Journal of Engineering and Applied Sciences 12 (Special Issue 9): 8714-8718, 2017
ISSN: 1816-949X

A Study on the Effect of Location-User Distances for Location Based Recommendation Systems

Wadhah R. Baiee
Department of Software, College of IT, University of Babylon, Hillah, Iraq

Abstract: Location based recommendation systems are used with many applications that support maps and user's interested locations. Recommendations must be accurate and near to user's needs, especially when many applications depend on maps and locations. This study studies the effect of distances between users and their location preferences, combined with our previous work 3D-SVDs algorithm. Here, Manhattan distance is used as a measurement's feature between location blocks against user's location, this technique is injected into our recommendation algorithm and evaluated by using information retrieval recall and precision for different distances. The results are differed according to location distance and the recommendation lists are varied between the accuracy and diversity.

Key words: GPS trajectories, Manhattan distance, location based recommendation, 3D-SVDs, evaluation, precision, recall

INTRODUCTION

The recommendation systems are different in multiple ways, according to their prediction level or recommended items which they support. Also, the type of recommendations and algorithms that are used in such systems make the comparison among them is one of the challenges in this field of research. Many applications are developed for that reason and datasets become vast and varied, these data must be organized, analyzed and processed to be meaningful to the user. Traveling is a significant field of mobile apps and an implausible number of services are now acquired to serve the users while they travel. It is essential to distinguish the abilities of this area of science and study the performance of mobile's users (Ricci *et al*., 2011). Trajectories of Global Positioning System (GPS) have offered exclusive information to realize moving items and places, calling for consistent research and development of different computing procedures to process, recover and mine trajectory data and determining its applications (Lee and Krumm, 2011).

Baltrunas *et al*. (2012) take a different method for modeling the association among related features and item-user ratings. Instead of using the traditional method to collecting needed data, they mimic contextual situations to more merely get data regarding how the context affects ratings. Zheng *et al*. (2010), they model the user's place and motion histories that are reserved as feedback to the system, they mine data such as the location features and activities interactions from the web and databases of GIS. Mac *et al*. (2009), this research presents a method that detects user's action and produces a profile for each user to reflect her requirements based on the user interaction with application, physical site and user actions. The system identifies the profile of users and adjusts settings to deliver suitable information. Savage *et al*. (2012) develop more comprehensive general algorithm for map based recommendation by gathering user's favorites and take into account topography of time and measurements features. This study studies the effect of the distances of the positions of interest that is called in recommendation systems places to the current location of the user, throw the usage of our previous research 3D-SVDs algorithm (Baiee, 2016a, b). The system used GeoLife Trajectories data by Zheng *et al*. (2008, 2009, 2010). This GPS dataset was collected in Microsoft Research Asia, it contains 182 users throw 5 years (2007-2012). GPS trajectory of this dataset is symbolized by point's sequences, it has the data of longitude, latitude and altitude. The results has been tested by using C#.NET and ArcMap GIS object oriented programming.

MATERIALS AND METHODS

Manhattan distance: It is a common distance measure and a special case of Euclidean. Here, the summation of the amounts of the differences in every dimension can be defined as the distance Manhattan distance between two points. The name "Manhattan distance" came from the distance can be calculated when a person would have to move between points in a way that he was constrained to

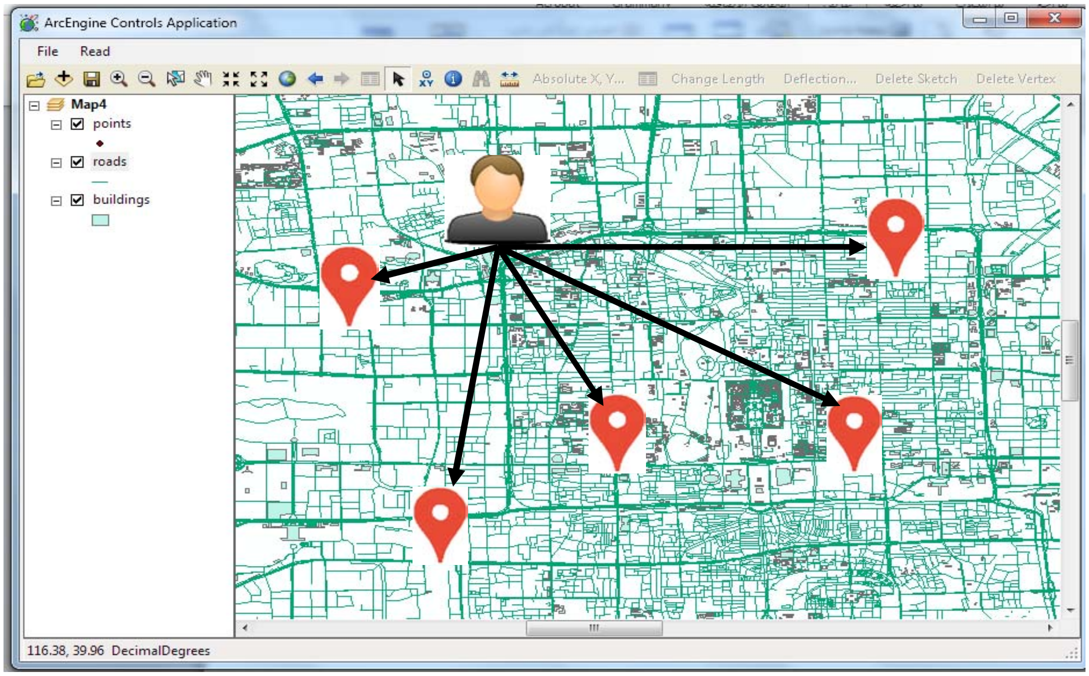

Fig. 1: User location against POI

move within grid lines, exactly like moving in the streets of a city such as Manhattan (Rajaraman and Ullman, 2011). The distance between the user x and y is defined by d (x, y) and is computed as follows (Ricci *et al*., 2011):

$$d(x, y) = \sum_{i=0}^{n} |x_i - y_i| \quad (1)$$

Recommendation of position of interest POIs: This part of the study is to discovering the list of recommended items of the current user. From our previous research (Baiee, 2016a, b), there are ordered vectors of similarities among users against the target user, these vector is used to find the nearest neighbor users to the current user, thus, after applying the most frequent item recommendation technique. Then, the system will have the most visited places for each similar user then the system can sort the resulted preference places to find the top N values which are the recommended places for the current user. The main goal is listing recommendation to the user from nearest users. The choice of the POIs depends on two main criteria:

- The relation of current user to other users and this will produced by features that extracted from rating matrix and user histories
- The current location of the user which has been got from the GPS sensor in his mobile device. Figure 1 illustrates distance between a user and POI

These two features should combined together to have more accurate and enhanced recommendations in the proposed system. This combination will produce recommended POIs near the user and the word near means after applying this technique it means two meanings: first is the places that other users who have high similarity to the current user are more likely to the current user as like as them, second, the near places should be taken into account to filter and recommend a really near places to the current user according to his current location.

The second criteria as mentioned above are calculated using Manhattan distance. The system uses Manhattan distance to calculate the similarity of current location of the user and the locations that are more likely to the users and produced from the most frequent item recommendation. Manhattan distance is used because the system depends on a map that is segmented to regions rows and columns. Thus, this distance measure is the right tool for this type of data illustrated as:

Algorithm of recommendation (Baiee, 2016a, b):

Inputs: Ordered list of similar users, map, user location
Outputs: List of recommended POIs near the user
Process:
Begin
    Get point p from user location
    For each period of time do
        For each similar user u_i in (List of similar users) do
            Sort the most visited places by u_i in descent
            Take the top 10 most frequent places
        For each near place do
            Dist←Manhattan distance from p to top 10 places
            Score←Rates of most visit places
          If this places (Dist<v1) and (Score>v2)
            Add this place to list placeList with dist and score
    Sort placeList descending for score and dist values
    Take top 10 places from placeList
    Draw their locations on map mobile application
End

RESULTS AND DISCUSSION

System evaluation against different distances: The proposed system is evaluated by calculating the precision and recall for recommended lists of POIs for each one of users against their preferences. The training set is the first 11 periods of the 3D-SVDs (Baiee, 2016a, b) and the last period is the test set, it represents the future movements of each user as well as their future preferences.

The precision and recall (Arora *et al*., 2016) are calculated gradually according to the number of items that are recommended to the user in the system. Table 1 shows the calculation of system recall at three level of Manhattan distance that around the user to resort the recommended items to the user according to his current place as mentioned previously.

There are some observations here while the distance value becomes large, the recommended list will have more matched items in the test period. In the same manner if the distance value and the filtering area become smaller, the recommended list will have new items to be recommended to the user and the recall will decrease. Thus, there is a relation can be concluded here between the Manhattan distance and the recall value of the system evaluation against matched and new items which are recommended to the user. Figure 2 illustrates the chart of recalls from Table 1 that is calculated for three different values of Manhattan distance.

The second level of evaluation is the graduated precision recall evaluation for different number of recommended items among the users then the average will be calculated for the system at all. The number of items are differ in each iteration of the test, thus, the system is evaluated with N recommended items equal to 1000 and descending to N = 5 items. The average recall and precisions are calculated for all system users in each test. Table 2 shows recalls and precisions according to the number of recommended items. Figure 3 illustrates the information retrieval recall and precision chart. The average value for recalls is 46.5% and the average precision is 30% for the proposed system.

Table 1: Recall values of the proposed recommended system for three different values of Manhattan distance from user's current location

User ID	Recall (large distance)	Recall (medium distance)	Recall (small distance)
1	0.98	0.798206278	0.342105263
2	0.98	0.772972973	0.777777778
3	0.944444444	0.840909091	0.366666667
4	0.97	0.536046512	0.223367698
5	0.96	0.403587444	0.352631579
6	0.9743	0.916666667	0.714285714
8	0.913793103	0.546296296	0.448275862
9	0.99453	0.645945946	0.487179487
10	0.9004	0.823076923	0.727272727
11	0.902272727	0.494117647	0.170454545
12	0.989583333	0.5678	0.291666667
14	0.978021978	0.769230769	0.362637363
15	0.975778547	0.615606936	0.235294118
16	0.96534	0.90625	0.423076923
17	0.9345	0.701492537	0.432098765
18	0.985355649	0.714566929	0.261506276
19	0.876190476	0.646341463	0.066666667
20	0.92344	0.565217391	0.23655914
23	0.962406015	0.601587302	0.338345865
25	0.978609626	0.505030181	0.090909091
26	0.994366197	0.715517241	0.233802817
29	0.985294118	0.708333333	0.25
30	0.9345	0.642105263	0.064516129
31	0.987755102	0.604395604	0.257142857
33	0.971014493	0.665517241	0.144927536
35	0.929078014	0.717910448	0.290780142
36	0.9442	0.897058824	0.8
37	0.990909091	0.622222222	0.227272727
38	0.94567	0.561151079	0.166666667
39	0.939086294	0.9375	0.314720812
40	0.974576271	0.491150442	0.279661017
Average	0.957594048	0.675284225	0.334782868

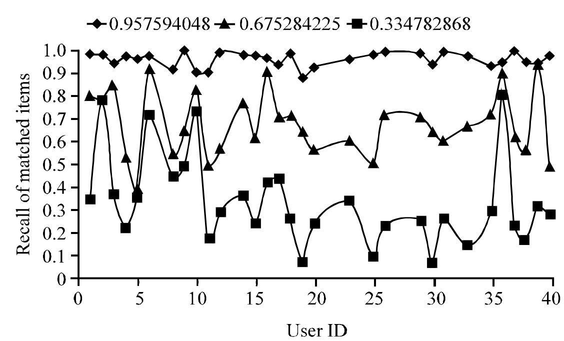

Fig. 2: Chart of recommendation average recalls for each user with three different Manhattan distance values

Table 2: Recall values of the proposed recommended system for three different values of Manhattan distance from user's current location

Recommended items count	Avgerage recall for system users	Avgerage precision for system users
1000	0.755729616	0.09255
900	0.732057639	0.098888889
800	0.708213933	0.10634375
700	0.685283983	0.116107143
600	0.651958633	0.126875
500	0.621419884	0.1418
400	0.576505480	0.1605625
300	0.513718113	0.185083333
200	0.447942712	0.227375
100	0.339005098	0.3135
80	0.314615401	0.35375
50	0.260413273	0.4325
25	0.185428150	0.535
10	0.105744736	0.6825
5	0.073309935	0.94
Average	0.464756439	0.300855708

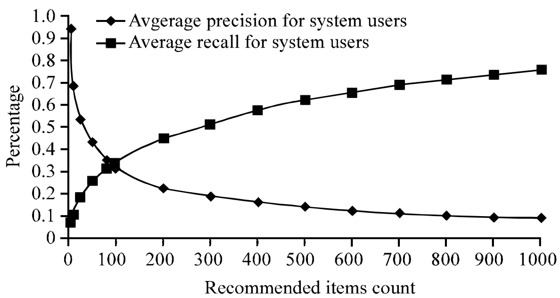

Fig. 3: Chart of recommendation average recalls and precisions for different number of recommended items

CONCLUSION

The recommendations are produced from the combinations of historical movement's behavior of the users that are extracted by 3D-SVDs (Savage *et al*., 2012; Baiee, 2016a, b) and the current location of user who uses the system to avoid the redundant results and to be more accurate recommended POIs to him.

The system is evaluated by using information retrieval recall and precision are calculated for the entire system, the evaluation is gradually calculated according to the number of recommended items in the list of recommendations. The values of the recall are between 7-75% and the average is 46.5%, the values of precision are between 9-94% and the average is 30% for the entire system.

Finally, the effect of Manhattan distance is evaluated for the system where the distance has large value, the recommended items match user needs and system recall becomes 95.7% with medium distance recall value is around 67%, finally if the distance becomes small the recall reduces to 33.4%. While the system recalls being smaller in such case the system will recommend new items to the user and vice versa.

REFERENCES

Arora, M., U. Kanjilal and D. Varshney, 2016. Evaluation of information retrieval: Precision and recall. Intl. J. Indian Cult. Bus. Manage., 12: 224-236.

Baiee, W.R., 2016a. A new technique for generating implicit feedbacks from users movements history. Proceedings of the International Conference on Change, Innovation, Informatics and Disruptive Technology ICCIIDT'16, October 11-12, 2016, American Research Foundation, London, England, UK., pp: 393-406.

Baiee, W.R., 2016b. Design 3D-SVDs algorithm for location based recommendation system. Res. J. Appl. Sci., 10: 1143-1150.

Baltrunas, L., B. Ludwig, S. Peer and F. Ricci, 2012. Context relevance assessment and exploitation in mobile recommender systems. Personal Ubiquitous Comput., 16: 507-526.

Lee, W.C. and J. Krumm, 2011. Trajectory Preprocessing. In: Computing with Spatial Trajectories, Yu, Z. and X. Zhou (Eds.). Springer, New York, USA., ISBN: 978-1-4614-1628-9, pp: 3-33.

Mac, A.E., G. McArdle, M. Petit, C. Ray and M. Bertolotto *et al*., 2009. Personalization in adaptive and interactive GIS. Annal. GIS., 15: 23-33.

Rajaraman, A. and J.D. Ullman, 2011. Mining of Massive Datasets. Cambridge University Press, UK., ISBN-13: 978-1107015357, Pages: 326.

Ricci, F., L. Rokach and B. Shapira, 2011. Introduction to Recommender Systems Handbook. In: Recommender Systems Handbook, Ricci, F., L. Rokach, L. Shapira and P. Kantor (Eds.). Springer, Berlin, Germany, ISBN:978-0-387-85819-7, pp: 1-35.

Savage, N.S., M. Baranski, N.E. Chavez and T. Hollerer, 2012. I'm Feeling Loco: A Location Based Context Aware Recommendation System. In: Advances in Location-Based Services, Gartner, G. and F.Ortag (Eds.). Springer, Berlin, Germany, ISBN:978-3-642-24197-0, pp: 37-54.

Zheng, V.W., Y. Zheng, X. Xie and Q. Yang, 2010. Collaborative location and activity recommendations with GPS history data. Proceedings of the 19th International Conference on World Wide Web, April 26-30, 2010, ACM, Raleigh, North Carolina, ISBN:978-1-60558-799-8, pp: 1029-1038.

Zheng, Y., L. Zhang, X. Xie and W.Y. Ma, 2009. Mining interesting locations and travel sequences from GPS trajectories. Proceedings of the 18th International Conference on World Wide Web, April 20-24, 2009, ACM, Madrid, Spain, ISBN:978-1-60558-487-4, pp: 791-800.

Zheng, Y., Q. Li, Y. Chen, X. Xie and W. Ma, 2008. Understanding mobility based on GPS data. Proceedings of the 10th International Conference on Ubiquitous Computing UbiComp '08, September 21-24, 2008, ACM, Seoul, Korea, ISBN:978-1-60558-136-1, pp: 312-321.

Zheng, Y., X. Xie and W.Y. Ma, 2010. GeoLife: A collaborative social networking service among user location and trajectory. IEEE. Data Eng. Bull., 33: 32-39.