Bag-of-Features Tagging Approach for a Better Recommendation with Social Big Data

Ming Cheung

HKUST-NIE Social Media Lab,
Department of Electronic and Computer Engineering
Hong Kong University of Science and Technology,
Hong Kong
Email: cpming@ust.hk

James She

HKUST-NIE Social Media Lab,
Department of Electronic and Computer Engineering
Hong Kong University of Science and Technology,
Hong Kong
Email: eejames@ust.hk

Abstract—**The interests of users are always important for personalized content recommendations on friendships, events and media content from the social big data. However, those interests may not be specified, which makes the recommendations challenging. One of the possible solutions is to analyze the user's interests from the shared content, especially images with manually annotated tags. They are shared on online social networks such as Flickr and Instagram. However, the accuracy of the recommendation is greatly affected by the accuracy of the tag, which is not always reliable. This paper demonstrates how a bag-of-features (BoF)-based tagging approach can help to improve the accuracy of recommendations using an unsupervised algorithm. A set of auxiliary tags is used to represent user interests and, hence, the recommendation. The approach is evaluated with over 500 user and 200k images from Flickr. It is proven that by BoF tagging (BoFT), friendship recommendation is possible without friendship/tag information and the recall and the precision rate are improved by about 50% over using user tags.**

Keywords–Image tagging; recommendation; online social network; bag-of-features; annotation; big data.

I. INTRODUCTION

Nowadays, sharing social content has become part of our lives, in which billions pieces of content are shared. Recommending content that matches the user's interests from the social big data is important for any social networks. However, the user's interests may be hidden, that is the interests are not specified in the user profile. With an incomplete set of data, the content recommendation may be inaccurate. On the other hand, the user's interests are reflected in the abundance of social content, especially image, shared on the networks. A good recommendation is possible through analyzing the users' interests reflected among shared images. One of the most important applications is friendship recommendation, the inference of the connection between two users [1]. One of the possible ways to analyze a user's interests from shared images is through tagging [2]. Tagging, the act of using text to annotate a social content, is one of the most basic and essential features in any social networks that helps content recommendation [3]. The tags describe the image and reflect the users' interests since users with similar interests are more likely to upload images with similar tags. Connections can be discovered with the tendency to make friends with someone who shares similar interests reflected in the tags. Fig. 1(a) to (d) is a set of images and their tags by different users in Fotolog, Flickr, Twitter and Instagram. The tags include the name of the object, location,

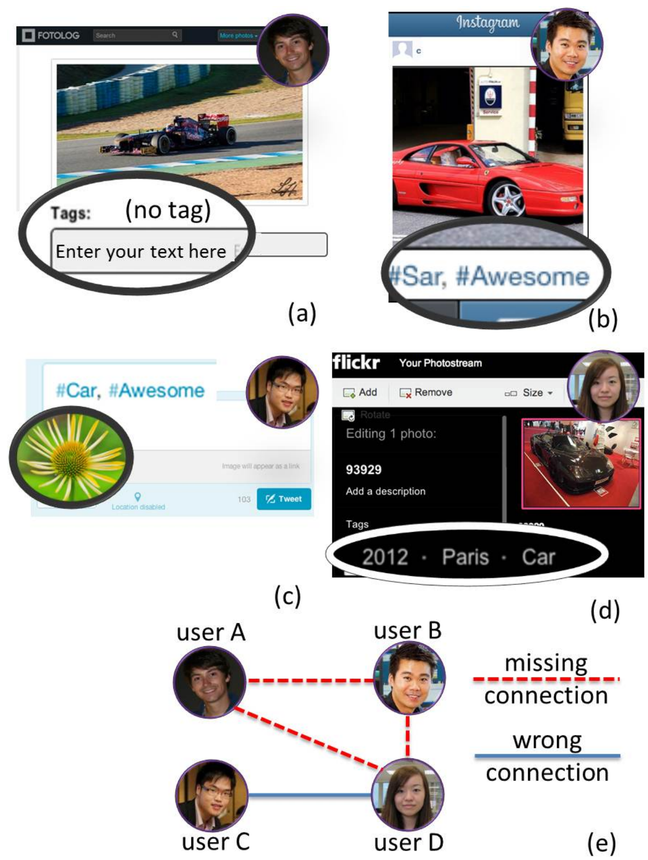

Figure 1: Examples of tagging on: (a) Fotolog, (b) Instagram, (c) Twitter, (d) Flickr, (e) and the corresponding social graph.

time or even the feeling felt at the time by the user. They are the major types of tags [4].

A tie between 2 people indicates that there is a connection between them such as friendship. The strength of connections among users can be measured by tie strength. People with higher tie strength such as best friends have a higher influence on the user. One of the important elements in measuring the tie strength is the common interests they share [5]. As tags reflect the interests of users, the tie strength can be estimated by the similarity of tags and hence, calculate possible connections. For example, if two users both upload images with the tag, "Car", the users are similar in terms of their interest and can be connected. Reliable tags that can reflect the nature of an image

 ISBN: 978-1-61208-364-3

are essential for this calculation. However, in most of the social networks, the tags are added manually and are not always reliable [6]. They may not be accurate, suitable for analysis, or, sometimes even available. Some users are not interested in annotating the images they upload because of the considerably longer time required than simply uploading the images, as the user shown in Fig. 1(a). For those tagged images, the tags may not be a good description of the content. Users may type the tag wrongly such as the tag, ”Sar” instead of ”Car”, in Fig. 1(b). Some irrelevant tags are added intentionally to increase the popularity. An example is the tag ”Car” for the picture of flowers of Fig. 1(c). For a good annotated image, analyzing the tags is still not an easy task. The tags may have different levels of details [8] or diverse details. For example, in Fig. 1(d), the user annotates the object, time and place. As a result, the social graph by the tags, as shown in Fig. 1(e) has wrongly connected user C with user D while leaving users A and B without any connection. These are some common examples of how a user annotates an image on social networks and how that annotation affects the discovery of connections.

This paper proposes a novel approach using BoF-based tagging that makes better recommendations through users' interests discovered from images uploaded. Instead of using a supervised approach in [9], this paper proposes an unsupervised approach in which images are grouped visually by BoF. The approach is also evaluated using a dataset of 542 users and 201006 images and the actual relationship among users. The results prove that the proposed approach can help to make a better recommendation. The main contributions are the following: 1) propose a novel way to represent user interest with auxiliary tags in an unsupervised manner; 2) introduce a friendship recommendation approach based on the auxiliary tags; and 3) verify the recommendation with the actual relationship from the scraped data. Section II in this paper discusses previous works. Section III is the general context of a BoF-based tagging, followed by how to connect people with similarity in Section IV. Section V shows the details of the experimental result and Section VI concludes the paper.

II. Previous Works

Recommending personalized content from billions of shared content is always a challenging task. Information overload may occur so that users have difficulty processing the huge amount of available content. A possible solution is a recommendation system based on a trust-based approach [10], where the user's social connection is considered for filtering the content. The interest shared is a way to measure tie, the strength of the social connection [7]. A hybrid approach can combine interests and the social connections. However, obtaining the user's interest is not always available. Although users can enter their interests for better recommendations, they may not want to spend time on the annotation. One of the possible solutions is to analyze the interests reflected in the content they have shared, especially images. This analysis can be based on the user annotated tags that describe the images. However, those tags may be inappropriate, wrong or have a different degree of details. One of the most well-known solutions is Collaborative Filtering (CF) [11][12], in which the same social content will be tagged by many users. The final tag quality can be improved [13] by analyzing the tags from different users. Although there is a promising result by applying CF, it is not suitable for systems with a large amount of images. The reason behind is that only small portion of images are popular and receive many tags. While it gives some of the images appropriate tags, most of them are left without proper tags for analysis. In image sharing platforms, such as Flickr, a user can upload hundreds of images at a time which makes CF inappropriate.

Another possible way is a content-based approach, in which the visual features are considered in order to annotate an image [14][15]. However, determining the relationship between the features and the tags is not a trivial task. The same object can be visually different among images. In this paper, BoF-based tagging [16] is applied. The proposed approach makes use of computer vision techniques in object recognition tasks to infer interests for friendship recommendations. BoF is an image-based approach that detects low-level features, and encodes an image into a feature vector. An unsupervised method is used for the learning. Images are grouped based on the similarity of their features vectors and hence the similarity of 2 users can be calculated. With this approach, it is possible to obtain the tag given to an image and, therefore, recommendation.

Among different types of content recommendation, friendships, or connections among people, is one the most important and fundamental functions. This problem has long been studied. One of the possible ways to make the recommendation is by the existing connections among people [17][18]. However, this may limit the recommendations from millions of users as the connections among users may not available. Friendship recommendation is also possible with user interests [18] inferred from user input [19] or user generated content [20] and other personal information [1]. Interests are combined with the existing connections with a machine learning algorithm in [21] for recommendation. In [22], the authors focus on how to make use of the group information on Flickr for friendship recommendation. The co-occurrence in images can also be a cue on friendship recommendation [23].

III. BoF-based Tagging

BoF has been a popular approach to many computer vision tasks because of its simplicity [16]. BoF is a method to represent images into feature vectors of local image descriptors. Fig. 2 is the process of the proposed approach in which Fig. 2(a) is the use of BoF in this work. The different parts of the BoF tagging are introduced in this section.

A. Feature Extraction

Feature extraction is a process to obtain the local features in step 1 of Fig. 2. These features can be detected by Harris Affine detector, or Maximally Stable Extremal Regions detector [16]. The extracted features are relatively consistent with viewing angles and lighting conditions. They are represented in a way that is independent of the size and orientation, such as scale-invariant feature transform (SIFT) [24].

B. Codebook Generation

Codebook generation (step 2 of Fig. 2) is a process to obtain the visual words that can represent the features obtained in the feature extraction in step 1 of Fig. 2. It is a clustering process that groups similar features. The mean vectors of each group are defined as the visual word, which can be used to represent

 ISBN: 978-1-61208-364-3

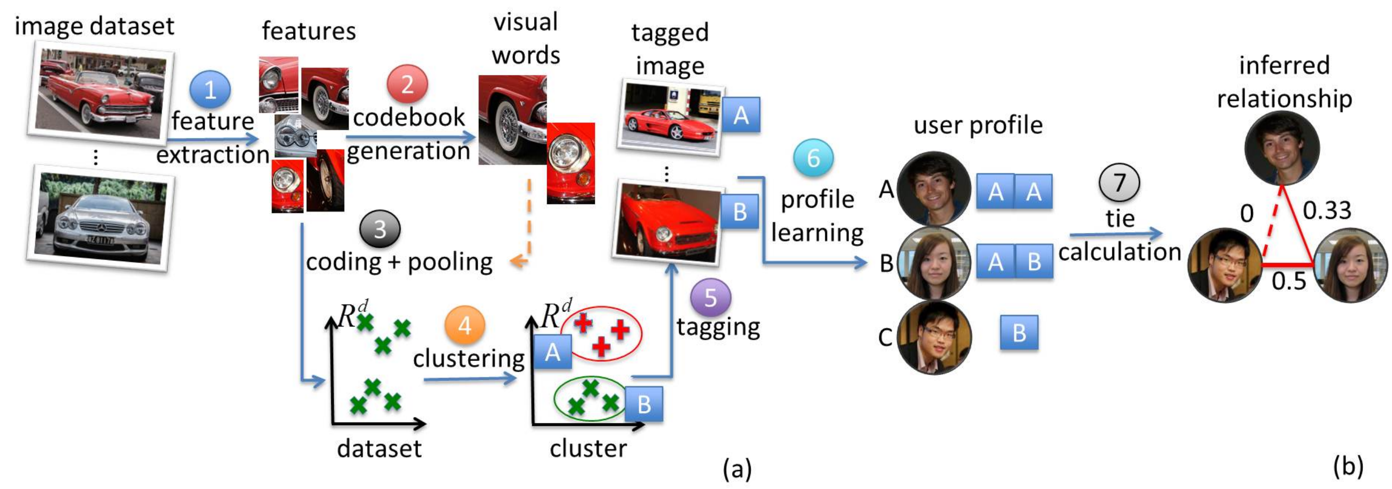

Figure 2: Flow chart of the proposed approach: (a) BoF-based tagging, (b) relationship recommendation.

the features in these images. One of the possible techniques to obtain those clusters is by k-mean clustering [25], which groups the visual features based on their visual similarity. The codebook generation is an offline process that does not need to be updated in real time.

C. Feature Coding and Pooling

Feature coding is to encode features with the visual words. Each feature in every image is represented by a visual word in feature coding. The images are then represented by a feature vector in the feature pooling. This process is carried out in encoding the images in the dataset (step 3 of Fig. 2). One of the most common approaches is using the histogram that counts the number of occurrences of each visual word in the image. The feature vector obtained is used in the clustering to group images that are visually similar.

D. Clustering and Tagging

The goal of clustering is to group images with similar feature vectors, that is, group images that are visually similar. Each cluster obtained in this operation corresponds to similar objects to which an auxiliary tag is assigned. After obtaining the cluster in step 4 of Fig. 2, the images in any cluster are assigned with the same auxiliary tag to reflect that they are visually similar and belong to the same group. It is an unsupervised operation no assumption is made or information on the image is known.

IV. Proposed BoF-Based Recommendation

This section introduces how to find similarities among people from the result of the BoF tagging (BoFT). The first part introduces how to learn the user profile from the result of BoF tagging, while the second part discusses how to make recommendations based on the user profile.

A. BoF Tagging and User Profile

The user profile, which reflects the interests of the users is the key in the content recommendation. A user profile can be obtained based on user manual input, in which the user manually inputs what kind of content is their favor. In the proposed approach, it is assumed that no user input is needed and the user profile is obtained from the image uploaded as in step 5 of Fig. 2. The histogram of the tags is used as the user profile in the proposed approach.

B. User Profile and Recommendation

When the user profile is obtained, the next step is to make a recommendation to the user. Recommendations are based on the tie, the strength of the relationship between two people. An item favored by a user may also be liked by friends of the user with strong ties. For example, user A likes Ferrari, while user B likes BMW and user A is user C's best friend. As a result, it is more likely that user C likes Ferrari. Content recommendation is then possible with the value of the tie calculated by user profile with the following formula:

$$S_{cosine}(A, B) = \frac{T_A \cdot T_B}{||T_A|| \cdot ||T_B||} \tag{1}$$

where T_A is the set of tags in the images uploaded by user A and T_B is the set of tags in the images uploaded by user B. The pairwise similarity is calculated based on the user profile. It is possible to obtain the tie between two people and find that people with similar interests have a higher similarity. The tie is assumed to be undirected, which means that the tie is the same from user A to B and user B to A. Different types of recommendation is possible with user ties, in particular, the focus of this paper is on friendship recommendation.

V. Experimental Results

In this section, the dataset, the experiments and the results are discussed. In this paper, the discussion focuses on discovering the connections among users by the tendency of people to make friends with people who share similar interests. The results show that it is possible to infer connections by using the BoF tagging.

A. Dataset and Experimental Setup

The setting of the experiment is shown in Fig. 3. A set of 201006 images uploaded by 542 users is scraped from Flickr,

 ISBN: 978-1-61208-364-3

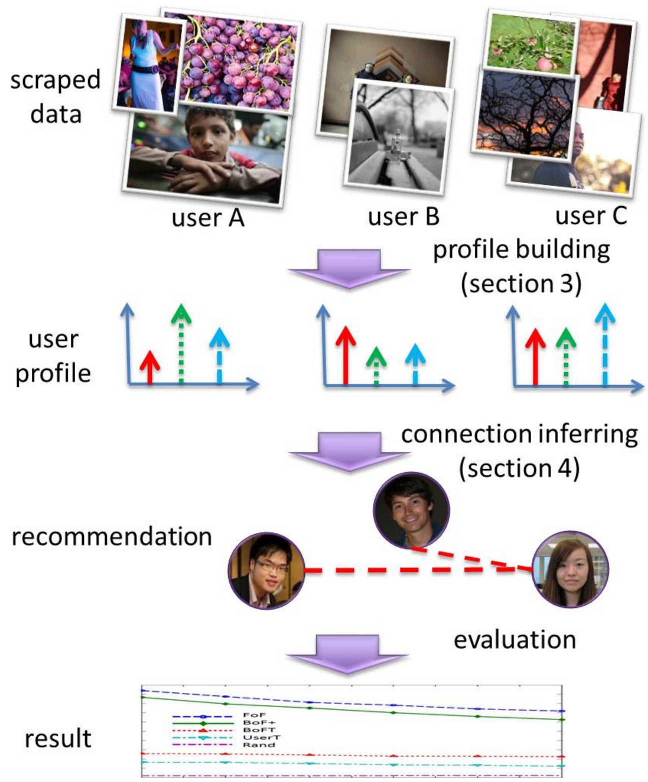

Figure 3: Experiment setting

an online social network for image sharing with millions of images uploaded and tagged, and the BoF tagging approach is used to tag those images. The 542 users are selected randomly from images under the same tag query page to provide diversity. The user profiles from the uploaded images are built with the tags obtained. Then connections among users are inferred with the tie calculation and evaluated with the actual connections scraped with the images. Tables I and II show the attributes scraped for the users and the images.

TABLE I: MAJOR ATTRIBUTES FOR IMAGES

Attribute	Description
ImageID	the unique ID for the image
Tag	the set of user annotated tags

TABLE II: MAJOR ATTRIBUTES FOR USERS

Attribute	Description
UserID	the unique ID for the user
ImageUploaded	the set of images uploaded by the user
FriendList	the user ID of the user friends

In the dataset, there are a total of 2827409 tags among the images, on average there are 14 tags per image and 5422 tags per user. There are 152938 unique tags. The average number of friends of a user is 170 for which there are 902 connections among the 542 users. The goal of the experiment is to infer those connections using the set of images uploaded by the users, even without using the friendship information.

The features of all the images are detected by the Harris-Affine key point detector. They are described by the 128-dimension SIFT descriptor. The visual words are obtained by k-mean and then used to represent all the images as feature vectors. A clustering operation is then used to group images with similar feature vectors. The images in the same cluster are assigned with the same auxiliary tag, in which each cluster has a unique auxiliary tag as in Fig. 4. The auxiliary tags are then used to build the user profile. the friendship recommendation is based on the similarity of the user profile of the users.

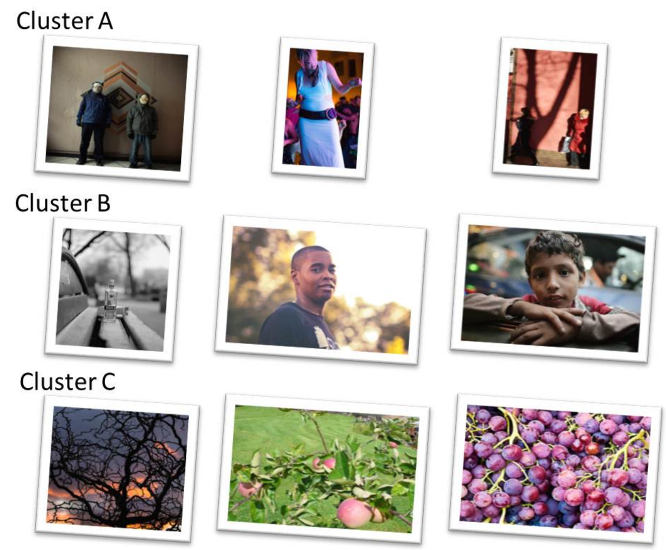

Figure 4: Examples of images in clusters

Different approaches are implemented for comparisons. The first approach is a random approach (Rand), in which users are recommended randomly. This is the baseline for the comparison to simulate the condition that user information such as friendship is not available. Two other approaches are also implemented to compare the proposed one. The first one is the Friend-of-Friend (FoF) [26], in which the similarity is based on the Jaccard similarity of the friend list. The similarity between two users with more common friends is higher. The FoF approach serves as the upper bound of the proposed approach to simulate the condition that information are available. The second approach is similar to the proposed approach, but instead of using auxiliary tag, the user annotated tags are used (UserT). The similarity is based on the tag they used for their images. It is also interesting to check if the additional information from BoFT can improve the preference of the upper bound, FoF. BoFT+ is defined as the following:

$$S_{BoFT+} = \beta * S_{BoFT} + (1 - \beta) * S_{FoF} \qquad (2)$$

where β is a constant, S_{BoFT}, S_{FoF} and S_{BoFT+} are the similarities of the BoFT, FoF and BoFT+. A study on the performance with different values of β is carried. It is measured by the area under curve (AUC) on the recall rate with 5 to 10 recommendation and shown in Fig. 5. A higher value in AUC implies a better performance. It is observed that AUC is maximum when β is 0.024. It implies that the majority of information is from FoF. In all approaches, no recommendation is made if the similarity between that user to others are all 0. For example, in the FoF approach, no recommendation is made if a user has no friend. It is a valid

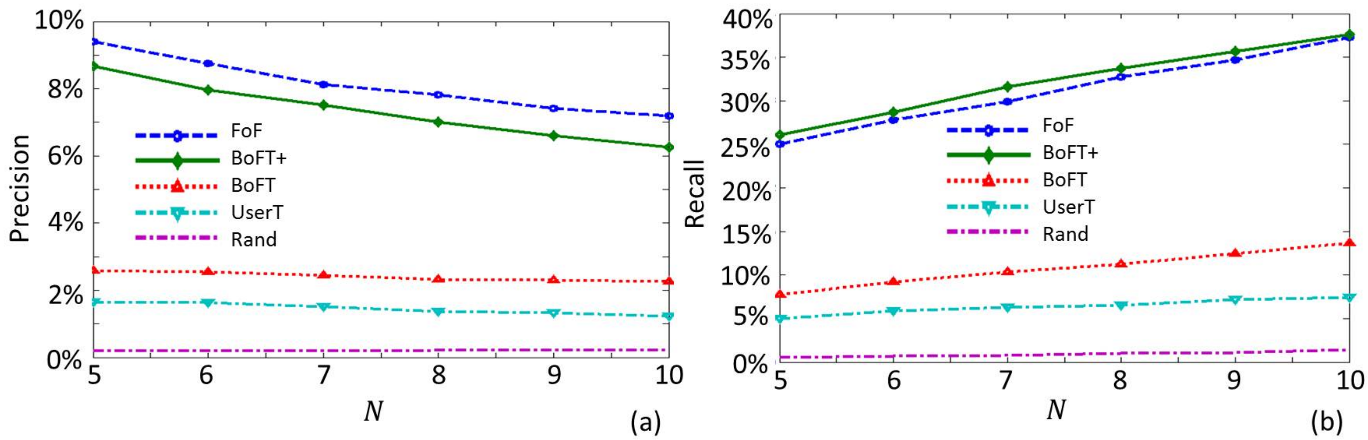

Figure 6: Result rate of different approaches: (a) precision, (b) recall.

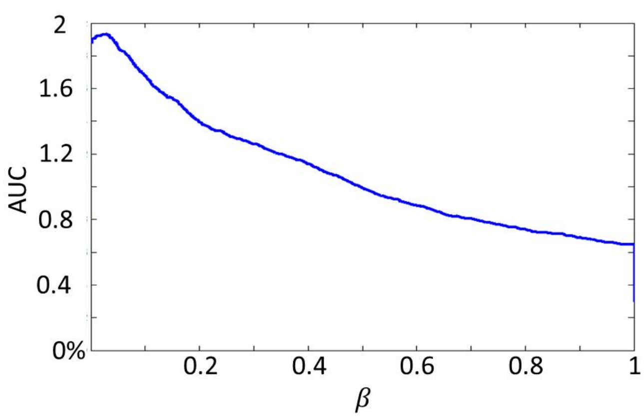

Figure 5: β vs. AUC for $N \subset [5, 10]$

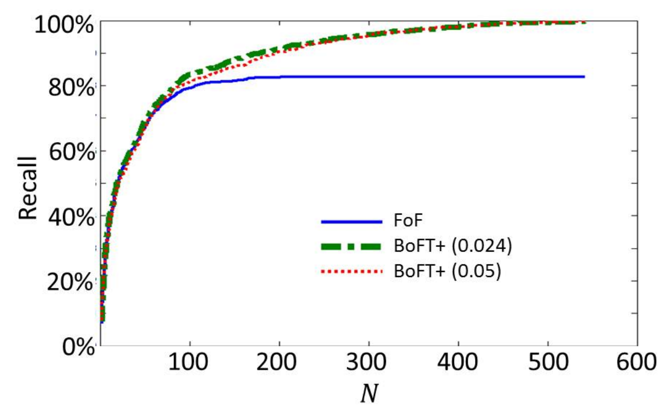

Figure 7: Recall rate for BoFT+ and FoF for N ⊂[1,542].

assumption when information is not available.

The results are evaluated by two popular rates, top N recall rate and top N precision rate as the following:

$$Precision = \frac{T_p}{(T_p + F_p)} \tag{3}$$

$$Recall = \frac{T_p}{(T_p + F_n)} \tag{4}$$

where T_p is the true positive (the recommended connection is an actual connection), while F_p and F_n are false positive and false negative respectively. F_p is the case that the recommended connection is not a connection, while F_n are the connections that are not recommended. The physical meaning for precision rate is the percentage of recommended items is actual connections. The recall is the percentage of actual connections that is recommended. The higher the values are, the better the approach is. When more items are recommended, recall rate is increased, but precision rate decreases. A list of recommendations is generated for each user with the proposed approaches. The approaches are evaluated with the top N per user recall rate and the top N precision rate, in which the top N users with the highest similarity, are recommended.

B. Results

Fig. 6 shows the top-N precision and recall rate of different algorithms for 5 to 10 recommendations per users. It is observed that FoF, BoFT, UserT and BoFT+ are all better than using user tags in terms of the two rates. It is clear that BoFT approach outperforms UserT and random approach and is much closest to the upper bound, FoF. BoFT+ can only improve the performance slightly. A detailed discussion can be found in the next subsection.

C. Discussion

In the experiment, it is observed that all approaches are better than the random one. The use of BoFT provides a significant improvement on the performance of the recommendation. By BoFT, the recall and the precision rate are improved by about 50% over UserT. Although the recall rate for BoFT+ is slightly higher than FoF, the precision rate of FoF is higher than BoFT+ as shown in Fig. 6. As discussed in previous section, the additional BoFT information can only slightly improve the FoF approach. It is also interesting to check the performance when N is large. The top N recall rate for BoFT+ with 2 values of β and FoF are shown in Fig. 7. When N is smaller than 100, the BoFT+ approach are only slightly better than FoF. However, when N is large, the improvement is more significant and the recall is increased by 13.3% with AUC by BoFT+. As mentioned in the previous section, the higher in

 ISBN: 978-1-61208-364-3

FoF approach in terms of the precision rate is that some users have no common friends with others. If two people have no common friend, no recommendation is possible. On one hand, users with no common friend are never recommended to each other in FoF approach and results in a higher precision rate. The number of common friends between 2 users can be small, or even zero. As a result, FoF approach may only capture relations with high confidence (with common friends) but those without any common friends. A high precision rate but a lower recall rate are obtained. On the other hand, BoFT can connect people through the images they have uploaded and therefore most users can be reached. The BoFT approach provides a new way to connect people. The next research challenge is how to improve the prediction performance such that it is close to the upper bound with the shared images.

VI. Conclusion

This paper proposes a novel BoF based Tagging approach to make better recommendations matching users interests discovered from images uploaded It demonstrated how a BoF-based approach can help discover hidden users' connections from the shared images on a social network for a better recommendation. As user friendship/interests are not always available, this paper proposes an unsupervised approach to classify image according to their visual features. Those visual features represent the user interest, and hence recommendation. The proposed approach is evaluated by friendship recommendation with a scraped dataset from Flickr with over 500 users and 200k images. It is proven that the proposed approach can recommend friendship to user based on the image shared by the users. With our proposed approach, BoFT, the recall and the precision rate are improved by about 50%.

Acknowledgment

This work is supported by HKUST-NIE Social Media Lab., HKUST

References

[1] J. Chen, W. Geyer, C. Dugan, M. Muller, and I. Guy, "Make new friends, but keep the old: Recommending people on social networking sites," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 2009, pp. 201-210.

[2] X. Li, L. Guo, and Y. E. Zhao, "Tag-based social interest discovery," in Proceedings of the 17th International Conference on World Wide Web, April 2008, pp. 675-684.

[3] T. C. Zhou, H. Ma, M. R. Lyu, and I. King, "UserRec: A user recommendation framework in social tagging systems." In Proceedings of AAAI, July 2010, pp. 1486-1491.

[4] T. Hammond, T. Hannay, B. Lund, and J. Scott, "Social bookmarking tools (I) a general review," D-Lib Magazine, vol. 2, 2005.

[5] S. A. Golder and S. Yardi, "Structural predictors of tie formation in twitter: Transitivity and mutuality," in Social Computing (SocialCom), 2010 IEEE Second International Conference on, August 2010, pp. 88-95.

[6] L. Kennedy, M. Naaman, S. Ahern, R. Nair, and T. Rattenbury, "How flickr helps us make sense of the world: Context and content in community-contributed media collections," in Proceedings of the 15th International Conference on Multimedia, September 2007, pp. 631-640.

[7] E. Gilbert and K. Karahalios, "Predicting tie strength with social media," in Proceedings of the SIGCHI Conference on Human Factors in Computing Systems, April 2009, pp. 211-220.

[8] D. G. Lowe, "Distinctive image features from scale-invariant keypoints," International Journal of Computer Vision, vol. 60, November 2004, pp. 91-110.

[9] M. Cheung and James She "Predicting Peoples Interests and Connections Using Their Shared Photos," unpublished.

[10] F. E. Walter, S. Battiston, and F. Schweitzer, "A model of a trust-based recommendation system on a social network," Auton. Agents Multi-Agent Syst., vol. 16, February 2008, pp. 57-74.

[11] X. Zhang, Z. Li, and W. Chao, "Tagging image by merging multiple features in a integrated manner," Journal of Intelligent Information Systems, vol. 39, August 2012, pp. 87-107.

[12] B. Sigurbjrnsson and R. Van Zwol, "Flickr tag recommendation based on collective knowledge," in Proceedings of the 17th International Conference on World Wide Web, April 2008, pp. 327-336.

[13] J. Sang, C. Xu, and J. Liu, "User-aware image tag refinement via ternary semantic analysis," IEEE Transactions on Multimedia, vol. 14, February 2012, pp. 883-895.

[14] E. Moxley, J. Kleban, J. Xu, and B. Manjunath, "Not all tags are created equal: Learning flickr tag semantics for global annotation," in Multimedia and Expo, 2009. ICME 2009. IEEE International Conference on, June 2009, pp. 1452-1455.

[15] X. Zhang et al., "Social image tagging using graph-based reinforcement on multi-type interrelated objects," Signal Processing, vol. 93 no. 8, August 2013, pp. 2178-2189

[16] Z. Wang, L. Sun, W. Zhu, S. Yang, H. Li, and D. Wu, "Joint Social and Content Recommendation for User-Generated Videos in Online Social Network," IEEE Transactions on Multimedia, April 2013, pp. 698-709

[17] I. Guy, N. Zwerdling, I. Ronen, D. Carmel, and E. Uziel, "Social media recommendation based on people and tags," in Proceedings of the 33rd International ACM SIGIR Conference on Research and Development in Information Retrieval, July 2010, pp. 194-201.

[18] I. Guy et al., "Personalized recommendation of social software items based on social relations," in Proceedings of the Third ACM Conference on Recommender Systems, October 2009, pp. 53-60.

[19] W. H. Hsu, A. L. King, M. S. Paradesi, T. Pydimarri, and T. Weninger, "Collaborative and structural recommendation of friends using weblog-based social network analysis," in AAAI Spring Symposium: Computational Approaches to Analyzing Weblogs, March 2006, pp. 55-60.

[20] X. Xie, "Potential friend recommendation in online social network," in Green Computing and Communications (GreenCom), 2010 IEEE/ACM Int'l Conference on & Int'l Conference on Cyber, Physical and Social Computing (CPSCom), December 2010, pp. 831-835.

[21] V. Leroy, B. B. Cambazoglu, and F. Bonchi, "Cold start link prediction," in Proceedings of the 16th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, July 2010, pp. 393-402.

[22] R. Parimi and D. Caragea, "Predicting friendship links in social networks using a topic modeling approach," in Advances in Knowledge Discovery and Data Mining Anonymous Springer, May 2011, pp. 75-86.

[23] H. Kim, J. Jung, and A. El Saddik, "Associative face co-occurrence networks for recommending friends in social networks," in Proceedings of Second ACM SIGMM Workshop on Social Media, October 2010, pp. 27-32.

[24] F. Tian, X. Shen, F. Shang, and K. Zhou, "Automatic image tagging by multiple feature tag relevance learning," in Pattern Recognition Anonymous Springer, 2012, pp. 505-513.

[25] D. A. Forsyth and J. Ponce, Computer Vision: A Modern Approach. Prentice Hall Professional Technical Reference, 2002.

[26] L. L and T. Zhou, "Link prediction in complex networks: A survey," in Physica A: Statistical Mechanics and its Applications, vol. 390, March 2011, pp. 1150-1170.

 ISBN: 978-1-61208-364-3