Context-based Friend Suggestion in Online Photo-Sharing Community *

Ting Yao [1], Chong-Wah Ngo [1], Tao Mei [2]
[1] City University of Hong Kong, Kowloon, Hong Kong
[2] Microsoft Research Asia, Beijing 100080, P. R. China
tingyao.ustc@gmail.com; cwngo@cs.cityu.edu.hk; tmei@microsoft.com

ABSTRACT

With the popularity of social media, web users tend to spend more time than before for sharing their experience and interest in online photo-sharing sites. The wide variety of sharing behaviors generate different metadata which pose new opportunities for the discovery of communities. We propose a new approach, named context-based friend suggestion, to leverage the diverse form of contextual cues for more effective friend suggestion in the social media community. Different from existing approaches, we consider both visual and geographical cues, and develop two user-based similarity measurements, i.e., *visual similarity* and *geo similarity* for characterizing user relationship. The problem of friend suggestion is casted as a contextual graph modeling problem, where users are nodes and the edges between them are weighted by *geo similarity*. Meanwhile, the graph is initialized in a way that users with higher *visual similarity* to a given query have better chance to be recommended. Experimental results on a dataset of 13,876 users and $\sim$1.5 million of their shared photos demonstrated that the proposed approach is consistent with human perception and outperforms other works.

Categories and Subject Descriptors

H.4.m [**Information Systems Applications**]: Miscellaneous

General Terms

Algorithms, Performance, Experimentation.

Keywords

Friend suggestion, social media, user similarity.

1. INTRODUCTION

With the rapid development of Web 2.0 technologies, online photo-sharing communities become increasingly popular. This offers users a distribution outlet for sharing their experience and interest with others [3]. Finding potential friends with similar interest in the social community will facilitate photo-sharing and browsing experience, as well as improve the effectiveness of user-target advertising.

In practice, most of the online communities allow users to self-create a set of interest groups, and then users can find friends and join their interest groups based on the group titles and tags. Since these interest groups are self-organized, plenty of groups usually fall into a similar topic [11], which leads to the difficulties for users to search for their interest groups, friends and photos. For example, 19,929 groups are returned by searching with the keyword *sport* in Flickr [6].

To address this problem, some online communities have provided the auto-suggestion function. For example, questionnaire based schemes ask users to do a set of psychological tests and then suggest friends according to the user-provided answers [15]. Usually, these tests require users to passively answer hundreds of questions. This is tedious and not succinct. A more common approach, for instance in [5], is by recommending common friends between users. This auto-suggestion approach is not fully automatic and requires human intervention.

Some works have also been done to explore the online rich media to automatically discover user relationship. Bhattacharyya *et al.* quantify user similarity based on user profiles such as location, hometown, activities, interest, favorites, and professional association [2]. Roth *et al.* suggest friends based on an implicit social graph, which is formed by user interaction with contacts and groups of contacts [12]. Li *et al.* propose to use location history to mine the user similarity and then recommend the potential friends, in which the location is only limited within tens of cities [9]. In brief, these applications focus on only one single cue. The rich media in the social communities are not fully exploited.

In real life, the chance that two persons will travel together to the same places and take similar photos would be higher when they share similar interest. Motivated by this observation, we propose a method to suggest friends for users, named *context-based friend suggestion*, by mining the user-shared photos and the geo-locations of their photos in the online photo-sharing community. Mathematically, we model the community as a graph in which the nodes represent users and the edges evaluate their relation strength. Specifically, given a *query user* (i.e., user who will be suggested friends), the graph is biased with the preference towards the users who take similar photos (higher similarity in visual appearance of the shared photos) as the query user, and further-

*Area chair: Mor Naaman

MM'11, November 28–December 1, 2011, Scottsdale, Arizona, USA.

more, the relation is weighted by the geo-location similarity. Therefore, we develop two user-based similarities, i.e., *visual similarity* and *geo similarity*, between a pair of users. The *visual similarity* indicates the similarity of users' shared photos in visual appearance, and it is filtered by photo tags and refined based on the representative photos which are selected based on photo comments and views. The *geo similarity* expresses the similarity of user experience and is defined based on the geographic distance of the shared photos.

The main contributions of this paper are two-fold. First, we propose a friend suggestion method in online photo-sharing community by leveraging multiple contexts, including user-contributed photos, their associated tags and geo-locations, as well as user behaviors like viewing and commenting. To the best of our knowledge, the proposed friend suggestion method represents one of the first attempts towards leveraging a variety of contexts for friend suggestion. Second, we develop two user-based similarity measurements, i.e., *visual similarity* and *geo similarity*, based on the multiple contexts of users' photos in online photo-sharing site.

The remainder of this paper is organized as follows. Section 2 provides a brief review of related work. Section 3 presents the proposed friend suggestion. Section 4 shows experiments and Section 5 concludes the paper.

2. RELATED WORK

Mining user relationship in the social network attracts more and more attentions in recent years [1][4][8][10]. Agrawal *et al.* analyze the social behavior of people on the newsgroups [1]. Li *et al.* develop a scalable community discovery solution for large-scale text document corpus [8]. To leverage more data types and their hidden relations, Cai *et al.* propose to learn an optimal linear combination of different relations, which can best meet the user expectation, to discover the hidden community [4]. Lin *et al.* propose a community discovery method through analysis of time-varying and multi-relational data in the rich media social network [10]. The ability to mine user relations also drive various applications. Bhattacharyya *et al.* quantify the user similarity based on user profiles [2]. Roth *et al.* suggest friends based on the implicit social graph formed by users' interactions with contacts and groups of contact [12]. Wu *et al.* present a friend recommendation system which focuses on people's visual appearances on portraits photos [15]. Li *et al.* propose to use location history to mine the user similarity and then recommend the potential friends [9].

3. CONTEXT-BASED FRIEND SUGGESTION

The basic idea of the proposed *context-based friend suggestion* is to recommend a list of friends that have similar interest with a given query user based on multiple contexts in online photo-sharing site. The framework is shown in Fig.1. In an online photo-sharing community, each user shares several photos with geo-locations, provides a set of textual tags to describe photo contents, and attracts some views and comments from the other users. For a given query user, the photo tags are first used to roughly select the potential users and filter out most of users, and then the selected users are ranked to obtain an initial friend list based on the user similarity in terms of visual appearance of their shared representative photos (*visual similarity*). The representative photos are obtained by considering the photo comments and views. Meanwhile, based on a large number of photos with geo-locations, we build a *geographic vocabulary* and represent each user by a set of *geographic words* obtained by vector quantizing the geo-locations into geographic vocabulary. Based on the *geographic words*, the *geo similarity* between each two users is obtained. Assuming that two users are likely to travel to the same places and take similar photos when they share similar interest, we build a graph to model the user friendship, in which, users are nodes biased with *visual similarity* to the query user, and edges between the nodes are weighted by their *geo similarity*. Through a propagation process, we can explicitly detect the user friendship strength and suggest a friend list to the query user.

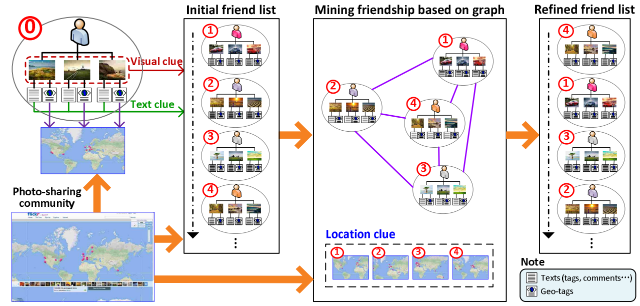


Figure 1: The flowchart of context friend suggestion.

3.1 Visual Similarity

The *visual similarity* is calculated based on the visual cues of users' representative photos. The representative photos are obtained according to the comments and view numbers of photos, assuming that the more the comments and views, the higher the quality of a photo. Specifically, if the number of views and comments in a photo exceeds the average view/comment number of the photo in an album, this photo will be selected as a representative one.

Denote R_u as the feature vector of representative photos of user u, then we can define the *visual similarity* between user u_i and u_j as follows:

$$s_v(u_i, u_j) = \min_{\mathbf{v}_i \in R_{u_i}, \mathbf{v}_j \in R_{u_j}} \cos(\mathbf{v}_i, \mathbf{v}_j), \quad (1)$$

where $\cos(\mathbf{v}_i, \mathbf{v}_j)$ is the cosine similarity by measuring the angle between the feature vectors of photos $\mathbf{v}_i$ and $\mathbf{v}_j$. We adopt the scale-invariant feature transform (SIFT) descriptor with a Difference of Gaussian (DoG) interest point detector [13]. The interest point is referred to as a local salient patch, each associated with a 128-dimensional feature vector. We further use K-means to cluster the similar patches into "visual words," and use the Bag-of-Word (BoW) model to represent each photo, as it has proven to be effective for object recognition [13].

3.2 Geo Similarity

The *geo similarity* is obtained based on the geographic distance of users' shared photos. First, we randomly collect a large number of photos with geo-locations (GPS coordinates). Then, we build a *geographic vocabulary* by using K-Means clustering and quantize each photo's geo-location into clusters like a "word" in a text document. In this way, all photos of a user will be represented as a set of *geographic words*.

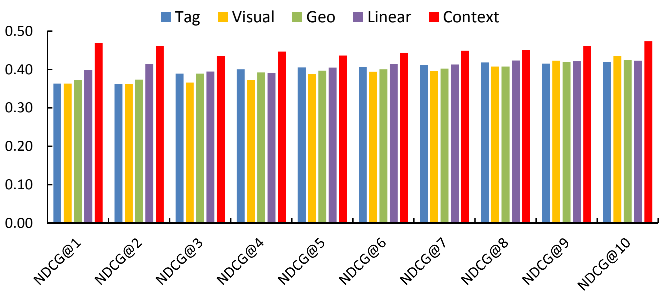


Figure 2: Subjective evaluation on friend suggestion.

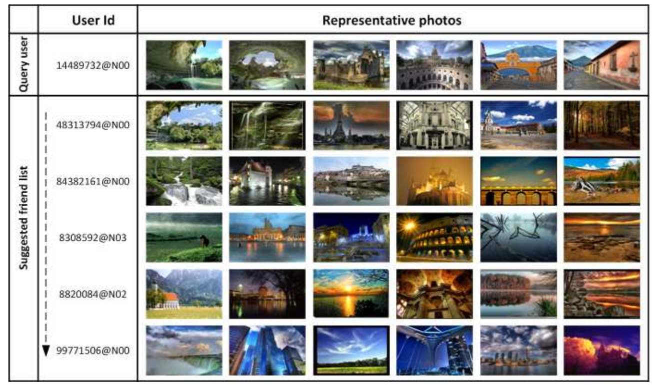


Figure 3: The representative photos of top five suggested friends for the query user with *Flickr* ID "14489732@N00."

Let K denote the size of the *geographic vocabulary*. Then each user can be represented by a vector of K dimensions, $\mathbf{g} = [g_1, g_2, \ldots, g_K]$, where each element g_i indicates the number of i^{th} words in a user's photo album. In other words, there are g_i photos fall into the i^{th} geographic words. Let $\mathbf{g}_i$ and $\mathbf{g}_j$ denote the geo representations of the user u_i and u_j respectively, then the *geo similarity* is calculated by the cosine similarity:

$$s_g(u_i, u_j) = \cos(\mathbf{g}_i, \mathbf{g}_j). \tag{2}$$

3.3 Context-based User Similarity

For a given query user q, let $\mathcal{U} = \{u_1, u_2, \ldots, u_N\}$ denote a set of N users. The objective is to rank these users according to the friendship with the query user q. The higher a user is ranked, the more similar interest they have. Let $G(\mathcal{V}, \mathcal{E})$ denote a connected graph, where nodes $\mathcal{V}$ correspond to the N users, initialized with the *visual similarity* between the users and query user, and edges $\mathcal{E}$ indicate the *geo similarity* between these nodes. We form the affinity matrix $\mathbf{S}_g = [s_g(u_i, u_j)]$ defined by an $N \times N$ symmetric similarity matrix on the edges of the graph, in which the element $s_g(u_i, u_j)$ is the *geo similarity* between user u_i and u_j, and $s_g(u_i, u_i) \equiv 0$. We construct the matrix $\mathbf{F} = \mathbf{D}^{-1/2}\mathbf{S}_g\mathbf{D}^{-1/2}$ in which $\mathbf{D}$ is a diagonal matrix with its (i,i)-element equal to the sum of the i^{th} row in $\mathbf{S}_g$. Following the method introduced in [16], we use the regularization framework and the cost function as follows

$$\begin{aligned} Q(\mathbf{s}) = \tfrac{1}{2}\big(\textstyle\sum_{i,j=1}^{N} s_g(u_i, u_j)\left\| \frac{s(u_i,q)}{\sqrt{D_{ii}}} - \frac{s(u_j,q)}{\sqrt{D_{jj}}} \right\|^2 + \\ \mu \textstyle\sum_{i=1}^{N} \|s(u_i, q) - s_v(u_i, q)\|^2\big), \end{aligned} \tag{3}$$

where $s(u_i, q)$ is the *context-based user similarity* of the user u_i and the query user q. The μ ($\mu > 0$) is the regularization parameter. The first term of Eq.(3) is the smoothness constraint, which defines the global consistency of the *geo similarity* over the graph. The second term is the fitting constraint that discourages large deviation from the initial scores of *visual similarity*.

Then the final *context-based user similarity* is given by

$$\mathbf{s}^* = \text{argmin}_{\mathbf{s}} Q(\mathbf{s}). \tag{4}$$

By differentiating $Q(\mathbf{s})$, we can obtain

$$\frac{\partial Q}{\partial \mathbf{s}}|_{\mathbf{s}=\mathbf{s}^*} = \mathbf{s}^* - \mathbf{F}\mathbf{s}^* + \mu(\mathbf{s}^* - \mathbf{s}^0) = 0, \tag{5}$$

where $\mathbf{s}^0$ represents the vector of *visual similarity*. We define $\alpha = \frac{1}{1+\mu}$ and matrix $\mathbf{I}$ as an identity matrix with diagonal element being 1 and the others 0. We can derive that

$$\mathbf{s}^* = (1-\alpha)(\mathbf{I} - \alpha\mathbf{F})^{-1}\mathbf{s}^0. \tag{6}$$

After obtaining the final *context-based user similarity* to the query user, the users are ranked accordingly to obtain the refined friend list.

4. EXPERIMENTS

4.1 Dataset

We conducted experiments on a dataset collected from *Flickr*, which includes 13,876 users and ∼1.5 million user-shared photos. Each photo has textual tags, geo-location, comments and view number from others. In this dataset, each user has at least five photos and each photo has geo-location information. In our experiments, the numbers of visual words and clusters based on GPS information are both set to 2,000 empirically [13]. The tradeoff parameter α is set to 0.9. The sensitivity of α will be analyzed in details in section 4.2.3.

4.2 Evaluations

We compared our context-based friend suggestion (Context) with the following four methods in both subjective and objective experiments:

- Tag: A *tag similarity* measured by the number of common tags on the users' photos. The higher the number of common tags, the more similar the two users.
- Visual: A *visual similarity* determined by estimating the visual similarity of user representative photos as described in Section 3.1.
- Geo: A *geo similarity* measured by exploring the similarity based on the geographic distance of users' photos as mentioned in Section 3.2.
- Linear: A linear fusion of Tag, Visual and Geo.

4.2.1 Subjective User Study

We randomly selected 1,000 users as the query users. The similarity of each user pair was manually labeled by three subjects on a scale of 1-3: (1) dissimilar (totally different interest), (2) similar (somewhat relevant interest), and (3) strongly similar (almost same interest). We adopted the Normalized Discounted Cumulative Gain ($NDCG$) as the performance metric [7].

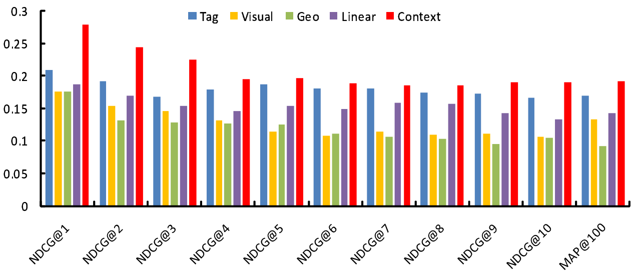
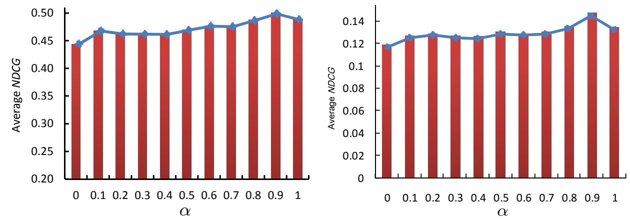


Figure 4: Objective evaluation on friend suggestion.



(a) Subjective user study (b) Objective evaluation

Figure 5: The performance curve with respect to the parameter α.

Fig.2 shows the experimental results. We can see that Context outperforms the others and achieves the best performance at different depth of *NDCG*. From the results, user similarity obtained by combining multiple contexts can express the user relationship more objectively and accurately, compared with the results using single cue. In contrast to the linear fusion, the superiority of Context indicates the effectiveness of jointly exploring user-based *visual similarity* and *geo similarity* over a graph model.

Fig.3 shows the top five users' representative photos to a query user according to the final user similarity. From the representative photos of each user, we can see that these photos are similar with each other. These users are likely to have similar interest and thus are potential candidates for friend suggestion.

4.2.2 Objective Evaluation of Friend Suggestion

The objective evaluation was conducted according to the *contacts* function in Flickr site, which is organized by users themselves. If a user exists in the other user's contact list, we view the two users as friends. We first filtered out the users who have no contacts in Flickr site or in our user set. After filtering, we obtained 2,975 users and took them as the query users. Since there are only two scales (in or not in contact list), we can also adopt the non-interpolated Mean Average Precision (*MAP*) [14] as another measurement. Fig.4 shows that the proposed Context also achieves the best performance at different depth of *NDCG* as well as *MAP*.

4.2.3 Parameter Sensitivity

In this section, we analyzed the sensitivity of the parameter α in Eq.(6). Fig.5 shows the *NDCG* performance with respect to different values of α. The *NDCG* score is obtained by averaging the top 20 results of all the query users. From the figure, we can see that the performance curve is smooth when α varies in a range from 0 to 1. The performance is not sensitive to the change of the parameter, which indicates the robustness of our method. The best performance is achieved when $\alpha = 0.9$.

5. CONCLUSIONS

We have presented our friend suggestion approach in online photo sharing community by leveraging multiple contexts, including user shared photos, associated geo-location, text cues (photo tags) and social behaviors (comments and views). Based on multiple contexts, we develop two types of user-based similarity, called *visual similarity* and *geo similarity*. Accordingly, we use a graph model to combine the two similarities. Future work includes the dynamic weighting of context cues according to their importance. We are also interested in exploring other applications of friend suggestion, such as interest group suggestion and auto-generation.

6. ACKNOWLEDGEMENTS

The work described in this paper was partly supported by a grant from the Research Grants Council of the Hong Kong Special Administrative Region, China (CityU 119709).

7. REFERENCES

[1] R. Agrawal, S. Rajagopalan, R. Srikant, and Y. Xu. Mining newsgroups using networks arising from social behavior. In *Proceedings of WWW*, Budapest, May 2003.

[2] P. Bhattacharyya, A. Garg, and S. F. Wu. Analysis of user keyword similarity in online social networks. *Social Network Analysis and Mining*, Oct 2010.

[3] D. Boyd and N. Ellison. Social network sites: Definition, history, and scholarship. *Journal of Computer-Mediated Communication*, 13(1):210–230, 2008.

[4] D. Cai, Z. Shao, X. He, X. Yan, and J. Han. Mining hidden community in heterogeneous social networks. In *Workshop on Link Discovery: Issues, Approaches and Applications*, Chicago, 2005.

[5] Facebook. http://www.facebook.com/.

[6] Flickr. http://www.flickr.com/.

[7] K. Jarvelin and J. Kekalainen. Cumulated gain-based evaluation of ir techniques. *ACM Transactions on Information Systems*, 20(4):422–446, 2002.

[8] H. Li, Z. Nie, W.-C. Lee, C. Giles, and J.-R. Wen. Scalable community discovery on textual data with relations. In *Proceedings of CIKM*, Napa Valley, Oct 2008.

[9] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma. Mining user similarity based on location history. In *Proceedings of GIS*, Irvine, 2008.

[10] Y.-R. Lin, J. Sun, P. Castro, R. Konuru, H. Sundaram, and A. Kelliher. Metafac: Community discovery via relational hypergraph factorization. In *Proceedings of KDD*, Paris, June 2009.

[11] R.-A. Negoescu, B. Adams, D. Phung, S. Venkatesh, and D. Gatica-Perez. Flickr hypergroups. In *Proceedings of ACM Multimedia*, Beijing, Oct 2009.

[12] M. Roth, A. David, D. Deutscher, and et al. Suggesting friends using the implicit social graph. In *Proceedings of KDD*, Washington DC, 2010.

[13] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In *Proceedings of ICCV*, 2003.

[14] TRECVID. http://www-nlpir.nist.gov/projects/trecvid/.

[15] Z. Wu, S. Jiang, and Q. Huang. Friend recommendation according to appearances on photos. In *Proceedings of ACM Multimedia*, Beijing, 2009.

[16] D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Schölkopf. Learning with local and global consistency. *Advances in Neural Information Processing Systems*, 2003.