place prediction, animal or bird movement modelling and prediction etc. Thus grouping of social data using spatiotemporal data technique will improve mobility pattern behavior and prediction.

REFERENCES

- [1] Atluri et al., Nov. 2017 Spatio- Temporal Data Mining: A survey of Problems and Methods, ACM Computing Surveys, Vol 1(1).
- [2] Chao Zhang, Keyang Zhang et al. August 13-17, 2016 GMove: Group-Level Mobility Modeling Using Geo-Tagged Social Media, KDD 2016, San Francisco, CA, USA, 2016 ACM. ISBN 978-1-4503-4232.
- [3] Jesus Cuenca-Jara, Fernando Terroso-Saenz Orcid, Mercedes Valdes-Vela and Antonio F. Skarmeta 2017 Fuzzy Modelling for Human Dynamics Based on Online Social Networks- Sensors 2017, 17(9). Licensee MDPI, Basel, Switzerland.
- [4] Leetaru, K., Wang, S., Cao, G.; Padmanabhan, A., Shook, E. 2013 mapping the global Twitter heartbeat: The geography of Twitter. First Monday 2013, 18, doi:10.5210/fm.v18i5.4366.
- [5] Muhammad Arslan, Muhammad Arslan et al. Oct. 2017 Using Spatio-temporal Trajectories to Monitor Construction Sites for Safety Management in ICIME 2017, Barcelona, Spain pp. 211-216.
- [6] Mathew W, Raposo R, Martins B. 2012 Predicting future locations with hidden markov models. UbiComp.911918.
- [7] Giannotti F, Nanni M, Pinelli F, Pedreschi D. 2007 Trajectory pattern mining. KDD: 330339.
- [8] Laube P, Imfeld S. 2002 Analyzing relative motion within groups of trackable moving point objects. GIScience.132144.
- [9] Li Z, Ding B, Han J, Kays R. 2010 Swarm: Mining relaxed temporal moving object clusters. PVLDB. 3(1):723734.
- [10] zheng K, Zheng Y, Yuan NJ, Shang S. 2013 On discovery of gathering patterns from trajectories. ICDE.:242253
- [11] Parent C, Spaccapietra S, Renso C, Andrienko GL, Andrienko NV, Bogorny V, Damiani ML, Gkoulalas-Divanis A, de Maçdo JAF, Pelekis N, Theodoridis Y, Yan Z. 2013 Semantic trajectories modeling and analysis. ACM Comput Surv. 45(4):42.
- [12] Wang Y, Yuan NJ, Lian D, Xu L, Xie X, Chen E, Rui Y. 2015 Regularity and conformity: Location prediction using heterogeneous mobility data. KDD.12751284.
- [13] Deb, B., Basu, P. Wireless Sensor Networks. Springer; 2015. Discovering latent semantic structure in human mobility traces; p. 84-103.
- [14] Ferrari, L.; Rosi, A.; Mamei, M.; Zambonelli, F. Nov.2011. Extracting Urban Pat-terns from Location-based Social Networks. In Proceedings of the 3rd ACM SIGSPATIAL International Workshop on Location-Based Social Networks, Chicago, IL, USA, 1; pp. 916
- [15] Chen, C.C.; Chiang, M.F.; Peng,W.C. 2016. Mining and clustering mobility evolution patterns from social media for urban informatics. Knowl. Inf. Syst., 47, 381403.
- [16] Sui, X.; Chen, Z.; Wu, K.; Ren, P.; Ma, J.; Zhou, F. December 2014 Social Media as Sensor in Real World: Geolocate User with Microblog. In Proceedings of the Natural Language Processing and Chinese Computing: Third CCF Conference, Shenzhen, China, 59; pp. 229237
- [17] Hawelka, B.; Sitko, I.; Beinart, E.; Sobolevsky, S. Kazakopoulos, P.; Ratti, C. 2014 Geo-located Twitter as proxy for global mobility patterns. Cartogr. Geogr. Inf. Sci., 41, 260271.
- [18] Cuenca-Jara, J.; Terroso-Saenz, F.; Valdes-Vela, M.; Gonzalez-Vidal, A.; Skarmeta, A.F. June 2017 Human Mobility Analysis based on Social Media and Fuzzy Clustering. In Proceedings of the 2017 IEEE Global Internet of Things Summit, Geneva, Switzerland; pp. 333338.
- [19] Terroso-Senz, F.; Cuenca-Jara, J.; Gonzlez-Vidal, A.; Skarmeta, A.F. 2016. Human Mobility Prediction Based on Social Media with Com-plex Event Processing. Int. J. Distrib. Sens. Netw., 12, doi: 10.1177/155014775836392



© 2023 by the Prachi Junwale. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license

(<http://creativecommons.org/licenses/by/4.0/>).