

An Efficient Personalized POI Recommendation using PCA-SVM based Filtering and Classification

Sunayna Sharma

Samrat Ashok Technological Institute
Department of Computer Science and Engineering,
Vidisha, India

Anil Suryavanshi

Asst. Prof. CSE department
Samrat Ashok Technological Institute,
Department of Computer Science and Engineering,
Vidisha, India

ABSTRACT

The rapid growth of cities has developed an increasing number of points of interest (POIs), e.g., restaurants, stores, hotels, etc; to enrich people's life, providing us with more choices of life experiences than before. People are willing to explore the city and neighborhood in their daily life and decide "where they should go" according to their personal interest and various choices of POIs. The Existing Methodology implemented for the Filtering of POI Recommendation is efficient but contains less Precision and Recall, hence a new and efficient technique for the POI Recommendation using Principle Component Analysis with Support Vector Machine Learning is proposed which provides more efficient results in comparison.

Keywords

Collaborative Filtering, POI Recommendation, Support Vector Machine, Principle Component Analysis, Mean Average Precision.

1. INTRODUCTION

Collaborative filtering (CF) is the efficient and common technique of predicting the interests of a user by collecting preference information from many users. In order to determine which items from the collection may be favored by individual users, conventional CF approaches take the ratings previously as-signed to items by a target user and use them together with ratings of the users with similar preferences to predict the ratings of yet-unseen items. Items are then, recommended in a descending order according to their predicted ratings.

Developing POI recommender systems requires observation of the human mobility with respect to real-world POIs, which is infeasible with traditional mobile data. However, the recent development of location-based social networks (LBSNs) provides such observation. The typical location-based social networking sites allow users to "check in" at POIs with smart phones, leave tips and share those experiences with their online friends. The increasing number of LBSN users has generated the large amounts of LBSN data, providing an unprecedented opportunity to study human mobility for the personalized POI recommendation in spatial, temporal, social, and content aspects. Recently, topic model method (TM) has been introduced into personalized travel recommendations [1], [2]. TM is similar to the content-based method in product recommendation systems [3]. TM analyzes tourist's travel preferences (such as culture, urban landscape, or landscape) and recommends POIs which match the themes of user preferences. Through interest category mapping, even if the user has been visited very few points of interest, we can still analyze user preferences.

At the same time, making a satisfying decision efficiently among the large number of POI choices becomes very tough problem for a user. To facilitate the user's exploration and decision making, POI recommendation has been introduced by location-based services such as Yelp and Foursquare. The category topics are usually determined by the naive category information from recommended systems in TM [1], [5]. For example, the original category information of social media websites, such as Foursquare [5], ODP [4], and Yelp [5], serve as topics. From the predetermined categories, it is convenient to calculate user preferences. Unfortunately, for rich photo sharing networks like Flickr and Panoramio, there is no such defined category information. However, such recommendation models are commonly based on majority users' preference on POIs, which ignore a user's personal preference. Comparing to visiting places that best fit a user's interest, visiting places against a user's taste may give him very terrible experience, especially in a situation when the user travels to a new place. Therefore, personalized POI recommendation is proposed to help user's filter out uninteresting venues according to their own taste and save their time in decision making. Before the Web 2.0 era, analyzing user's mobility for personalized POI recommendation is infeasible even the mobile devices are widely adapted with large amount of cell phone-based GPS data available, as there is no indication of POI information from the GPS data other than longitude and latitude records. For example, we could observe a set of locations in terms of longitude and latitude pairs that a user has been to, while there is no easy way to figure out whether a specific pair of longitude and latitude is corresponding to a restaurant, or a hotel, or just a point on highway, since all these information are passively recorded by mobile devices. With the developing of Web 2.0 technology, a number of location-based social networking services, e.g., Foursquare, Yelp, and Facebook Places, have emerged in recent years, making the study of personalized POI recommendation possible. Typical location-based social networking services maintain a large POI database and allow a user to "check-in" at POI with his smartphone regarding to his current physical location. The user can also leave tips and share the "check-in" experience with his online friends, along with creating the opportunity to make new friends. According to a recent survey from the Pew Internet and American Life Project, over the past year 18% of smartphone owners use geosocial services to "check in" at certain locations and share them with their friends, while this percentage has risen from 12% in 2011 [6].

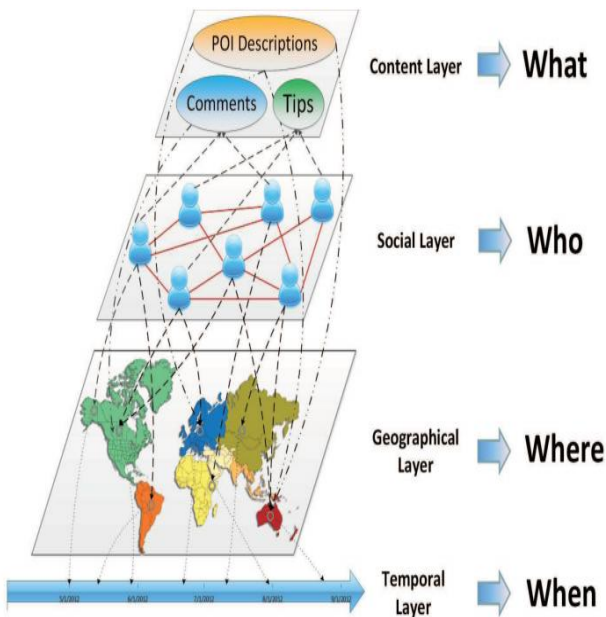


Fig 1: The Information Layout of Location-Based Social Networks

Traditional social network analysis mainly studies network structure and properties without the consideration of geographical distance between nodes. Although the idea of “Death of Distance” proposed in 2011 claims that geographical distance plays a less important role due to the communication revolution and the rapid development of the Internet, which could make of our world a “global village” studies on spatial structure of networks demonstrated that there is a strong correlation between geographical attributes and network properties, indicating the significance of considering the spatial properties of networks for future applications [9]. Researchers have further studied the distinctions between online and offline social networks [7], and discovered that geographical property does play important roles when constructing the social connection between two users especially in explaining their mobility in the physical world [8, 10].

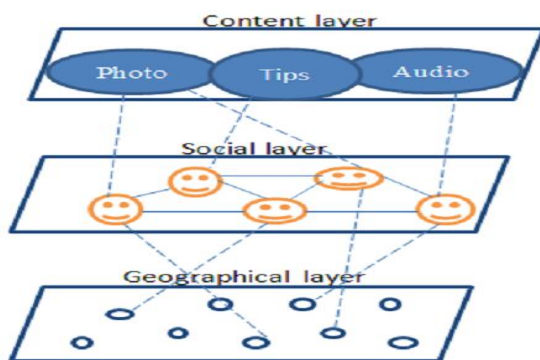


Fig 2: Location Based Social Networks

CONTEXTS-AWARE RECOMMENDER SYSTEMS

Context-Aware recommender systems (CARSSs) extend the traditional formulation of the recommendation problem by incorporating also the context of user-item interactions. Therefore, CARSSs estimate the rating of a target user for an item, not only based on a data set of ratings (of users for items), but they also exploit both the contextual information under which the ratings were acquired and the contextual situation of the target user asking for a recommendation.

More formally, given a set of possible contextual situations S that users can encounter while experiencing items, the rating estimation function is formulated as $\hat{r}: U \times I \times S \rightarrow R$ and estimates the rating of a given user for a given item in a specific contextual situation.

The main assumption supporting CARS techniques is that in many domains users can experience items differently depending on their contextual situation, and thus user’s ratings can also be influenced by the context. For instance, depending on the weather, a user may prefer to visit a beach or a museum. Therefore, CARSs’ goal is to enhance the effectiveness of recommendations by incorporating context into the recommendation process. Depending on how context is exploited, three main types of contextual paradigms can be identified [11].

COLLABORATIVE FILTERING

Collaborative Filtering (CF) techniques [12], [13] base their predictions on the ratings of other users. Differently from CB recommenders, CF approaches are domain-independent in the sense that no items descriptions are required to facilitate generate recommendations, only ratings. This technique emulates a simple but effective social strategy called “word-of-mouth”, which relies on the added credibility of person-to-person recommendations. The basic assumption behind CF is that a user may be interested in items that have been positively rated by other users with similar interests.

2. LITERATURE SURVEY

In this paper [13], here they proposed an author topic model-based collaborative filtering (ATCF) technique is proposed to make easy complete points of interest (POIs) recommendations for social users for personalized travel recommendations. User’s preference themes, i.e. cultural, cityscape, or landmark, are extracted from the geo-tag constrained can be mined from the textual descriptions attached with his/her photos through author topic model (ATM) as an alternative of only from the geo-tags (GPS locations). Through ATM travel topics and a user’s topic preference can be extracted at the same time. In ATCF, POIs are ranked according to related users, who share similar travel topic preferences as a replacement for of raw GPS (geo-tag) data as is the case of earlier works. Unlike location-based collaborative filtering even without GPS records related users can still be mined correctly according to the comparison of users’ topic preferences.

In this paper [14], we present a review of existing POI recommendation algorithms and discuss some research directions for POIs recommendation. According to the type of additional information integrated with check-in data by POI recommendation algorithms, we classify POI recommendation algorithms into four categories: pure check-in data based POI recommendation approaches, geographical influence enhanced POI recommendation approaches, and social influence enhanced POI recommendation approaches and temporal influence enhanced POI recommendation approaches. Pure check-in data based POI recommendation approaches take check-in frequency as ratings and make an assumption that two users are similar if they have checked in a lot of common POIs. Then, conventional collaborative filtering approaches are adopted to make POI recommendations by averaging most similar users’ preferences on candidate POIs.

Despite some research works have studied the problem of POI recommendation in LBSNs, POI recommendation just emerges recently and several interesting research directions

are worthy of exploring. First, in LBSNs, the frequencies of check-in for POIs vary dramatically and users' check-in frequency intuitively reflects the degree of users' preferences for POIs. However, the reviewed works reported that making POI recommendation based on 0/1 rating matrix is better than on check-in frequency matrix. Hence, it is desirable for POI recommender systems to adopt suitable approaches to model check-in frequency data. Rank-based collaborative filtering approaches [15] may be applicable to POI recommendation since rank-based collaborative filtering approaches infer users' preferences from pair wise comparisons rather than numerical ratings. Second, it is reported that social influence enhanced POI recommendation approaches have not achieved important improvements compared with the state-of-art POI recommendation methods. Beside the decision process of POI selection is influenced by the geographical property of POIs, a possible reason is that social influence enhanced POI recommendation approaches take all social relations as homogeneous social connections and ignore different types of social relations.

Garcia et al. identified some features that might be useful for recommending followee [16]. The intuition of the paper was that if a target user has many popular and active followees, other popular and active followees should be recommended to the user. If the target user has only popular followees, only popular followees should be recommended. A similar approach can be applied for target users with active followees. They found that the popularity (the followers and followees count ratio), and the activity of the user (the number of tweets he posted since the creation of his account), are the most relevant features used for recommendation.

Hannon et al. presented Twittomender system that recommends followees using both content-based and collaborative-based approaches [17]. In the content-based approach, users are represented by their own tweets, their followers' tweets, their followees' tweets, or combination of all of them. Recommendation is then made based on the similarity between user and targeted user's tweets. In the collaborative-based approach, the users are represented by followees IDs, followers IDs or combination of them. Each user is then represented by a set of his follower/followee IDs. Then, TF-IDF weighting scheme is used to find users with similar follower/followee IDs.

Zheng et al. proposed a collaborative-based approach to extract the features for the locations, and to provide activity recommendation in LBSNs [18]. They relied on three matrices: location activity matrices, location-feature matrix and activity-activity matrix. The location activity matrix is used to correlate the user's activity to a spatial location. Location-feature matrix is used to connect locations and categories (e.g. restaurants, cafes and Movie Theater). The basic idea in this matrix is that locations of the same category are likely to have the same activity possibilities. The activity-activity matrix shows the correlations between different activities. The probability that certain activity will be performed at a certain location given that a user has performed some other activity can then be predicted.

Yuan et al. developed a collaborative recommendation model to recommend POIs for a given user at a specified time in a day [19]. They defined a new problem, the time-aware POI recommendation, that considers the temporal influence in user activities. In addition to the temporal influence, they also enhanced the recommendation model by considering geographical information and the social influence (i.e. users tend to visit nearby POIs). The authors found that if two users

have similar temporal behavior, they are likely to visit similar POIs at the same time.

3. PROPOSED METHODOLOGY

The Proposed Methodology implemented here works on following stages:

1. Take an input Geo-Tags Sets or User's Photo Set.
2. Mining of POI based on City Level or Author Topics using Recommendation by Popularity.
3. Applying Collaborative Filtering.
4. Recommendation POI using Principle Component Analysis.

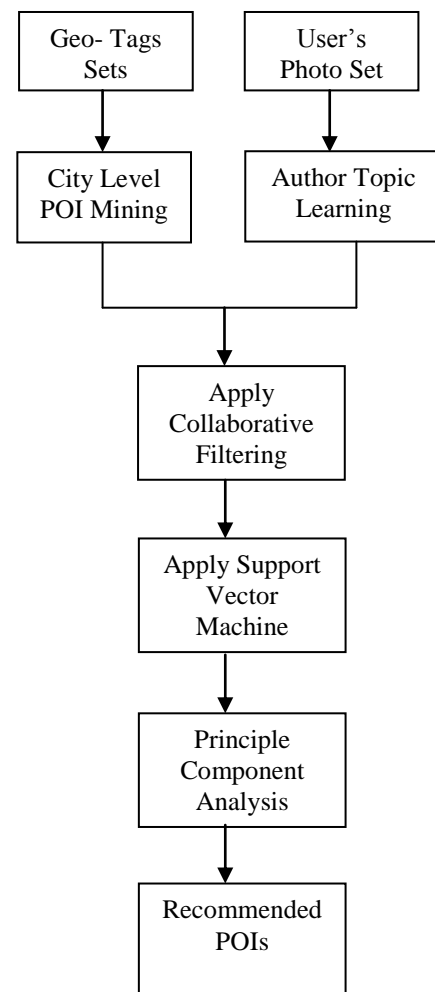


Fig 3: Flow Chart of Proposed Methodology

The proposed methodology implemented here works in 3 stages:

1. *Recommendation by Popularity:* All the POIs of one city are ranked according to degree of popularity. The popularity of each POI is measured by how many users have uploaded photos of this POI. Then to each user, we recommend the top ranked POIs. To all users, the recommendation results are the same ones.
2. *Collaborative Filtering:* Location-based Collaborative Filtering (LCF) is the most popular method that can be easily realized.

First a user-POI matrix is generated to record user travel history. Then for a given two users, their similarity is calculated by the cosine standard measurement towards their corresponding vectors in the user-POI matrix. Finally, locations are recommended based on similar users' visiting histories.

3. Recommendation by PCA With SVM Learning

POI recommendation approach consists of the following two steps: 1) similar user detection, and 2) POI ranking.

Recommendation by Popularity

All the POIs of one city are ranked according to degree of popularity. The popularity of each POI is measured by how many users have uploaded photos of this POI. Then to each user, we recommend the top ranked POIs. To all users, the recommendation results are the same ones.

Collaborative Filtering

Location-based Collaborative Filtering (LCF) is the most popular method that can be easily realized. First a user-POI matrix is generated to record user travel history. Then for a given two users, their similarity is calculated by the cosine standard measurement towards their corresponding vectors in the user-POI matrix. Finally, locations are recommended based on similar users' visiting histories.

Principal Component Analysis

- Probably the most widely-used and well-known of the "standard" multivariate methods.
- Invented by Pearson (1901) and Hotelling (1933).
- First applied in ecology by Goodall (1954) under the name "factor analysis" ("principal factor analysis" is a synonym of PCA).
- Takes a data matrix of n objects by p variables, which may be correlated, and summarizes it by uncorrelated axes (principal components or principal axes) that are linear combinations of the original p variables.
- The first k components display as much as possible of the variation among objects.
- Objects are represented as a cloud of n points in a multidimensional space with an axis for each of the p variables.
- The centroid of the points is defined by the mean of each variable.
- The variance of each variable is the average squared deviation of its n values around the mean of that variable.

$$V_i = \frac{1}{n-1} \sum_{m=1}^n (X_{im} - \bar{X}_i)^2$$

- Degree to which the variables are linearly correlated is represented by their covariances.

$$C_{ij} = \frac{1}{n-1} \sum_{m=1}^n (X_{im} - \bar{X}_i)(X_{jm} - \bar{X}_j)$$

SUPPORT VECTOR MACHINE

Consider training sample $\{(x_i, d_i)\}$, where x_i is the input pattern, d_i is the desired output:

$$W_0^T X_i + b_0 \geq +1, \text{ for } d_i = +1$$

$$W_0^T X_i + b_0 \leq -1, \text{ for } d_i = -1$$

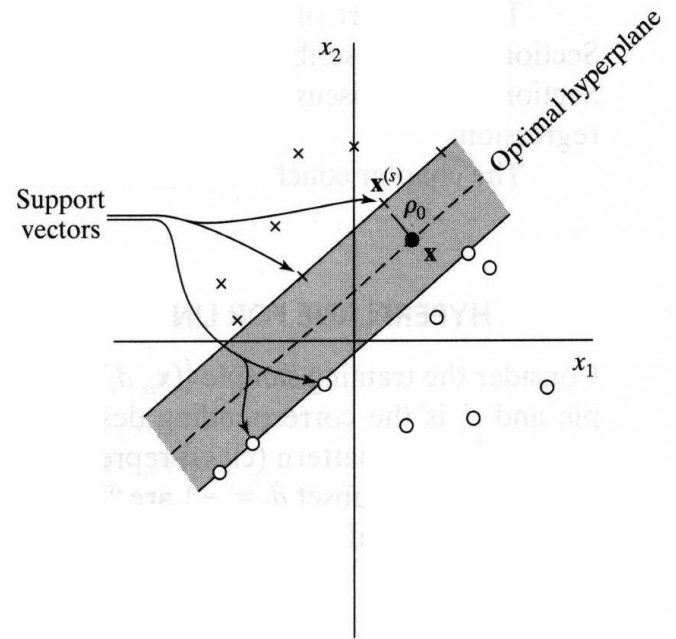


Fig 4: Basic Architecture of SVM

The data point which is very near is called the margin of separation ρ

The main aim of using the SVM is to find the particular hyperplane of which the margin ρ is maximized

Optimal hyperplane $W_0^T X + b_0 = 0$

For example, if we are choosing our model from the set of hyperplanes in R^n , then we have:

$$f(x; \{w; b\}) = \text{sign}(w \cdot x + b)$$

We can try to learn $f(x; _)$ by choosing a function that performs well on training data:

$$R_{emp}(\alpha) = \frac{1}{m} \sum_{i=1}^m l(f(x_i, \alpha), y_i)$$

4. RESULT ANALYSIS

The Table shown below is the analysis of POI Recommendations Performance of various existing Collaborative filtering techniques and the proposed methodology. The Analysis done here is on the basis of different MAP@n is the mean Average Precision in which the proposed methodology has high Recommendation Performance.

$$MAP@n = \frac{(\sum_{i=1}^m AP_i)}{m}$$

$$AP@n = \frac{(\sum_{i=1}^n (\sum_{j=1}^i rel_j) / i)}{n}$$

Where, n is the number of POIs which recommend the users, MAP@n is the mean Average Precision, m is the number of users, rel_i is a relevance value and is equal to 1 if user have actually visited the recommended POI, otherwise 0.

Table 1. Analysis of POI Recommendation Performance

Performance	PO	CF	LDA	ATCF	Proposed
MAP	0.3408	0.4137	0.4166	0.4225	0.4362
MAP@1	0.4861	0.5595	0.5678	0.5876	0.5962
MAP@5	0.3557	0.4312	0.4361	0.4483	0.4537
MAP@10	0.3076	0.4059	0.4005	0.4115	0.4275
MAP@20	0.2642	0.3519	0.3545	0.3545	0.3629
MAP@30	0.2438	0.3151	0.3163	0.3184	0.3264

The Table shown below is the analysis of Precision & Recall of existing Collaborative filtering techniques and the proposed methodology. The Analysis done here is on the basis of different Threshold values in which the proposed methodology has higher Precision & Recall rate.

$$\text{Precision} = \frac{\text{No. of Actual POI Recommendation Filtered}}{\text{Total No. of POI Recommendations}}$$

$$\text{Recall} = \frac{\text{No. of Actual POI Recommendation Filtered}}{\text{Available POI Recommendations in Dataset}}$$

Table 2. Analysis of Precision & Recall of POI mining with respect to Vt.

Threshold	Existing Work		Proposed Work	
	Precision	Recall	Precision	Recall
1	0.38	0.78	0.4	0.8
4	0.5	0.83	0.53	0.85
7	0.74	0.85	0.78	0.87
10	0.65	0.87	0.67	0.89
13	0.63	0.89	0.68	0.91
16	0.6	0.91	0.63	0.93
19	0.57	0.93	0.59	0.95
22	0.54	0.95	0.56	0.97
25	0.54	0.97	0.56	0.99
28	0.54	0.97	0.56	0.99
31	0.54	0.97	0.56	0.99
34	0.54	0.97	0.56	0.99
37	0.54	0.97	0.56	0.99
40	0.54	0.97	0.56	0.99
43	0.54	0.97	0.56	0.99
46	0.54	0.97	0.56	0.99
49	0.54	0.97	0.56	0.99

The Table shown below is the analysis of Mean Average Precision of existing Collaborative filtering techniques and

the proposed methodology. The Analysis done here is on the basis of different no. of POI Available in which the proposed methodology has Mean Average Precision.

Table 3. Analysis of MAP

# of POIs	MAP	
	Existing Work	Proposed Work
5	0.24	0.28
10	0.36	0.39
15	0.38	0.42
20	0.4	0.46

The Figure shown below is the analysis of POI Recommendations Performance of various existing Collaborative filtering techniques and the proposed methodology. The Analysis done here is on the basis of different MAP@n is the mean Average Precision in which the proposed methodology has high Recommendation Performance.

$$\text{MAP@n} = \frac{(\sum_{i=1}^m AP_i)}{m}$$

$$AP@n = \frac{(\sum_{i=1}^n (\sum_{j=1}^i rel_j) / i)}{n}$$

Where, n is the number of POIs which recommend the users, MAP@n is the mean Average Precision, m is the number of users, rel_j is a relevance value and is equal to 1 if user have actually visited the recommended POI, otherwise 0.

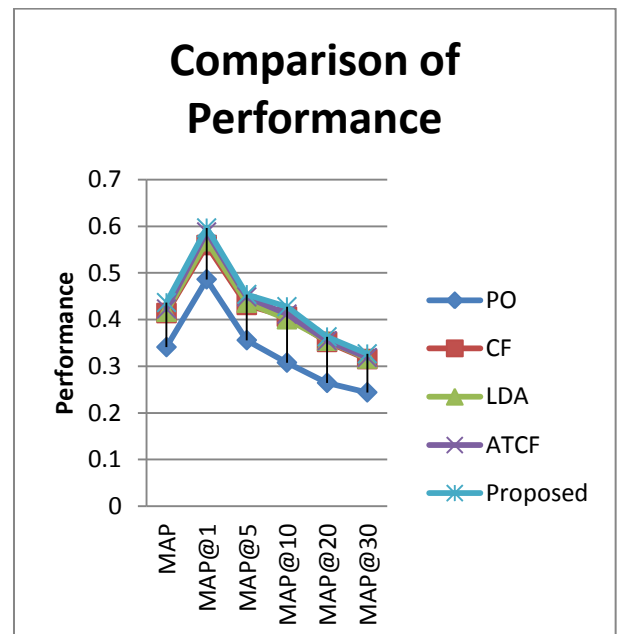


Fig 5: Comparison of POI Recommendation Performance

The Figure shown below is the analysis of Mean Average Precision of existing Collaborative filtering techniques and the proposed methodology. The Analysis done here is on the basis of different no. of POI Available in which the proposed methodology has Mean Average Precision.

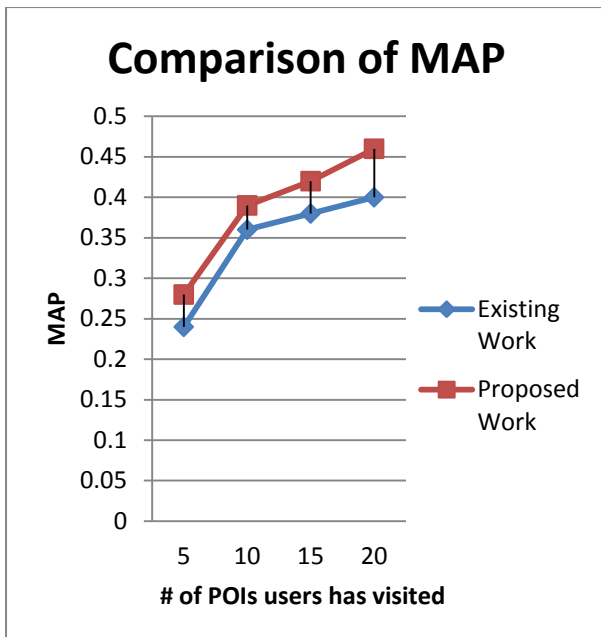


Fig 6: Comparison of POI Recommendation Performance

5. CONCLUSION & FUTURE WORK

Collaborative filtering (CF) is the most common technique of predicting the interests of a user by collecting preference information from many users. In order to determine which items from a collection may be favored by individual users, conventional CF approaches take the ratings previously assigned to items by a target user and use them together with ratings of users with similar preferences to predict the ratings of yet-unseen items.

The proposed methodology implemented for the Personalized POI Recommendations using Principle Component Analysis provides efficient results in comparison to the existing methodology implemented for POI Recommendations.

Although the proposed methodology implemented here for the Collaborative Filtering is efficient, but further enhancements can be done for the improvement of the methodology such as by using some enhanced framework or by implementing Filtering for High Dynamic Ranging Images.

6. REFERENCES

- [1] J. Bao, Y. Zheng, and M. F. Mokbel, "Location-based and preference-aware recommendation using sparse geo-social networking data," in Proc. 20th Int. Conf. Adv. Geographic Inf. Syst., 2012, pp. 199–208.
- [2] T. Kurashima, T. Iwata, G. Irie, and K. Fujimura, "Travel route recommendation using geo-tags in photo sharing sites," in Proc. 19th ACM Int. Conf. Inf. Knowl. Manage., 2010, pp. 579–588.
- [3] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in Proc. 10th Int. Conf. WWW, 2001, pp. 285–295.
- [4] H. Feng and X. Qian, "Mining user-contributed photos for personalized product recommendation," Neurocomput., vol. 129, pp. 409–420, 2014
- [5] X. Qian, H. Feng, G. Zhao, and T. Mei, "Personalized recommendation combining user interest and social circle," IEEE Trans. Knowl. Data Eng., vol. 26, no. 7, pp. 1763–1777, Jul 2014
- [6] K. Zickuhr. Three-quarters of smartphone owners use location-based services. Pew Internet & American Life Project, May 11, 2012.
- [7] J. Cranshaw, E. Toch, J. Hong, A. Kittur, and N. Sadeh. Bridging the gap between physical location and online social networks. In Proceedings of the 12th ACM international conference on Ubiquitous computing, pages 119–128. ACM, 2010.
- [8] E. Cho, S. Myers, and J. Leskovec. Friendship and mobility: user movement in location-based social networks. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pages 1082–1090. ACM, 2011.
- [9] M. Gastner and M. Newman. The spatial structure of networks. The European Physical Journal B-Condensed Matter and Complex Systems, 49(2):247–252, 2006.
- [10] S. Scellato, A. Noulas, R. Lambiotte, and C. Mascolo. Socio-spatial properties of online location-based social networks. Proceeding of the 5th International AAAI Conference on Weblogs and Social Media, 11, 2011.
- [11] Adomavicius, G., & Tuzhilin, A. (2011). Context-aware recommender systems. In Recommender Systems Handbook (pp. 217–256).
- [12] Koren, Y., & Bell, R. Advances in collaborative filtering. Recommender Systems Handbook, 145–186, 2011.
- [13] Shuhui Jiang, Xueming Qian, Jialie Shen, "Author Topic Model-Based Collaborative Filtering for Personalized POI Recommendations" IEEE Transactions On Multimedia, Vol. 17, No. 6, June 2015.
- [14] Yonghong Yu, Xingguo Chen, "A Survey of Point-of-Interest Recommendation in Location-Based Social Networks" Trajectory-Based Behavior Analytics: Papers from the 2015.
- [15] Yi, J.; Jin, R.; Jain, S.; and Jain, A. Inferring user's preferences from crowd sourced pair wise comparisons: A matrix completion approach. In First AAAI Conference on Human Computation and Crowd sourcing, 2013.
- [16] A. X. Garcia, R. Weighted content based methods for recommending connections in online social networks. In In: The 2nd ACM Workshop on Recommendation Systems and the Social Web, Barcelona, Spain, June 2010.
- [17] J. Hannon, K. McCarthy, and B. Smyth. Finding useful users on twitter: twittomender the followee recommender. In Proceedings of the 33rd European conference on Advances in information retrieval, ECIR'11, pages 784–787, 2011.
- [18] V. W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative location and activity recommendations with gps history data. In Proceedings of the 19th international conference on World wide web, WWW '10, pages 1029–1038, New York, NY, USA, 2010. ACM.
- [19] Q. Yuan, G. Cong, Z. Ma, A. Sun, and N. M. Thalmann. Time-aware point-of-interest recommendation. In Proceedings of the 36th international ACM SIGIR conference on Research and development in information retrieval, SIGIR '13, pages 363–372, New York, NY, USA, 2013.