Geo-social Recommendations

Alexis Papadimitriou
Aristotle University
Thessaloniki, Greece
apapadi@csd.auth.gr

Panagiotis Symeonidis
Aristotle University
Thessaloniki, Greece
symeon@csd.auth.gr

Yannis Manolopoulos
Aristotle University
Thessaloniki, Greece
manolopo@csd.auth.gr

ABSTRACT

Social networks have evolved with the combination of geographical data, into Geo-social networks (GSNs). GSNs give users the opportunity, not only to communicate with each other, but also to share images, videos, locations, and activities. The latest developments in GSNs incorporate the usage of location tracking services, such as GPS to allow users to "check-in" at various locations and record their experience. In particular, users submit ratings or personal comments for their location/activity. The vast amount of data that is being generated by users with GPS devices, such as mobile phones, needs efficient methods for its effective management. In this paper, we have implemented an online prototype system, called GeoSocial Recommender System, where users can get recommendations on friends, locations and activities. In order to provide recommendations, we represent this data by a 3-order tensor, on which latent semantic analysis and dimensionality reduction is performed using the Higher Order Singular Value Decomposition (HOSVD) technique. Also, as more data is accumulated to the system, we use incremental solutions to update our tensor. We perform an experimental evaluation of our method with a real data set and measure its effectiveness through recall/precision.

Keywords

tensor, geographical, social, geo-social, recommendations

1. INTRODUCTION

Over the past few years, social networks have attracted a huge attention after the widespread adoption of Web 2.0 technology. Social networks combined with geographical data, have evolved into Geo-social networks (GSNs). GSNs such as Facebook Places, Foursquare.com, etc., which allow users with mobile phones to contribute valuable information, have increased both in popularity and size. These systems are considered to be the next big thing on the web [3]. An interesting statistic is that more than 250 million users are daily accessing Facebook through their mobile devices and they are twice as active than non-mobile users.

GSNs allow users to use their GPS-enabled device, to "check-in" at various locations and record their experience. In particular, users submit ratings or personal comments for the location/activity they visited/performed. That is, they "check-in" at various places, to publish their location online, and see where their friends are. These GSN systems, based on a user's "check-in" profile, can also provide activity and location recommendations. For an activity recommendation, if a user plans to visit some place, the GSN system can recommend an activity (i.e. dance, eat, etc.). For a location recommendation, if a user wants to do something, the GSN system can recommend a place to go. Recently, Zheng et al. [9] proposed a User Collaborative Location and Activity Filtering (UCLAF) system, which is based on Tensor decomposition. However, as the authors claim, they do not update their system online as more users accumulate data continuously over time. Moreover, even though their system provides location and activity recommendations to users, it does not consider the case of providing also friend recommendations.

Our prototype system GeoSocial is an online recommender system that relies on user "check-ins" to provide friend, location and activity recommendations. The "check-in" procedure involves selecting the location he is currently at, the activity he is performing there, and finally rating that activity. Based on the users' "check-in" history and friendship network, GeoSocial provides friend, location and activity recommendations. Friends are recommended based on the Friendlink algorithm presented in [4] and the geographical distances between user "check-ins", which are used as link weights. Users, locations and activities are also inserted into a 3-order tensor, which is then used to provide location and activity recommendations.

The remainder of this paper is organized as follows. Section 2 summarizes the related work, whereas Section 3 describes the GeoSocial recommender system and its components. Section 4 explains the main steps that are followed when performing the tensor reduction to detect latent associations between the user, location and activity dimensions and also the way we update the tensor data by implementing the Incremental Tensor Reduction (ITR) algorithm. In Section 5 we study the performance of ITR and Friendlink in terms of friend, location and activity recommendations. Finally, Section 6 concludes the paper and proposes possible future work.

ACM PeMA '11 October 27, 2011. Chicago, IL, USA

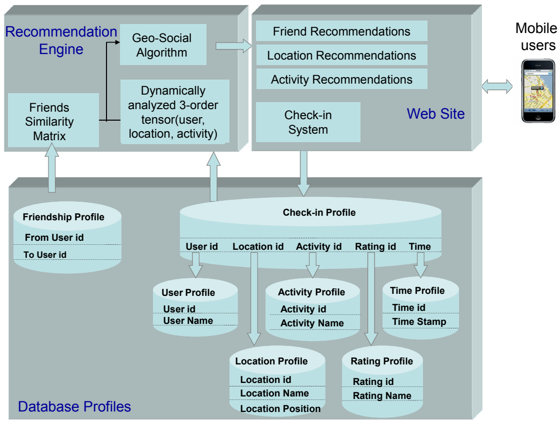


Figure 1: Components of the Geo-social recommender system.

2. RELATED WORK

Recently emerged GSNs (i.e. Gowalla.com, Foursquare.com, Facebook Places etc.) provide to users activity or location recommendation. For example, in Gowalla.com a target user can provide to the system the activity he wants to do and the place he is (e.g. coffee in New York). Then, the system provides a map with coffee places which are nearby the user's location and were visited many times from people he knows. Moreover, Facebook Places allows users to see where their friends are and share their location in the real world.

Scellato et al. [7] proposed a graph analysis based approach to study social networks with geographic information. They also applied new geo-social metrics to four large-scale online Social Network data sets (i.e. Liveljournal, Twitter, FourSquare, BrightKite). Quercia et al. [5] address the mobile cold-start problem when recommending social events to users without any location history. Zheng et al. [12] proposed a personalized friend and location recommender for the Geographical Information Systems (GIS) on the Web, as well as a framework, namely "Hierarchical-graph-based similarity measurement (HGSM)" to uniformly model each individual's location history and effectively measure the similarity among users. Finally, Zheng et al. [11] perform two types of travel recommendations by mining multiple users' GPS traces. The first is a generic one that recommends a user with top interesting locations and travel sequences in a given geospatial region. The second is a personalized recommendation that provides an individual with locations matching her travel preferences.

Moreover, there are tensor-based approaches. For example, Biancalana et al. [1] implemented a social recommender system based on a tensor that provides points of interest (POI) recommendations. Furthermore, Zheng et al. [10] proposed a method, where geographical data is combined with social data to provide location and activity recommendations. Moreover, Zheng et al. [9] proposed a User Collaborative Location and Activity Filtering (UCLAF) system, which is based on Tensor decomposition.

In contrast to the aforementioned tensor-based methods, our GeoSocial recommender system provides (i) location and activity recommendations (ii) friend recommendations by combining Friendlink algorithm proposed in [4] with the geographical distance between users. Moreover, our tensor method includes an incremental stage, where newly created data is inserted into the tensor by incremental solutions [6, 2].

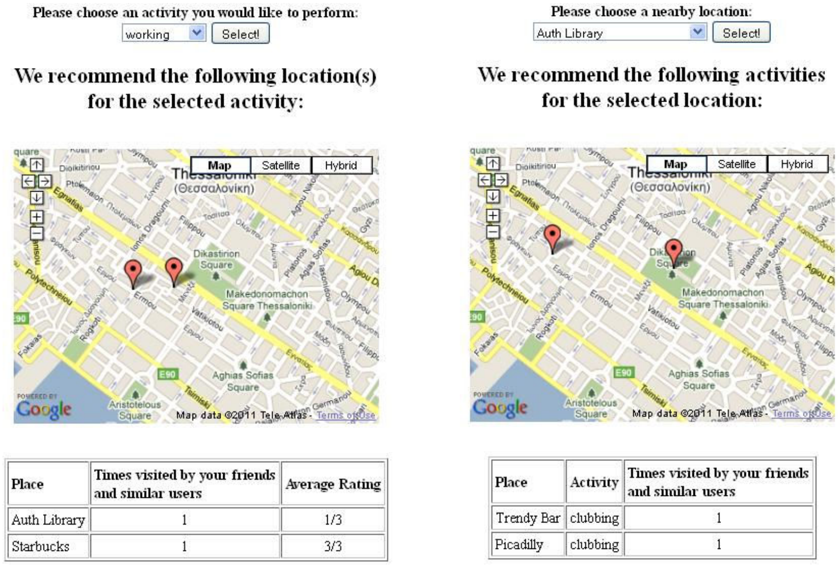
3. GEOSOCIAL SYSTEM DESCRIPTION

Our GeoSocial system consists of several components. The system's architecture is illustrated in Figure 1, where three main sub-systems are described: (i) the Web Site, (ii) the Database Profiles and (iii) the Recommendation Engine. In the following sections, we describe each sub-system of GeoSocial in detail.

3.1 GeoSocial Web Site

The GeoSocial system uses a web site [1] to interact with the users. The web site consists of four sub-systems: (i) the friend recommendation, (ii) the location recommendation, (iii) the activity recommendation and (iv) the "check-in" system. The friend recommendation sub-system is responsible for evaluating incoming data from the Recommendation Engine of GeoSocial and providing updated friend recommendations. In order to provide such recommendations, the web site sub-system implements the Friendlink algorithm presented in [4] and also considers the average geographical distance between each pair of users based on their "check-in" points. The same applies to the location and activity recommendation sub-systems where new and updated location and activity recommendations are presented to the user as new "check-ins" are stored in the Database profiles. Finally, the "check-in" system is responsible for passing the data inserted by the users to the respective Database profiles. Figure 2a shows a location recommendation while Figure 2b depicts an activity recommendation. As shown in Figure 2a, the user selects an activity that he would like to perform, in this case working, and the system provides location recommendations where he could perform his selected activity, in this case either Starbucks or the Aristotle University of Thessaloniki (Auth) Library. As shown in Figure 2b, the user selects a nearby location, i.e. Auth Library and the system provides activities that he could perform. In this case the user's location is near the Auth Library and the system proposes clubbing at the "Trendy bar" or the "Picadily" as possible activities.

[1]http://delab.csd.auth.gr/geosocial



(a) (b)

Figure 2: Location and activity recommendations made by the GeoSocial recommender system.

3.2 GeoSocial Database Profiles

The database profile sub-system contains five profiles where data about the users, locations, activities and their corresponding ratings is stored. As shown in Figure 1, this data is received by the "Check-In" profile and along with the Friendship profile, they provide the input for the Recommendation Engine sub-system.

3.3 GeoSocial Recommendation Engine

The recommendation engine is responsible for collecting the data from the database and producing the recommendations which will then be displayed on the web site. As shown in Figure 1, the recommendation engine constructs a friends similarity matrix by implementing the Friendlink algorithm proposed in [4]. The geographical distances between user "check-ins" are used as link weights. It also produces a dynamically analyzed 3-order tensor, which is firstly constructed by the HOSVD algorithm and is then updated using incremental methods [6, 2], both of which are explained in later sections.

4. OUR INCREMENTAL TENSOR REDUCTION APPROACH

Our Tensor Reduction algorithm initially constructs a tensor, based on usage data triplets $\{u, l, a\}$ of users, locations and activities. The motivation is to use all three entities that interact inside a geo-social system. Consequently, we proceed to the unfolding of $\mathcal{A}$, where we build three new matrices. Then, we apply SVD in each new matrix. Finally, we build the core tensor $\mathcal{S}$ and the resulting tensor $\hat{\mathcal{A}}$. All these can be summarized in 6 steps, which we describe as follows (A more detailed description of the method can be found in [8]):

1. The initial construction of tensor $\mathcal{A}$.
2. Matrix unfolding of tensor $\mathcal{A}$.
3. Application of SVD in each matrix.
4. The core tensor $\mathcal{S}$ construction.
5. The tensor $\hat{\mathcal{A}}$ construction.
6. The generation of the location/activities suggestions.

The reconstructed tensor $\hat{\mathcal{A}}$ measures associations among users, locations and activities, so that the elements of $\hat{\mathcal{A}}$ represent a quadruplet $\{u,\ l,\ a,\ p\}$ where p is the likeliness that user u will visit location l and perform activity a. Therefore, locations/activities can be recommended to u according to their weights associated with $\{u,\ a\}$ and $\{u,\ l\}$ pairs, respectively.

4.1 Inserting new users, locations, or activities over time

As new users, locations, or activities are being introduced to the system, the $\hat{\mathcal{A}}$ tensor, which provides the recommendations, has to be updated. The most demanding operation for this task is the updating of the SVD of the corresponding unfoldings. We can avoid the costly batch recomputation of the corresponding SVD, by considering incremental solutions [6, 2]. Depending on the size of the update (i.e., number of new users, locations, or activities), different techniques have been followed in related research. For small update sizes we can consider the *folding-in* technique [6], whereas for larger update sizes we can consider Incremental SVD techniques [2].

5. EXPERIMENTAL CONFIGURATION

In this Section, we study the performance of our approach in terms of friend, location and activity recommendations. To evaluate the aforementioned recommendations we have chosen a real data set from our newly developed site. There are 1,173 triplets in the form user–location–activity. To these triplets correspond 102 users, 46 locations and 18 activities.

The numbers c_1, c_2, and c_3 of left singular vectors of matrices $U^{(1)}$, $U^{(2)}$, $U^{(3)}$ for our approach, after appropriate tuning, are set to 25, 12 and 8. Due to lack of space we do not present experiments for the tuning of c_1, c_2, and c_3 parameters. The core tensor dimensions are fixed, based on the aforementioned c_1, c_2, and c_3 values.

5.1 Evaluation Metrics

We perform 4-fold cross validation and the default size of the training set is 75% – we pick, for each user, 75% of his "check-ins" and friends randomly. The task of all three recommendation types (i.e. friend, location, activity) is to predict the friends/locations/activities of the user's 25% remaining "check-ins" and friends, respectively. As performance measures we use precision and recall, which are standard in such scenarios. For a test user that receives a list of N recommended friends/locations/activities (top-N list), the following are defined:

– ***Precision*** is the ratio of the number of relevant friends/locations/activities in the top-N list relative to N.

– ***Recall*** is the ratio of the number of relevant friends/locations/activities in the top-N list relative to the total number of relevant friends/locations/activities, respectively.

5.2 Comparison Results

In this Section, we study the accuracy performance of our method in terms of precision and recall. We examine the top-N ranked list, which is recommended to a test user, starting from the top friend/location/activity. Figure 3 shows a precision versus recall curve. As shown, activity recommendations are more accurate than location recommendations. A possible explanation could be the fact that the number of locations is bigger than the number of activities. It is therefore easier to make an accurate activity prediction than a location prediction. Notice also that for the task of friend recommendation, the performance of Friendlink is not so high. The main reason is data sparsity. In particular, we have calculated that the friendship network has average nodes' degree equal to 2.7 and average shortest distance between nodes 4.7. This means that the friendship network can not be considered as a "small world" network, which results to lower accuracy for the friend recommendation task.

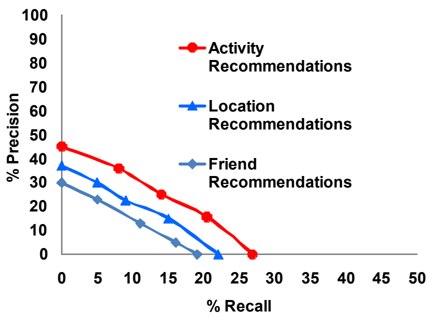


Figure 3: Precision Recall diagram of ITR for activity, location and friend recommendations

6. CONCLUSION AND FUTURE WORK

In this paper we have proposed a Geo-social recommender system which is capable of recommending friends, locations and activities. We used a tensor, which is updated by incremental tensor approaches, as new users, locations, or activities are being inserted into the system.

As future work, we plan on conducting a user study concerning the recommendations in our Geo-social web site to measure user satisfaction. We are also planning on comparing our method to other state-of-the-art methods in terms of effectiveness and efficiency.

References

[1] C. Biancalana, F. Gasparetti, A. Micarelli, and G. Sansonetti. Social tagging for personalized location-based services. In *2nd International Workshop on Social Recommender Systems*, 2011.

[2] M. Brand. Incremental singular value decomposition of uncertain data with missing values. In *European Conference on Computer Vision (ECCV2002)*.

[3] Economist. A world of connections: A special report on networking. *The Economist: Editorial Team*, 2010.

[4] A. Papadimitriou, P. Symeonidis, and Y. Manolopoulos. Predicting links in social networks via bounded local path traversal. In *Proceedings of the 3rd Conference on Computational Aspects of Social Networks (CASON'2011)(to appear)*, 2011.

[5] D. Quercia, N. Lathia, F. Calabrese, G. Di Lorenzo, and J. Crowcroft. Recommending Social Events from Mobile Phone Location Data. In *Proceedings of IEEE ICDM 2010*, Dec. 2010.

[6] B. Sarwar, J. Konstan, and J. Riedl. Incremental singular value decomposition algorithms for highly scalable recommender systems. In *International Conference on Computer and Information Science*, 2002.

[7] S. Scellato, C. Mascolo, M. Musolesi, and V. Latora. Distance Matters: Geo-social Metrics for Online Social Networks. In *Proceedings of the 3rd Workshop on Online Social Networks (WOSN 2010)*, June 2010.

[8] P. Symeonidis, A. Nanopoulos, and Y. Manolopoulos. A unified framework for providing recommendations in social tagging systems based on ternary semantic analysis. *IEEE Transactions on Knowledge and Data Engineering*, 22(2):179–192, 2010.

[9] V. Zheng, B. Cao, Y. Zheng, X. Xie, and Q. Yang. Collaborative filtering meets mobile recommendation: A user-centered approach. In *Proceedings of the international Conference on Artificial Intelligence (AAAI'10)*, 2010.

[10] W. Zheng, Y. Zheng, X. Xie, and Q. Yang. Collaborative location and activity recommendations with gps history data. In *WWW '10: Proceedings of the 19th international conference on World wide web*, pages 1029–1038, New York, NY, USA, 2010. ACM.

[11] Y. Zheng and X. Xie. Learning travel recommendations from user-generated gps traces. *ACM Trans. Intell. Syst. Technol.*, 2:2:1–2:29, January 2011.

[12] Y. Zheng, L. Zhang, Z. Ma, X. Xie, and W.-Y. Ma. Recommending friends and locations based on individual location history. *ACM Trans. Web*, 5:5:1–5:44, February 2011.