IOSR Journal of Computer Engineering (IOSR-JCE)
e-ISSN: 2278-0661,p-ISSN: 2278-8727
PP 10-12
www.iosrjournals.org

A Distributed Representation Framework for Mining Human Trajectory Data

Revathy S.B [1], Remya R[2]
[1](*Department of Computer Science, College of Engineering Perumon, Kollam, India*)
[2](*Department of Information Technology, College of Engineering Perumon, Kollam, India*)

Abstract: *Trajectory data represents the traces of moving objects. Nowadays, there exists different technologies which provides a variety of methods to track human mobility, including user generated geo-tagged contents, check-in services, and mobile apps. Trajectory data has different applications such as location recommendation, social link prediction, behavior analysis, path discovery etc. Trajectory data mining is a challenging task due to the complex characteristics of human mobility. Trajectory data is a kind of sequential data, and such surrounding contexts are necessary to consider for trajectory mining. Here propose a Multi-Context Trajectory Embedding Model, called MC-TEM, to explore the different contexts involved in mining trajectory data. The different contextual information included in the model are user-level, trajectory-level, location-level, and temporal-level contexts. A distributed representation framework is used to characterize the various contexts ,which represents all the contextual features in the same embedding space. Here also proposes a rating based recommendation system by using advanced deep learning methods.*
Keywords: *Trajectory data , location recommendation, location prediction, contextual information*

I. Introduction

Recent progress on satellite and wireless technologies has made it possible to systematically track movements of objects and collect trajectory data. The location- based services provides an opportunity to deeply understand people's behavior. The users have different spatial and temporal activity preference according to their lifestyles. Since trajectory data represents human mobility, it can be utilized for different applications, including personalized location prediction ,social link prediction, behavior analysis etc. It is a challenging task to mine and analyze trajectory data due to the complex characteristics of human mobility. Even for a single check-in record, there are several important factors, including user's regional and categorical preference have to be considered.. There exists many studies related to trajectory data mining. The most recent study has proposed a tensor factorization model, which captures the direct interactions among multiple dimensions, but surrounding contexts involved in a trajectory cannot be directly modeled. Since trajectory data is a kind of sequential data, and such surrounding contexts also need to be consider for trajectory modeling .

Here propose a general Multi-Context Trajectory Embedding Model (MC-TEM), which characterizes different contexts in a general and flexible way. Distributed representation learning method is used for modeling trajectory data. Each contextual feature is represented with a unique distributed vector . A contextual vector can be concatenated using embedding vectors of contextual features. A softmax probability function is used to generate embedding vectors associated with location. The different contexts included in the model are user-level, trajectory-level, location-level and temporal contexts. Different from traditional representations, the dense embedding feature of distributed representation makes it capable of uncovering hidden knowledge in trajectory data. Moreover, it is highly resistant to noise in many fields and yield better performance. It is also convenient to analyze the relationship among multiple contexts. The model is flexible to characterize check-in sequences in a generative process. Location recommendation aims to recommend locations to a user based on user's historical trajectory information. Link prediction using trajectory data aims to identify the likelihood for a pair of users to be friends. There exists many other wide spectrum of applications for trajectory data mining, such as path discovery, movement behavior analysis for individual or a group of moving objects and other applications of urban service.

II. Existing System

A point-of- interest recommendation system [1] was proposed to identify geographical impact, user preference and social impact by using a collaborative filtering approach based on Naïve Bayesian. CF is not promising when the number of visitors is insufficient in a particular location.

A tensor factorization based model [2] was proposed based on collaborative recommendation by considering multiple dimensions including location, time and activities. The relationships between these dimensions are modeled by using higher order matrices called tensors. The model shows significant improvement in runtime and outperforms well.

An unsupervised algorithm called paragraph vector [3] was proposed which returns fixed length representation of features from variable length texts including documents and paragraphs. The algorithm is mainly trained to overcome the disadvantage of bag-of-words models. It can be used in wide range of applications such as sentiment analysis, text classification etc.

A location content-aware system [4] was proposed which recommends locations to a user by considering their personal interests and location preferences. It exploits trajectory patterns and category labels. It also provides home city recommendation and new city recommendation to a user. The classic Threshold algorithm is extended in order to speed up the recommendation process.

An entropy based method [5] for inferring the social connection among users was proposed which analyses people's location information. The model is applicable to a variety of domains. Weighted frequency and diversity are the two factors considered in the model. In cases, where only limited amount of location information is available, the characteristics of each location are also taken into consideration.

A hierarchical graph-based similarity measurement [6] was proposed to model individual's location history for geographic information system. In this model, people's movement behavior and sequence property of mobility is considered. It can explore the correlation between geographic regions and also the relationship among users. It is significant to communities and business to retrieve the information with high relevance.

The time-aware POI recommendation [7] was proposed which recommends point of interest for a user at a particular time in a day. Time plays a vital role in POI recommendation because most users visit different places at different times in a day. In this, the conventional collaborative recommendation system is incorporated with temporal contexts. The model is also enhanced with geographic information.

User based collaborative filtering method using time function [8] was proposed which present a novel algorithm for computing the time weights of items such that decreasing weight is assigned to old data. A user's purchase habits vary. The algorithm uses clustering to discriminate different kinds of items. It shows the different attitudes of same user towards different items. The algorithm improves precision without introducing complexity.

Human mobility analysis by taking the trajectories of mobile phone users was proposed [9] which considers the distance between user's visited location, spatial cosine similarity, co-location rate to predict the social relationship. The similarity between two users movement correlates with their proximity in the social networking systems. Compared to traditional network based factors, mobility measures alone give better predictive power. A supervised classifier is used to improve the prediction accuracy. The method offers new perspective to network dynamics and link prediction.

Trajectory pattern mining [10] was proposed together with different methods for analyzing T patterns from the trajectory data. The background knowledge was also included at the level of preprocessing of trajectory data. It is an extension of sequential pattern mining algorithm. Two notions are considered in this method including region of interest and travel time of users from one location to another. Trajectory patterns are concise representation of mobility.

III. Proposed System

A general Multi-Context Trajectory Embedding Model (MC-TEM) aims to provide a flexible way to characterize various contextual information associated with trajectory data. The context features are modeled using distributed representation framework. Here, a four level hierarchy is adopted to organize contextual features in a top-down way.

Modeling the Contexts

User-Level Contexts*:* User preference is an important factor to be considered for location choices in trajectories. It reflects the behavioral patterns, habits and the interests of users. U dimensions are reserved for user features, and the *u*th dimension specifies whether location *l* is generated by *u* or not. User-level contexts are shared by all trajectories from each user.

Trajectory-Level Contexts: Trajectory-level contexts will be used by the locations in a particular trajectory. Trajectory intent is important for the generation of current location. T dimensions are reserved for trajectory features, and the *t*th dimension specifies whether current location *l* is in trajectory t.

Location-Level Contexts: The current location of a user is influenced by its surrounding locations in a trajectory. Here introduce a context window to denote all previous and successive locations associated with current location. M dimensions are reserved for location features, and the *m*th dimension specifies whether *m* is in the location list.

Temporal Contexts: Temporal contexts consider daily routine and weekly periodicity. Daily routine indicates the overall regular behavior of users. Weekly periodicity indicates the periodic activities by week. Seven dimensions are reserved for week features.

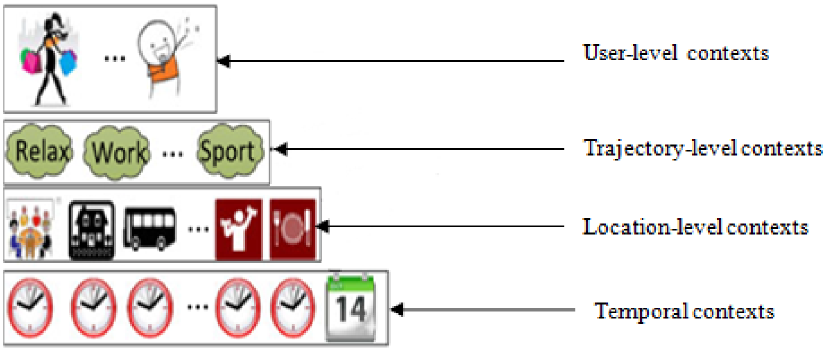

Fig: Different contexts used in the MC-TEM model

The contexts for each check-in location are characterized by using dense vectors in the same embedding space. A hierarchical softmax algorithm is used for learning the parameters of MC-TEM. Stochastic Gradient Descent method is also used to train the parameters. The hierarchical softmax uses a binary tree representation for computing the relative probability of taking a location. The algorithm iterates through the trajectories and updates the embedding vectors of contextual features until the procedure converges. Here MC-TEM is applied to two tasks , namely location recommendation and social link prediction. Location recommendation includes both general and time-aware recommendations. In link prediction , the trajectory data from two users are analyzed and aims to predict whether there exists a link betwesen them. First extract useful features from both users and then construct a feature vector for the pair. Here also proposes to include a rating based recommendation system by using advanced deep learning methods such as convolutional neural networks.

IV. Conclusion

In this paper, we have proposed a general Multi-Context Trajectory Embedding Model to mine human trajectory data. Multiple contexts are included in the model such as user-level, trajectory-level, location level and temporal contexts. A distributed representation framework is used to represent all the contextual information in the same embedding space. The model is flexible to characterize different contexts. It is also convenient to analyze the association among multiple contexts. Here we apply MC-TEM to two challenging tasks namely social link prediction and location recommendation. It uncovers the hidden knowledge in trajectory data and also uncovers noise in many fields.

References

[1]. M. Ye, P. Yin, W.-C. Lee, and D.-L. Lee, Exploiting geographical influence for collaborative point -of- interest recommendation, *in Proc. 34 Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2011, pp. 325–334.

[2]. P. Bhargava, T. Phan, J. Zhou, and J. Lee, Who, what, when, and where: Multi-dimensional collaborative recommendations using tensor factorization on sparse user-generated data, *in Proc. Int. Conf. World Wide Web*, 2015, pp. 130–140.

[3]. Q. V. Le and T. Mikolov , Distributed representation of sentences and documents, *in Proc. Int .Conf Mach. Learn.*, 2014, pp. 1188–1196.

[4]. H. Yin , Y. Sun, B. Cui, Z. Hu, and L. Chen, Lcars: A location content- aware recommender system, *in Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2013, pp. 221–229.

[5]. H. Pham, C. Shahabi, and Y. Liu, Ebm: An entropy-based model to infer social strength from spatiotemporal data, *in Proc. ACM SIGMOD Int. Conf. Manage. Data*, 2013, pp. 265–276.

[6]. Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma, Mining user similarity based on location history, *in Proc. Annu. ACM Int. Symp. Adv. Geographic Inf. Syst.*, 2008, pp. 34:1–34:10.

[7]. Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann , Time aware point-of-interest recommendation, *in Proc. 36th Int. ACM SIGIR Conf. Res. Develop. Inf. Retrieval*, 2013, pp. 363–372.

[8]. Y. Ding and X. Li, Time weight collaborative filtering, *in Proc. ACM Conf. Inf. Knowl. Manage.*, 2005, pp. 485–492.

[9]. D. Wang, D. Pedreschi, C. Song, F. Giannotti, and A.-L. Barabasi, Human mobility, social ties, and link prediction, *in Proc. ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining*, 2011, pp. 1100–1108.

[10]. F. Giannotti, M. Nanni, F. Pinelli, and D. Pedreschi, Trajectory pattern mining, in Proc. ACM SIGKDD *Int. Conf. Knowl. Discovery Data Mining*, 2007, pp. 330–339.