International Journal of Scientific Research in Computer Science, Engineering and Information Technology
© 2018 IJSRCSEIT | Volume 3 | Issue 2 | ISSN : 2456-3307

# A Survey on Location Based News Recommendation

**Mrunalini Bhosale, Prof. U. A. Mande**

Department of Computer Networks, Sinhgad College of Engineering, Savitribai Phule, Pune University, Pune, Maharashtra, India

## ABSTRACT

On the Internet, where the number of choices is overwhelming, there is need to filter, prioritize and efficiently deliver relevant information in order to alleviate the problem of information overload, which has created a potential problem to many Internet users. Recommender systems solve this problem by searching through large volume of dynamically generated information to provide users with personalized content and services. This paper explores the different characteristics and potentials of different filtering techniques in recommendation systems in order to serve as a compass for research and practice in the field of recommendation systems, basically we focus on the system developed for location aware based news recommendation system, we also discussed the brief survey on semantic analysis used for recommendation.

**Keywords :** Recommender System, Information Retrieval, Collaborative Filtering, Semantic Analysis.

## I. INTRODUCTION

Content-based news article recommendation aims to offer news articles to a reader based on his/her interests. To reflect the interests of a specific reader into article recommendation, the interests are predicted from the user profile that is a collection of articles he/she has read in the past or explicit information about the reader such as age and gender. This kind of user profile is static or at least almost static. The profile remains same regardless of the user location, even if a large number of users of handheld devices (i.e. smart phone or tablet PC) access online news providers anywhere, and their interests depend on their location. The fact that the interests of a user are strongly related to their location, implies that his/her geographical context is important for localized news article recommendation. For instance, assume that a user is reading an article "After Delays, Wireless Web Comes to Parks" of The New York Times. This article delivers news about establishing wireless networks at the prominent parks in New York city. Therefore, it has two major topics, a new facility of public parks and a company that won a contract for the wireless net-work service. The people who are strolling at a public park will focus on the former, while those working at Wall Street would prefer the latter.

The key of news article recommendation is obviously a representation of the articles. Word frequencies or topics are the most widely used representation of articles [11, 18], but recently, topic representation is preferred because a topic is a good proxy for article contents [23]. For instance, Wang and Blei proposed a collaborative topic model using Latent Dirichlet Allocation (LDA) to recommend scientific articles [2]. Egozi et. al [5] projected articles onto a topic space con- structed with Explicit Semantic Analysis (ESA) [7]. In ESA, each Wikipedia concept is regarded as a possible topic, and an article is represented as a topic vector with Wikipedia concepts. Because Wikipedia has a large volume of concepts, articles can be expressed efficiently and accurately. On the other

hand, such topic representations of articles are geo-neutral. Therefore, a topic model that reflects the geographical context of a user is needed to localize an article recommendation.

When the location at which a user is standing is the only information available for the user profile, the geographical context of a user is equivalent to the geographical context of his/her location. The geographical context of a location can also be expressed as a topic vector, because most locations have their own geographical topics that are defined as spatially coherent meaningful themes [25]. Most topic models proposed for extracting geographical topics from a location identify topics actually from a set of geo-tagged documents associated with the location [13]. Nevertheless, they have a problem in that the geo-tagged documents include not only the topics that are related directly to the location but also some other topics that have no relationship with the location.

### 1.2 Localized Recommendation

The popularity of handheld devices with a GPS makes the user location available for various kinds of recommendations[12, 17, 19]. Most previous studies on localized recommendation have focused on the physical attributes of locations. Dao et al. proposed a context-aware collaborative filtering for location-based advertising [4]. In their study, the item scores for a specific user are determined by considering user's location. That is, the user-item matrix of this method is expanded with the user locations. Thus, the user similarity can be obtained using the items shared by other users at the same location.

Zheng et al. proposed the user-centered collaborative location and activity filtering (UCLAF) to recommend tourism spots and activities [27]. In UCLAF, the similar users of a specific user are obtained based on their GPS trajectories and activities at tourism spots. UCLAF recommends tourism spots for the user by analyzing similar users. CityVoy- ager system proposed by Takeuchi and Sugimoto [22] uses the shops previously visited by a user to recommend new shops for him/her. Ye et al. proposed Geo-Measured Friend- based Collaborative Filtering (GM-FCF) to recommend places such as stores, movie theaters, etc [24]. The recommendation by GM-FCF is made using the places visited by his/her friends in the social network. GM-FCF ranks the places by their physical distance from his/her location. Yu and Chang suggested a tour planning system [26]. This system also recommends sightseeing spots, hotels, and restaurants based on user's location.

As shown in these studies, the user's location helps recommendation systems improve their performance. This is because it reveals valuable information of the user. This was recently taken into account for news recommendation. GeoFeed is a location-aware news feed system [1]. This system provides its users with the news that are spatially related with the users. The spatial relationship between the user and a news in GeoFeed is determined using the distances from his/her to the locations tagged in the news. Mokbel et al. employed user locations to recommend local news [14]. In their study, they considered the distance between the user and location in which the news article was published. Although these studies proved the feasibility of their idea by implementing practical localized recommender systems, only the physical attributes of the user location was considered but its latent attributes are ignored. To the best of our knowledge, no study has used the latent attributes of the user locations in news article recommendation.

### 1.3 Geographical Topic Model

A location is described physically by its longitude and latitude. This is because it is difficult to transform a location into other forms but the coordinate of the longitude and latitude, even though most locations have their own geographical context and the context can be expressed with topics at these

locations. For instance, the topics 'shopping'at a department store and 'trip' at a sightseeing spot describe their locations well. One of the efforts to extract the geographical topics of a location is the geographical topic model. Therefore, range of geographical topic models have been proposed [20, 28].

## II. Related Work

In [1] presents the GeoFeed system; a location- aware news feed system that provides a new platform for its users to get spatially related message updates from either their friends or favorite news sources. GeoFeed distinguishes itself from all existing news feed systems in that it enables users to post message with spatial extent rather than static point locations, and takes into account their locations when computing news feed for them. GeoFeed is equipped with three different approaches for delivering the news feed to its users, namely, spatial pull,spatial push ,and shared push. We design a smart model for GeoFeed to decide about using these approaches in a way that:(a) minimizes the system overhead for delivering the location- aware news feed, and (b) guarantees a certain response time for each user to obtain the requested location-aware news feed. GeoFeed also supports location-aware news feed function for its mobile users.

In [3] argue that the performance of content-based news recommender systems has been hampered by using relatively old and simple matching algorithms. Using more current probabilistic retrieval algorithms results in significant performance boosts.

In [4] propose a new recommendation model, which we termed Context-Aware Collaborative Filtering using genetic algorithm (CACF-GA), for location-based advertising (LBA) based on both user's preferences and interaction's context. We first defined discrete contexts, and then applied the concept of ''context similarity'' to conventional CF to create the context-aware recommendation model. The context similarity between two contexts is designed to be optimized using GA. We collect real-world data from mobile users, build a LBA recommendation model using CACF-GA, and then perform an empirical test to validate the usefulness of CACF-GA.

In [5] introduce a new concept-based retrieval approach based on Explicit Semantic Analysis (ESA), a recently proposed method that augments keyword-based text representation with concept-based features, automatically extracted from massive human knowledge repositories such as Wikipedia. These approach generates new text features automatically, and we have found that high-quality feature selection becomes crucial in this setting to make the retrieval more focused. However, due to the lack of labeled data, traditional feature selection methods cannot be used, hence we propose new methods that use self-generated labeled training data.

In [6] present a multi-level generative model that reasons jointly about latent topics and geographical regions. High-level topics such as "sports" or "entertainment" are rendered differently in each geographic region, revealing topic-specific regional distinctions. Applied to a new dataset of geotagged microblogs, our model recovers coherent topics and their regional variants, while identifying geographic areas of linguistic consistency. The model also enables prediction of an author's geographic location from raw text, outperforming both text regression and supervised topic models.

In [7] propose Explicit Semantic Analysis (ESA), a novel method that represents the meaning of texts in a high-dimensional space of concepts derived from Wikipedia. We use machine learning techniques to explicitly represent the meaning of any text as a weighted vector of Wikipedia-basedconcepts. Assessing the relatedness of texts in this space amounts to comparing the corresponding vectors using conventional metrics (e.g., cosine).

In [8] show that PageRank, computed on a semantic network constructed from word-association data, outperformed word frequency and the number of words for which a word is named as an associate as a predictor of the words that people produced in this task. We identify two simple process models that could support this apparent correspondence between human memory and Internet search, and relate our results to previous rational models of memory.

In [9] focus on Twitter and present an algorithm by modeling diversity in tweets based on topical diversity, geographical diversity, and an interest distribution of the user. Furthermore, we take the Markovian nature of a user's location into account. These model exploits sparse factorial coding of the attributes, thus allowing us to deal with a large and diverse set of covariates efficiently. Our approach is vital for applications such as user profiling, content recommendation and topic tracking. Author show high accuracy in location estimation based on our model. Moreover, the algorithm identifies interesting topics based on location and language.

In [12] propose a location based reminder system with image recognition technology. With this system, mobile phone users can actively capture pictures from their favorite product or event promotional materials. After the phone user sends the picture to a computer server, location based reminders will be downloaded to the phone. The mobile phone will alert the user when he/she is close to the place where the product is selling or the event is happening. Kd-tree image matching and geometric validation are used to identify which product the user is interested in. A mobile client application is developed to take pictures, conduct GPS location tracking and pop up the reminder.

In [14] present our vision of Location-based Services (LBS) 2.0, where users can generate significant location-based content and have meaningful location-aware interaction with both the system and other users. There are two ways to look at LBS 2.0,either as embedding location-awareness into existing Web 2.0 infrastructures, or embedding Web 2.0 functionality inside the core of existing location-based services. We take the former approach, which makes use of the existing Web 2.0 model, thus building upon an already-successful infrastructure. This approach is in contrast to reinventing Web 2.0 modules inside a location-based service environment. In other words, our approach is similar to the story of spatial databases over the last two decades, where the spatial functionalities were embedded inside existing database systems, making use of the existing infrastructure including query operators, optimizers, indexing, and transaction processing.

In [15] Recommender systems improve access to relevant products and information by making personalized suggestions based on previous examples of a user's likes and dislikes. Most existing recommender systems use social filtering methods that base recommendations on other users' preferences. By contrast, content-based methods use information about an item itself to make suggestions. This approach has the advantage of being able to recommended previously unrated items to users with unique interests and to provide explanations for its recommendations. We describe a content-based book recommending system that utilizes information extraction and a machine-learning algorithm for text categorization. Initial experimental results demonstrate that this approach can produce accurate recommendations. These experiments are based on ratings from random samplings of items and we discuss problems with previous experiments that employ skewed samples of user-selected examples to evaluate performance..

## III. Conclusions

In this work, we have provided a survey of location-aware recommendation systems . We first described the basics of LARS and some generic approaches. Then, we presented a number of location-aware recommendation systems for several scenarios. In the

last decade, location-aware recommendation approaches made an important progress thanks to significant efforts developed by the research community. Nevertheless, more research is needed to solve existing difficulties and design systems able to obtain more effective recommendations.

## IV. REFERENCES


[1]. J. Bao, M. Mokbel, and C. Chow. GeoFeed: A location-aware news feed system. In Proceedings of the 28th IEEE International Conference on Data Engineering, pages 54–65, 2012. 301

[2]. D. Blei, A. Ng, and M. Jordan. Latent Dirichlet Allocation. Journal of Machine Learning Research,3:993–1022, 2003.

[3]. T. Bogers and A. Bosch. Comparing and evaluating information retrieval algorithms for news recommendation. In Proceedings of ACM conference on Recommender systems, pages 141–144, 2007.

[4]. T. Dao, S. Jeong, and H. Ahn. A novel recommendation model of location-based advertising: Context-aware collaborative filtering using ga approach. Expert Systems with Applications,39(3):3731–3739, 2012.

[5]. O. Egozi, S. Markovitch, and E. Gabrilovich. Concept-based information retrieval using Explicit Semantic Analysis. ACM Transactions on Information Systems, 29(2):8:1–8:34, 2011.

[6]. J. Eisenstein, B. O'Connor, N. A. Smith, and E. P. Xing. A latent variable model for geographic lexical variation. In Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing ,pages 1277–1287, 2010.

[7]. E. Gabrilovich and S. Markovitch. Computing semantic relatedness using Wikipedia-based Explicit Semantic Analysis. In Proceedings of the 20th International Joint Conference on Artificial Intelligence, pages 1606–1611, 2007.

[8]. T.Griffiths,M.Steyvers,andA.Firl. Google and the mind: Predicting fluency with PageRank. Psychological Science, 18(12):1069–1076, 2007.

[9]. L. Hong, A. Ahmed, S. Gurumurth, A. Smola, and K. Tsioutsiouliklis. Discovering geographical topics in the twitter stream. In Proceedings of the 21st International Conference on World Wide Web, pages 769–778, 2012.

[10]. J. Kamps and M. Koolen. The importance of link evidence in Wikipedia. In Proceedings of the 30th European Conference on IR Research , pages 270–282, 2008.

[11]. R. Krestel, P. Fankhauser, and W. Nejdl. Latent Dirichlet Allocation for tag recommendation. In Proceedings of the 6th ACM Conference on Recommender Systems, pages 61–68, 2009.

[12]. Y.Li,A.Guo,S.Liu,Y.Gao,andY.Zheng.A location based reminder system for advertisement. In Proceedings of the 18th International Conference on Multimedia , pages 1501–1502, 2010.

[13]. Q. Mei, C. Liu, and H. Su. A probabilistic approach to spatiotemporal theme pattern mining on weblogs. In Proceedings of the 15th International Conference on World Wide Web , pages 533–542, 2006.

[14]. M. Mokbel, J. Bao, A. Eldawy, J. Levandoski, and M. Sarwat. Personalization, socialization, and recommendations in location-based services 2.0. In Proceedings of the International Workshop on Personalized Access, Profile Management, and Context Awareness in Databases 2011, collocated with VLDB , pages 1–6, 2011.

[15]. R. Mooney and L. Roy. Content-based book recommending using learning for text categorization. In Proceedings of the SIGIR-99 Workshop on Recommender Systems: Algorithms and Evaluation 1999.

[16]. L. Page, S. Brin, R. Motwani, and T. Winograd. The pagerank citation ranking: Bringing order to the webTechnical report, Computer Systems Laboratory, Stanford University, 1998.

[17]. M. Park, J. Hong, and S. Cho. Location-based recommendation system using bayesian user's preference model in mobile devices. In

Proceedings of the 4th International Conference on Ubiquitous Intelligence and Computing, pages 1130–1139, 2007.

[18]. D. Ramage, D. Hall, R. Nallapati, and C. Manning. Labeled LDA: a supervised topic model for credit attribution in multi-labeled corpora. In Proceedings of the 2009 Conference on Empirical Methods in Natural Language Processing, pages 248–256, 2009.

[19]. N. Savage, M. Baranski, N. Chavez, and T. Hollerer. I'm feeling loco: A location based context aware recommendation system. In Advances in Location-Based Services, pages 37–54, 2012.

[20]. S. Sizov. GeoFolk: Latent spatial semantics in web 2.0 social media. In Proceedings of the 3rd ACM International Conference on Web Search and Data Mining, pages 281–290, 2010.

[21]. J. Son, Y. Noh, H. Song, S. Park, and S. Lee. Location Comparison through geographical topics. In Proceedings of the 2012 IEEE/WIC/ACM International Conference on Web Intelligence , pages 311–318, 2012.

[22]. Y. Takeuchi and M. Sugimoto. CityVoyager: An outdoor recommendation system based on user location history. In Proceedings of the 3rd International Conference on Ubiquitous Intelligence and Computing, pages 625–636, 2006.

[23]. X. Wei and B. Croft. LDA-based document models for ad-hoc retrieval. In Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pages 178–185, 2006.

[24]. M. Ye, P. Yin, and W. Lee. Location recommendation for location-based social networks. In Proceedings of the 18th SIGSPATIAL International Conference on Advances in Geographic Information Systems, pages 458–461, 2010.

[25]. Z. Yin, L. Cao, J. Han, C. Zhai, and T. Huang. Geographical topic discovery and comparison. In Proceedings of the 20th International Conference on World Wide Web , pages 247–256, 2011.

[26]. C. Yu and H. Chang. Personalized location-based recommendation services for tour planning in mobile tourism applications. In Proceedings of the 10th International Conference on E-Commerce and Web Technologies, pages 38–49, 2009.

[27]. V. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang. Collaborative filtering meets mobile recommendation: A user-centered approach. In Proceedings of the 24th AAAI Conference on Artificial Intelligence, pages 236–241, 2010.

[28]. Y. Zhou and J. Luo. Geo-location inference on news articles via multimodal pLSA. In Proceedings of the20th ACM International Conference on Multimedia, pages 741–744, 2012.