International Journal of Advanced Technology in Engineering and Science
Vol. No.6, Issue No. 10, October 2018
www.ijates.com
ijates
ISSN 2348 - 7550

Topic Modeling based Recommendation in Location-Aware Social Network

Madhu K P[1], D Manjula[2]

[1]*Assistant Professor, Department of Computer Science and Engineering,*
Government Engineering College Idukki, Kerala, India - 685603
[2]*Professor, Department of Computer Science and Engineering, Anna University*
Chennai, India – 600025

ABSTRACT

This article is aimed at developing a recommendation system in a Location Based Social Network with location-tagged media contents such as photos, videos and texts. The tags associated with the items as well as the user profile are modeled to extract topics which helps to give a deeper inference of the user behavior and preferences. The recommendation is done based on this comprehension. A comprehensive study is performed on the popular Foursquare LBSN. The empirical analysis shows that social recommendation using Topic Modeling gives better and accurate results when compared with the systems that use Collaborative Filtering technique.

Keywords: Topic Modeling, Social Recommendation, Location Based Social Networks

INTRODUCTION

With the latest technologies, there is a radical change in the way the information is being stored and exchanged. The task is to retrieve and manipulate the relevant knowledge from the appropriate source in an efficient manner. Social network is a proficient means of managing the information irrespective of user's location and the type of the knowledge ingrained. The recommendation in Location-Based Social Network (LBSN) refers to the recommendation of locations, users, activities or social media [2]. This recommendation must be buoyed by regimented information retrieval techniques.

The organization of topics in this article is as follows: the next section discusses about social networks. The subsequent section briefly explains the concept of topic modeling. This section is followed by a study about the Social Recommendation. Followed by this is the proposed methodology. Subsequent to that are the experimental results and the conclusion.

LOCATION BASED SOCIAL NETWORKS

Location-Based Social Networks (LBSNs) are those social networks which duly create geo-tagged or location-tagged data. Facebook Places, Google Plus, Google Latitude, Gowalla, Brightkite, Twitter, Flickr, Foursquare etc are some of the popular LSBNs. The services given by LBSNs in this context allow users to share information in the form of text, image and video. Though there are constraints on the availability and usage of location information because of security measures, there are still a large number of users who gracefully use their personal information in geo-tagged data.

The technique of location-tagging or geo-tagging facilitates the user to share the data in the social network along with location as an additional attribute.

Foursquare is considered in this article for analysis. In Foursquare, users can check-in at any venue and post their current locations. Users are allowed to add tips to the checked-in venues. This can be shared and can be read by others. These user generated contents like tips and badges are associated with the venues or the point locations. This could enable other users to gain location information about the data. This feature of LBSN is exploited for social recommendation. The recommendations to users could be like things to do, places to visit, malls to shop and so on.

The opportunity of forming potentially useful groups among the users of the network is an advantage of such networks. Users with similar tastes can be dragged into the group using such recommendations. Even it is possible for the users to have group places in private lists named to-dos. The information in the recommendations or tips, venues and to-dos can be characterized as tags (terms, key-words). These tags facilitate the sharing of content taking into account the semantics contained in them.

People use the Foursquare service to share their location with friends, meet new people and get coupons. Users are allowed to connect and publish their check-ins to Facebook and Twitter. Even people who are not a friend of the user on Foursquare can still track or follow the users' whereabouts through Facebook. Foursquare allows the users to check-in at their actual location with the objective of announcing it in a dynamic environment and through text information expressing opinions about certain attended or not places. The Foursquare Social Network was used in this work as a dynamic environment propitious to the system development. The working of the system will be discussed in the proposed methodology section of the article.

TOPIC MODELING - GENERAL APPROACH

Topic Modeling is a technique used to label documents and words with topics. Topic Models facilitate the inference of latent (hidden) topics from visible words in documents. Probabilistic Latent Semantic Indexing (PLSI) is a probabilistic generative topic model. It is given by the following equation.

$$P(d,w) = P(d)P(w \mid d) = P(d)\sum_{z \in Z} P(w \mid z)P(z \mid d) \qquad (1)$$

where *P(d,w)* is the probability of occurrence of a word w in a document d and it can be calculated as the product of *P(d)* and *P(w|d)*, where *P(d)* is the probability distribution of documents and *P(w|d)* is the probability distribution of words given a document. A word selection for a document is given by the above equation. Here the document is selected first and then a word in that document is selected. A document and eventually a whole document corpus could be generated by repeating the above process many times. Let there be a latent topic z in the corpus. It is possible to rewrite the equation as the product of *P(w|z)* and *P(z|d)*, where *P(w|z)* is the probability distribution of words given a topic and *P(z|d)* is the probability distribution of topics given a document. The new equation contributes to the task by adding a new step which is meant for topic selection between the document selection step and the word selection step. As there could be many latent topics in the source from where a word is selected, the products are added over Z, which is the set of all the independent topics *z*.

PLSI and other probabilistic topic models support multiple memberships using the probabilities *P(w|z)* and *P(z|d)*. The strength of the association between a word w and a topic z and that between a topic z and a document d can well be measured by using *P(w|z)* and *P(z|d)*, respectively.

PLSI suffers from the overfitting problem and is not apt for unobserved words. Blei et al. [3] has given Latent Dirichlet Allocation (LDA) to address the above mentioned problem. LDA comes with Dirichlet priors α and β introduced to PLSI to constrain *P(z|d)* and *P(w|z)*, respectively. α is a vector of dimension |Z|, where |Z| represents the number of topics. Every element in the vector α is a prior value for a matching element in *P(z|d)*. A larger value for a particular vector element α_i indicates that the corresponding topic z_i occurs more frequently than other topics in the corpus. Likewise, β is a vector of dimension |W|, where |W| represents the number of words. Every element in the vector β is a prior value for a matching element in *P(w|z)*. A larger value for a particular vector element β_j indicates that the corresponding word w_j occurs more frequently than other words in the corpus. The Dirichlet distribution can be used to simplify the statistical inference. With the use of the Dirichlet priors α and β on *P(z|d)* and *P(w|z)*, these multinomial distributions could be smoothed by the amount of α and β and they become safe from the overfitting problem of PLSI. PLSI could be regarded as a specific instance of LDA, in this regard.

The LDA model is Bayesian model where each document is represented as a mixture of topics while each topic is a discrete probability distribution mostly by an array that defines how common is each word in each topic. LDA considers a document as a collection of weighted topics from which words can be generated. LDA algorithm is as given below.

<u>*LDA's generative model*</u>:

1. *for each topic: decide what words are likely*
2. *for each document,*
   a. *decide what proportions of topics should be in the document,*
   b. *for each word,*
      i. *choose a topic,*
      ii. *given this topic, choose a likely word*

This process can be inverted using standard techniques. Here, the group of topics those were responsible for generating a collection of documents can be inferred using such techniques.

THE LOCATION RECOMMENDER SYSTEMS

Commonly used recommendation systems work by learning the profiles, interests and activity history of the users and recommend products, services or other items depending on the context. LDA is applied to extract the topics from the profiles, interests and activity history of the users in the LBSN. The reviews and the descriptions about different items given by the user is also included in the source on which the topic modeling is applied. Geo-tagged data is the main input for the system.

One of the early recommendation systems is given by Goldberg et al. [10]. It was inclined towards Collaborative Filtration. The system was developed with the objective of information filtration through interactive intervention by the users. That is, recommendation is done by the cooperation among a group of interested persons who are the real beneficiaries of the recommendation. This cooperation creates a Social Recommendation which depends on the personal interests, products, services or recommendation of other users in the. This requires the collection of recommendations by a group of users or required services in similar groups. Also, it contributes to the formation of personalized or customized aggregations in which the diversity in the source or the input tastes will enhance the recommendation.

The technique used in the article makes use of content-based methodology for location recommendation. The system is robust against cold start problem for both users and locations. But, the system needs to maintain structured information about both users and locations. Also, the system does not consider the aggregated community opinions. Thus, some of the short comings of the existing methodologies that are addressed by the proposed system include Cold Start Problem, Similarity / gray sheep and Overspecialization.

THE PROPOSED METHODOLOGY

A methodology based on Topic Modeling is proposed to provide the recommendation in the location-aware social networks which work with geo-located data. Foursquare has been adopted as a domain to the creation of the methodology. Foursquare is one of the emerging location based social networks in this sphere with more than 10 million users and numerous activities happening in the domain. It supports more of a collaborative approach in its social relations and recommendations, which often raises the need for more effective techniques of content-based approach to generate expressive results in the social recommendation. The proposed solution tries to give better results taking into account the limitations listed at the end of the last section.

5.1 Related Works

With the development of technologies that facilitate location-acquisition, the location-based social networking services like Foursquare, Twinkle, and GeoLife are getting flourished [22][23]. Studies by Li et al. [15] and Xiao et al. [18] show that users with similar location history are more likely to have similar interests and preferences. DeScioli et al. [7] states that users who live close to each other are more likely to be friends. Another significant study by Levandoski et al. [14] states that the likelihood that a user would be interested in a location is influenced greatly by the user-location distance. This study is emphasized by the analysis done on Foursquare data. For example, users in Foursquare visit restaurants close to their homes more frequently than others. Ye et al. [21] concludes that location-location distance affects the correlations between locations. The example sites that shopping malls are often located close to each other.

A prominent study by Noulas et al.[17] on the point-location data gathered from Foursquare reveals numerous patterns like 20% of the user check-ins occur within a distance of 1 km, 60% occur between 1 km and 10 km, 20% occur between 10 km and 100 km and a small percentage extend beyond 100 km. Another pattern observed is like the user's activities vary during weekdays and weekends. These kinds of studies together with the analysis on the user and location correlations provide information about user preferences which can be utilized for recommendations in LBSNs.

An important article by Eagle and Pentland [8][9] justifies that the user's historical behaviors is a strong indicator of the user's preferences. This suggestion is more strongly supported by the work by Zheng and Zhou [24] which finds out that a user's historical behavior gathered in an LBSN is more effective and accurate than his online behavior in giving a user's preferences, patterns, interests and experience.

Location features are found to be equally important in profile-based location recommendation systems. Ye et al. [19] developed a method to extract location features based on the temporal distribution of the users' check-ins. Location recommendation systems using Collaborative Filtering (CF) models were proposed by [5][6][12][20]. These systems give personalized recommendations for locations by taking into consideration other users' ratings. Each work gives a different version or addition to the basic CF method.

Bao et al. [1] suggest three main components in a location recommendation system. They are like the user's current location, the user's location histories and the location histories from the other users. Ye et al. [21] presents a location recommendation system that incorporates the user's preferences, the user's social connections and the geographic distance between the user and the candidate locations. The experiments were done on large data set from Foursquare and Whrrl.

The CF-based location recommendation systems works by similarity inference, candidate selection and recommendation score prediction. They suffer from the following problems.

- Because of the large number of the users and items in the system, the similarity model construction is time consuming. This may also pose a problem to the scalability. This may be further worsened by the rapid growth and evolution of LBSNs.
- When the rating matrix is sparse the system would fail in making the accurate or effective recommendations. This may happen when the number of user ratings is low.

- The system performs poorly with the cold start problem.
- The system is constrained if the number of visitors is comparatively less.

So, to overcome these problems, a system may be designed which makes content-based recommendations. The system deploys Topic Modeling technique to find the topics in the user generated contents in the LBSNs. These user generated contents can be the tags associated with the locations, user profiles, user activity history etc.

5.2 Tag Recommendation

Geo-tagging is the process of adding special identification to the media content in the LBSN sites. It can be geographical or location identification. Tags can also be the text associated with the media files such as text, photo or video. For example, photos can be tagged with location. Topic Modeling can be applied to these tags and tips associated with the media files such as text, photos or videos. This is nothing but the content. The Topic Modeling technique will yield the topics associated with these contents. Vectors could be formed with these topics to represent the location for suggestion. Similarly, the same method can be applied on to the user profiles too. This also would yield the topics which will be nothing but the input for the recommendation. Then, recommendation can be done by finding best matches for the profiles.

A study by Heymann et al. [11] shows that the tag-based methods can produce better results. A thorough and comprehensive analysis of the tag based recommendation system follows that the reason for the boom of such recommendation systems are the effective utilization of the methods like filtration and cooperation. The items or the services which are to be found can be associated with tags or it can be used in any other useful manner that the user wishes. These tags can then be used propitiously even as part of the search process to make it efficient and proper [4].

The article proposes a Topic Modeling based approach for the Social Recommendation. This approach makes use of the tags and the user profiles to provide a more accurate recommendation. Tags are nothing but the listed possible candidate terms that the user might be looking for. The aggregation and ranking for the filtration can be done and then recommend socially the most similar or appropriate results to the user. The merits of this research include the possibility to increase the capacity of social recommendation through an association of consult objects between semantics and textual information in the form of tips given to venues.

Foursquare makes it possible for the companies to craft pages containing tips. It also allows the users to follow these companies and receive exceptional and adept tips from them as and when they check-in at certain locations or venues. Even some of the companies allow users to unlock special badges provided they satisfy enough number of check-ins. Most of the companies provide their Facebook, Twitter and other related website links with Foursquare. Tips and lists generated by the companies are also provided to support the users.

5.3 Proposed Social Recommendation System

The various steps involved in the system are as follows.

1) Insertion of User Credentials in the Social Network, Foursquare
2) Comprehending the terms given by the users
3) Data Collection from Foursquare through check-ins shared on Twitter
4) Apply Topic Modeling technique for extracting the topics from the tags and profiles
5) The Social Recommendation

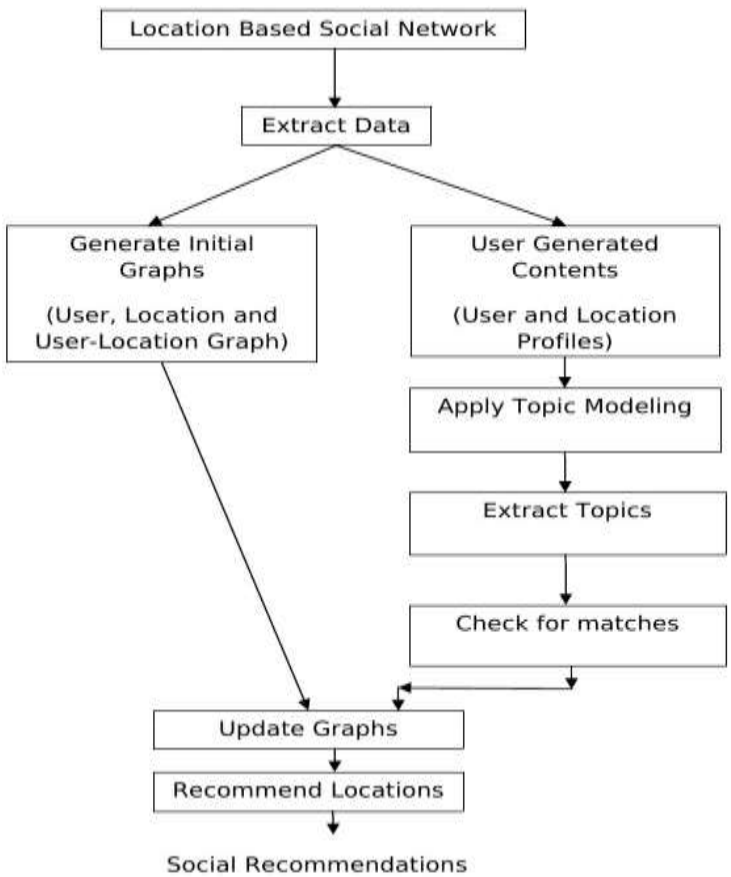

FIG.1: SOCIAL RECOMMENDER SYSTEM

Fig. 1 depicts the social recommender system. The users are required to insert their profiles and access credentials of the Social Network data. This is followed by the automatic generation of credentials: ID and access code. Accessing and manipulation of a profile are achieved by means of an API where there is a basis of data web with user information in the social network. This Foursquare data base is distributed through unstructured data on the

ijates
ISSN 2348 - 7550

configuration JSON (Java Script Object Notation). That means, any information, sharing, comment or profile update is discriminated in a JSON structure ready to be used in the development of future applications.

Now there is the need for the mechanism to extract the data from the structure: the crawler. The crawler mechanism collects data referenced on the corporation of JSON structure. It also serves as parameters on the consultation action done by the user. Combined with the corporation data which are collected there is also geo-located data incorporated as: latitude, longitude, city etc. which help on the mapping process of information patterns in certain places. The hierarchical data organization brings forth the co-occurrence of tags or collaboration tagging. The tags that co-occur on the same resource suggest a semantics similarity. At the end the social recommendation is done to the user with an analysis that goes beyond a collaborative filtration and it generates a group of triple: *User/ Tag / Location* and/ or *Item/ Tag / Location*.

The system is meant to recommend friends and locations based on location matching of individual preferences. User similarity can also be inferred based on location history as an individual's location history implies an individual's interests and behaviors to some extent. The work can be potentially extended to any LBSN which stores activity history of the users.

EXPERIMENTAL RESULTS

The data collected spans over Foursquare, the popular LBSN. Individual user check-in data is not directly accessible from Foursquare. But, it facilitates the users to share their check-ins on Twitter publicly. This feature was made use of, in this work. The publicly available check-ins were crawled using the Twitter's streaming API. Those check-ins which were shared on Twitter could only be crawled. Certain assumptions were made viz. if a pair of users follow each other on Twitter, then a social link exists between them. That is, if two users who are connected on Twitter then, they share similar location preferences. This collected data is just a fair sample of the corresponding data available from the Foursquare LBSN. Filters were used to avoid replicated crawls. A weak social connectivity has been observed in Foursquare. Also, recommendations were not sufficient and accurate when CF-based techniques were deployed. So our model is preferred over the CF-based model.

The proposed recommendation system using topic modeling technique is experimentally compared with the system which is based on Collaboration Filtering technique. Significant improvement in performance is obtained. As part of a comprehensive analysis, particular nodes with known activity history, interests and profiles were selected and the system was made to recommend items for the known user. The item here could be either friends, products, services or even locations where they could be. The system which uses topic modeling was found to give more accurate recommendations compared to the counterpart. The evaluation method used is normalized Discounted cumulative Gain (nDCG) [16]. The effectiveness of the recommendation can be found by using nDCG. Discounted Cumulative Gain (DCG) is commonly used in information retrieval systems. DCG can be used to measure the effectiveness, usefulness or gain of a document on the basis of its position in the result list. Here, the recommendation with respect to a query can be regarded as similar to a document in information retrieval. DCG uses a graded relevance scale of documents in the result set. DCG at a particular rank position *p* is given by the following equation.

$$DCG_p = gr_1 + \sum_{i=2}^{p} \frac{gr_i}{\log_2(i)} \tag{2}$$

where gr_i is the graded relevance of the document in the result set. The length of the result lists differs based on the query. Thus, for getting an accurate result, the DCG is normalized. The normalized DCG is given by the following equation.

$$nDCG_p = \frac{DCG_p}{IDCG_p} \tag{3}$$

where *nDCGp* is the normalized *DCG* at a particular rank position *p*, *DCGp* is the *DCG* at a particular rank position *p* and *IDCGp* is the ideal *DCG* at a particular rank position *p*. Here, the ideal *DCG* refers to the ideal ordering of the recommendations for the given query.

The users were asked to give a query. The recommendations were made based on the query given by the user. The users were asked to judge the relevance of each of the recommendations with respect to the query. Each of the recommendation was judged on a scale of 0-3 with 0 representing irrelevant, 3 representing completely relevant and 1 and 2 representing the in between values. The graded relevance value of each of the recommendations is taken as gr1 through gr_n. Here, n is assigned a maximum value of 10. Then the *DCG*, *IDCG* and *nDCG* values are calculated for result set of each of the queries given by the users and it is used to calculate the effectiveness of the system.

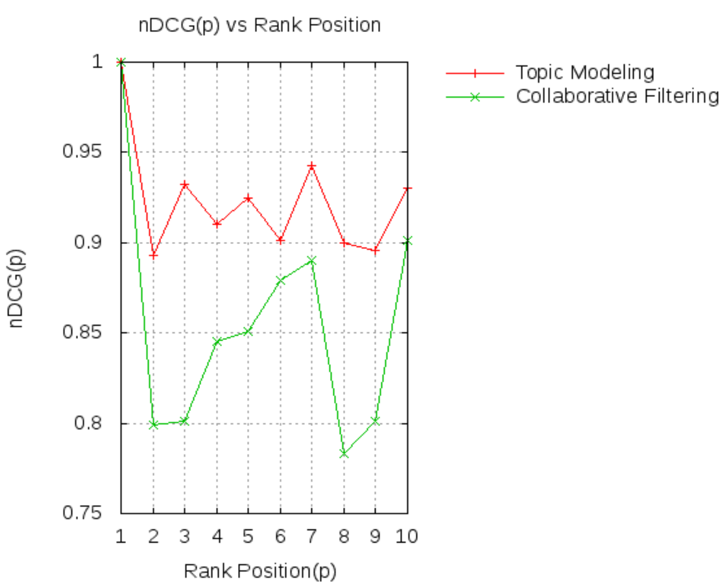

Fig. 2: nDCG(p) versus Rank Position

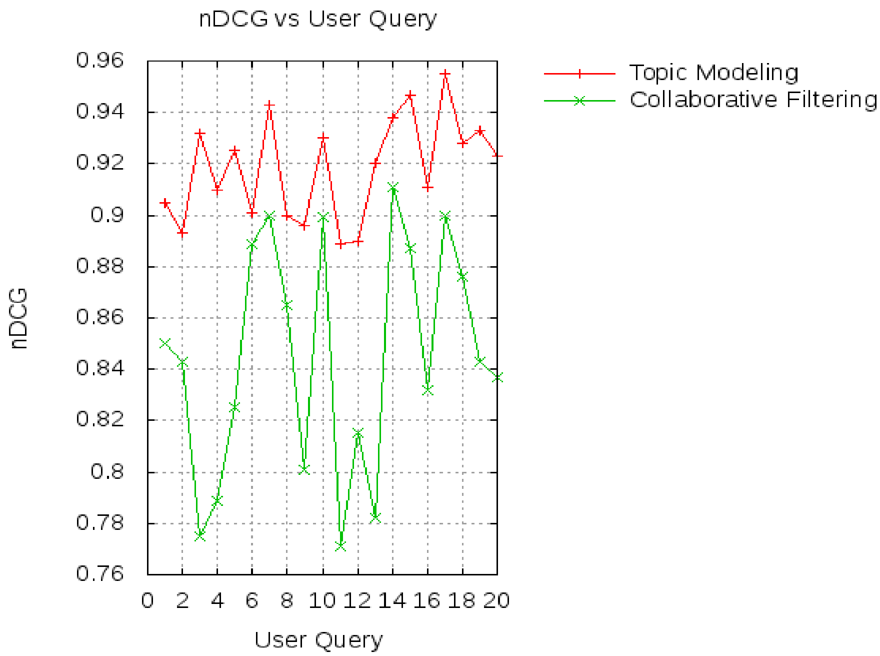

Fig. 3: nDCG versus User Query

Fig. 2 is a sample output graph representing the *nDCG(p)* versus the rank position (*p*) of a particular query and its corresponding recommendations. The experiment was done with both Topic Modeling and Collaborative Filtering based techniques. The relevance judgement was done and *nDCG(p)* is calculated at each rank position for both the cases. The result shows that the recommendation done in the case of Topic Modeling yield better result. That is, recommendations are more accurate in the case of Topic Modeling techniques. The analysis was done on each of the query based recommendations. Another sample graph is shown in Fig. 3. It represents *nDCG* versus user queries. The proposed system using Topic Modeling technique shows a consistent and high *nDCG* value for the recommendations compared to the system using Collaborative Filtering technique.

CONCLUSION

The article presents a methodology to perform the Social Recommendation on Location based Social Networks. The location recommendation is done based on the user generated contents such as the tags associated with the items and the profiles of the users and the items. The development of a system using topic modeling technique for the recommendation in Location Based Social Networks is the main contribution in this work. The technique tries to reduce the ambiguities and redundancies that are found in terms of semantic relations. As a future work, it is planned to extend this model to find the social correlation among the users of the LBSN.

REFERENCES

[1] J. Bao, Y. Zheng and M. F. Mokbel, “Location-based and preference-aware recommendation using sparse geo-social networking data”, in *proceedings of the 20th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, Redondo Beach, CA, USA, November 6-9, 2012.

[2] J. Bao, Y. Zheng, D. Wilkie and M. F. Mokbel, “A Survey on Recommendations in Location-based Social Networks”, *ACM Transactions on Intelligent Systems and Technology*, in press, 2013.

[3] D. M. Blei, A. Y. Ng and M. I. Jordan, “Latent Dirichlet Allocation”, *Journal of Machine Learning Research*, 3, 2003, p.993-1022.

[4] D. Carmel, H. Roitman and E. Yom-Tov, “Social Bookmark Weighting for Search and Recommendation”, *VLDB Journal*, Vol. 19, No. 6, 2010, pp. 761-775.

[5] C.-Y. Chow, J. Bao and M. F. Mokbel, “Towards Location-Based Social Networking Services”, in *proceedings of the 2nd ACM SIGSPATIAL International Workshop on Location Based Social Networks*, 2010. pp. 31-38.

[6] L. Del Prete and L. Capra, “differs: A mobile recommender service”, in *proceedings of the Eleventh International Conference on Mobile Data Management (MDM), IEEE*, Kanas City, Missouri, USA, May 23-26, 2010, pp. 21-26.

[7] P. Descioli, R. Kurzban, E. N. Koch and D. Liben-nowell, “Best Friends Alliances, Friend Ranking, and the MySpace Social Network”, *Perspectives on Psychological Science*, Vol. 6, No. 1, January 2011, pp. 6-8.

[8] N. Eagle and A. Pentland, “Reality mining: sensing complex social systems”, *Journal of Personal and Ubiquitous Computing*, Vol. 10, No. 4, 2006, pp. 255-268.

[9] N. Eagle and A. Pentland, “Eigenbehaviors: Identifying structure in routine”, *Behavioral Ecology and Sociobiology*, Vol. 63, No. 7, April 2009, pp. 1057-1066.

[10] D. Goldberg, D. Nichols, B. M. Oki and D. Terry, “Using Collaborative Filtering to Weave an Information Tapestry”, *Communications of the ACM - Special issue on information filtering*, Vol. 35, No. 12, December 1992, pp. 61-70.

[11] P. Heymann, G. Koutrika and H. Garcia-Molina, “Can Social Bookmarking Improve Web Search?”, in *proceedings of the International Conference on Web search and data mining*, Palo Alto, California, USA, February 11-12, 2008, pp. 195-206.

[12] T. Horozov, N. Narasimhan and V. Vasudevan, "Using location for personalized poi recommendations in mobile environments", *International Symposium on Applications and the Internet, SAINT 2006*, Phoenix, AZ, January 23-27, 2006, pp. 124-129.

[13] L. Jin, X. Long and J.B.D. Joshi, "Towards understanding Residential Privacy by Analyzing User' Activities in Foursquare", *Workshop on Building Analysis Datasets and Gathering Experience Returns for Security (BADGERS'12), held in conjunction with CCS 2012*, Raleigh, NC, USA, October 15, 2012.

[14] J.J. Levandoski, M. Sarwat, A. Eldawy and M. F. Mokbel, "LARS: A Location-Aware Recommender System", in *proceedings of the 28th IEEE International Conference on Data Engineering (ICDE),* Washington, DC, April 1-5, 2012, pp. 450-461.

[15] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu and W.-Y. Ma, "Mining user similarity based on location history", in *proceedings of the 16th ACM SIGSPATIAL international conference on Advances in geographic information systems*, Irvine, CA, USA, November 05 - 07, 2008, Article No. 34.

[16] C. D. Manning, P. Raghavan and H. Schutze, *Introduction to information retrieval*, Vol. 1, Cambridge University Press, Cambridge, 2008.

[17] A. Noulas, S. Scellato, C. Mascolo and M. Pontil, "An empirical study of geographic user activity patterns in foursquare", in *proceedings of the Fifth International AAAI Conference on Weblogs and Social Media, ICWSM*, Poster Paper. Barcelona, Spain, July 2011, pp. 570-573.

[18] X. Xiao, Y. Zheng, Q. Luo and X. Xie, "Finding similar users using category-based location history", *Proceedings of the 18th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, New York, NY, USA, 2010, pp. 442-445.

[19] M. Ye, K. Janowicz, K., C. Mulligann and W.-C. Lee, "What you are is when you are: The temporal dimension of feature types in location based social networks", in *proceedings of the 19th ACM SIGSPATIAL International Conference on Advances in Geographic Information Systems,* ACM, New York, NY, USA, 2011a, pp. 102-111.

[20] M. Ye, P. Yin and W.-C. Lee, "Location recommendation for location-based social networks", in *proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems*, ACM, New York, NY, USA, 2010, pp. 458-461.

ijates
ISSN 2348 - 7550

[21] Ye, M., Yin, P., Lee, W., and Lee, D., "Exploiting geographical influence for collaborative point-of-interest recommendation", in *proceedings of the 34th ACM International conference on Research and development in Information Retrieval (SIGIR)*, New York, NY, USA, 2011c, pp. 325-334.

[22] Y. Zheng, Y. Chen, X. Xie and W.-Y. Ma, "GeoLife2.0: A Location-Based Social Networking Service", in *proceedings of the 10th IEEE International Conference on Mobile Data Management: Systems, Services and Middleware*, MDM'09, IEEE, Taipei, May 18-20, 2009c, pp. 357-358.

[23] Y. Zheng, X. Xie and W.-Y. Ma, "Geolife: A collaborative social networking service among user, location and trajectory", *IEEE Data Engineering Bulletin*, Vol. 33, No. 2, June 2010b, pp.32-40.

[24] Y. Zheng and X. Zhou, *Computing with Spatial Trajectories*, Springer, 1 November 2011, Chapter 8: Location-based social networks: Users.