ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 16, March 2016

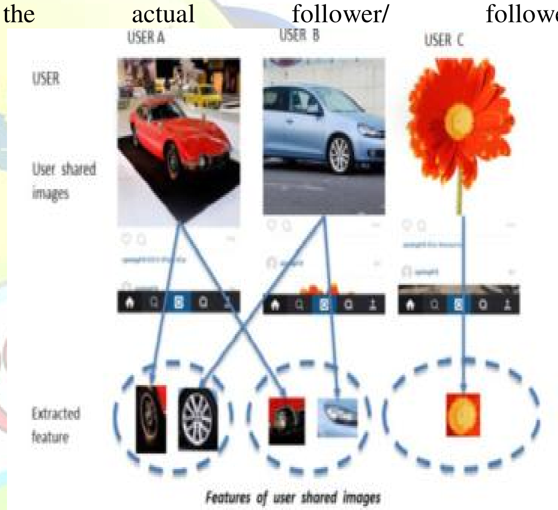
# Associating User Shared Images on Social Media

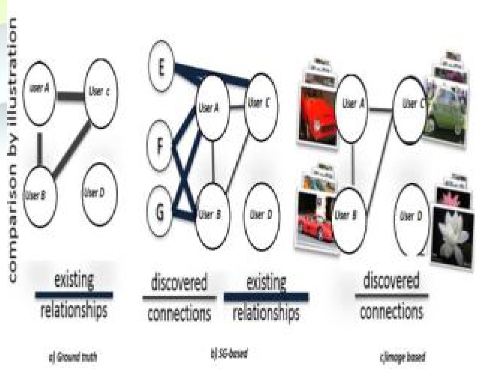
M. Suganya [1],U. Akil priyadharshini[2],D. Kanagalakshmi[3], R. Raja prabha[4]
Research Scholar-Madurai Kamaraj University[1],UG scholar[2,3,4],
Ultra-college of engineering and technology for women, Madurai,
Tamil Nadu,India
Email:akilpriyadharshini@gmail.com

***Abstract*---Social tagging systems have emerged as an effective way for users to annotate and share objects on the Web. However, with the growth of social tagging systems, users are easily overwhelmed by the large amount of data,through this 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, using user social graph we may find their interest easily.This social graph includes users detail.In our project we are going to post add directly on user account completely based on user interest.Because if the advertisement is based on their interested user will surely like or share it,through this a higher number of recommendations occur.**

## 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. User connections are the major information,such connection may be any type of online relationship based on user interaction in social media. Social tagging systems have emerged as a popular way for users to annotate, organize and share resources on the Web Social tagging systems enjoy the advantages that users can use free-form tags to annotate objects, which can ease sharing of objects despite vocabulary differences. As a form of users individual behavior, tagging activity not only can represent users judgments on the resources but also can indicate users personal interests. Some users also hide or limit the information of their connections from the public in social media platforms due to privacy concerns At a fundamental level, we gain insights into utilizing information in social tagging systems to provide personalized service for each user. Both users and shared images about cars and user shared an image about a flower . The follower/ followee relationship between users A and B can be possibly detected from the higher similarity of visual features in their shared images. When more shared images from each of users, andare accessible for evaluation, the actual follower/ followee




relationships should become reliably and accurately detectable though becoming challenging to process when the number of shared images and user connections grows bigger and faster every day in

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 16, March 2016

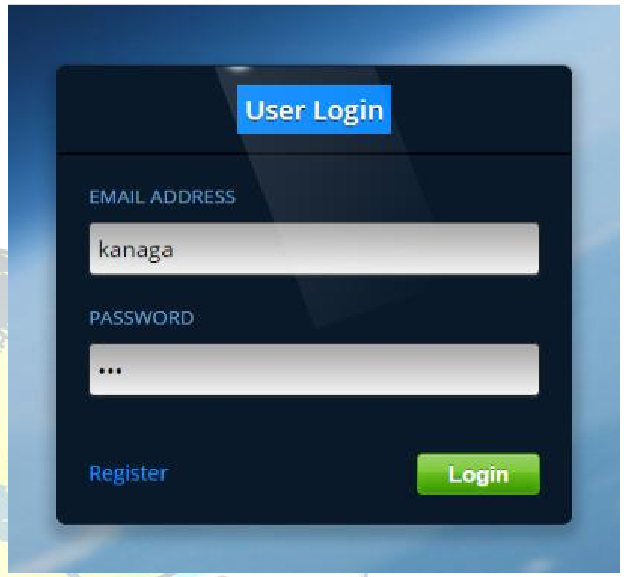
social network. Another possible way is a content-based approach, in which the visual features are considered in order to annotate an image. However, determining the relationship between the features and the tags is not a trivial task. The same object canbe visually different among images. In this paper, BoF-based tagging is applied. The proposed approach makes use of computer vision techniques in object recognition tasks to infer interests for friendship recommendations. BoF is an imagebasedapproach that detects low-level features, and encodes an image into a feature vector.

## II.RELATED WORKS

User behaviors in online social networks have been recently studied through the use of SGs and it was concluded that relationships (such as follower/followee and online friendships) in an SG are not formed randomly, but follow the power law distribution. An example where connection discovery is made via existing relationships,such as follower/followee, thatusers share in common. Users ,and share common related users, and connections amongthem can be obtained, while user is alone. Without access to SGs, follower/followee recommendation is also possible with the connection discovered by user common interests inferred from user input or user generated content and other personal information. Analyzing shared images can help to understand users, and hence discover the connections among them. Another common method to discover user connectionsis to analyze user annotated tags on shared images,in which a tag in the form of textual wording is provided by a user as meta-data to describe the shared image.
However, only some popular images are annotated by many users, while the restare either not correctly annotated or missing annotation, which leads to a poor connection discovery performance. Fig. 2(c) is an example of how connections are discovered by BoFT similarity distribution to explain how they are related to the connection discovery and follower/followee relationships; and formulated follower/followee recommendation based on the measurements and models, and verified that connection discovery is useful for recommendation and gender identification, even without using SGs and user annotated tags. user generated images. Users , and generate car images, and their connections can be discovered by the similarity among their images.

*USER PROFILE GENERATION*

This section introduces the proposed method, BoFT, that labels images with non-user generated labels, BoFT labels, and how BoFT similarity, the pairwise similarity among users based on BoFT labels, is calculated.




A.Bag of feature-Based Tagging

Images are analyzed using BoFT, which annotates each image with a BoFT label. BoF is a popular computer vision approachfor analyzing images. It shows the key steps involved: Fig. 3(a) is the steps for BoF and it is the method for connection discovery based on user shared images. The different steps of BoFT are introduced in this section below.


*B.Extracting Feature:*

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 16, March 2016

Feature extraction is a process to obtain the unique local features in step 1 ofFig(a). These unique features can be detected by feature detection, such as the Harris Affine detector, Maximally Stable Extrenal Regions detector and KadirBrady saliency detector. The extracted features are relatively consistent across images taken under different viewing angles and lighting conditions

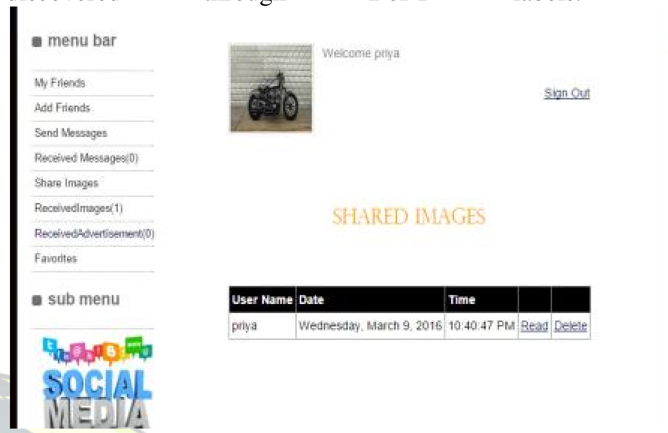

*C.Generating Codebook:*

Codebook generation in step 2 of Fig. (a) is a clustering process to obtain a set of visual words, a representative and distinct set of unique visual features. This step starts with clustering extracted visual features into groups by clustering techniques, such as -means clustering, based on their visual similarity, and the mean vectors of each group are defined as a visual word. Other possible techniques are the Canopy clustering algorithm and LindeBuzoGray algorithm. A -means clustering is used in our work.

D.Coding And Pooling: Feature coding represents each visual feature by the closest visual word. Each image is represented by a feature vector in the feature pooling, as shown in step3 of Fig. 3(a). One of the most common approaches is counting the number of occurrences of each unique visual word on an image as the feature vector.

*E.Connections Discovery with bOft Lables*

This section introduces how connections can be discovered through BoFT labels.




and User Profile: A user profile, which reflects the content of a user's shared images, is the key in connection discovery. The proposed method uses the number of occurrences of the BoFT labels in step 5 of Fig. 3(a) of the shared images of a user as his/her user profile, as in step 6 of Fig. 3(b). A user is represented by his/her user profile, , and the distribution of the BoFT labels that the user has is defined as (1) where is the number of occurrences of the -th label among the shared images of user , and is the total number of labels which is set to 500

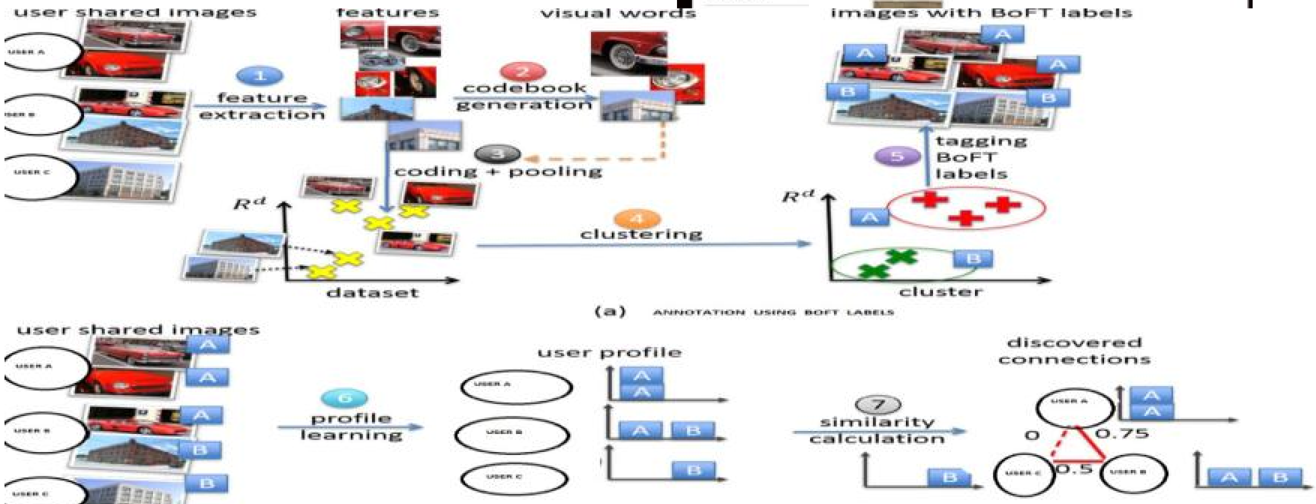
III.CLUSTERING AND BoFT LABELING:

Clustering groups images that are visually similar through the similarity in their feature vectors, which is shown in step 4 of Fig. 3(a). Christo Ananth et al. [4] proposed a system in which OWT extracts wavelet features which give a good separation of different patterns. Moreover the proposed algorithm uses morphological operators for effective segmentation. From the qualitative and quantitative results, it is concluded that our proposed method has improved segmentation quality and it is reliable, fast and can be used with reduced computational complexity than direct applications of Histogram Clustering. The main advantage of this method is the use of single parameter and also very faster. While comparing with five color spaces, segmentation scheme produces results noticeably better in RGB color space compared to all other color spaces.

More discussion of this can be found in Section VI. The next step, BoFT labeling, assignseach cluster a BoFT label so that those images with the same BoFT label are visually similar, and this is shown in step 5 of Fig. 3(a). The set of BoFT labels of user shared images of user is obtained. is a vector, with each element being the set of occurrences of a BoFT label in the shared images of

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 16, March 2016

user . The step is an unsupervised operation that analyzes user shared images without any manual input and process.




(a) ANNOTATION USING BOFT LABELS

(b) CALCULATING USER SIMILARITY

A.User Profile And Boft Similarity

When the user profile of each user is established, the next step is the connection discovery based on the BoFT similarity, , of users and , in which users who share highly similar images will have a high BoFT similarity. This requires a pairwise similarity comparison among user profiles based on the number of occurrences of BoFT labels

B.Measurements On User Shared Images

This section first describes the dataset, followed by the characteristics of the user shared images and follower/followee relationships. The third part of this section analyzes the BoFT similarity distribution by BoFT [23].

C.Datasets

Skyrock is a Western social network site that allow users to create blogs, follow other users and exchange messages. Most of the users of Skyrock are from European countries like France, England, German, Holland, etc.163Weibo was a microbloggingsocial network applicationfrom China with a similar mechanismto Twitter. As the user bases of the two social networks are different, it is interesting to observe whether the user behaviors in these social networks are similar. Fig. 4 shows the user interfaces of the two social networks, Skyrock and 163 Weibo. Skyrock, as shown in Fig. 4(a), users can share blogswith text, images and even video. On 163 Weibo, users share text and image content, as shown in Fig. 4(b). Similar to anysocial network, users of these networks can follow others to receive notifications of newly shared content from those they follow.

D.Characteristics Of User Shared Images

This section describes the characteristics of the user shared images and the follower/followee relationships. Fig. 5(a) andFig. 5(b) show the distribution of the number of user shared images a user has, and the frequency of this number, on Skyrockand 163 Weibo, respectively. It is observed that a few users share a large number of images, while most of the users share a fewimages only, and the same trend can be observed on both social networks. Fig. 6(a) and Fig. 6(b) show the distribution of thenumber of follower/followee relationships a user has, and the frequency of this number on Skyrock and 163 Weibo, respectively.The same trend that a few users have a large number of follower/followee relationships, while most of the users have a few follower/followee relationships only can also be observed. The same observation can be

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 16, March 2016

found in both social networks. It is obvious that the distribution of the number of shared images and follower/followee relationships on both social networks followsthe power law distribution, as do most social networks. It is concluded the selected users are a good representation of the users in the two social networks.

E. Similarity Distribution Boft

In this paper, there are two types of user pairs: related pairs, which are the pairs of users that are follower/followee, and nonrelatedpairs, which are the pairs in which a follower/followee relationship does not exist between the two users. Related pairs and non-related pairs can be considered as two classes,

F.Follower/ Followee Recommendation Using Connection Discovery

This section introduces the system flow and formulation of how follower/followee recommendation can be made with discovered connections. This is a 3-stage (stages A to C) systems as shown in Fig. 11. The first part is image collection, followed by connection discovery using BoFT. The third part focuses onhow to recommend follower/followees based on the discovered connections and the BoFT similarity distribution. The stages areintroduced one by one in this section.

G.Collecting Images

The proposed system carries out data collection as shown in step A of Fig. 11, which shows the process to collect user generatedimages from social media applications, such as Skyrock and 163 Weibo. The images can be provided by the operators of the social media and mobile applications or collected through the API of the social networks. The user generated images can be shared in various forms, such as posted images on social media or images shared through instant messaging applications.
On social networks such as Skyrock and 163 Weibo, user generated images are those images shared by users. This process is ongoing, which means that user shared images are collected continuously.

H. Connection Using Boft

The objective of the image understanding is to annotate user generated images with non-user annotated labels, as shown instep B of Fig. 11. The proposed system applies a computer vision approach to give a label to user generated images, which is not affected by the language, culture or other characteristicsof the user who shares the image, but is based on the image's visual appearance only. The accuracy of the user generated tagsis unreliable, sometimes even unavailable, and the performanceof connection discovery is affected. The proposed system appliesBoFT to annotate user generated images with non-user annotated labels, called BoFT labels. The set of user shared imagesof user is processed by the proposed method, and a set of BoFT labels, , is generated to represent user . As discussed,millions of images are generated every day, so a system that can process big data with scalable storage design is needed for collecting and processing these user shared images, such as a cloud-assisted system to handle profile learning and similarity calculation [38] for a scalable system. The feature vectors are first split into multiple blocks in the Hadoop Distributed FileSystem (HDFS) and distributed to virtual machines (VMs) for the -means clustering process. Each VM is in charge of computing the distribution of different labels for several users, and the BoFT similarity is also calculated in a distributed way.

I.Recommendations Of Follower

Follower/followee recommendation is one of the most popular applications on social media. The probability that two usersare a related pair, or , given the BoFT similarity of userand , can be calculated by (2) based on and Follower/followee recommendation should be made based on , from the highest to the lowest. This sectionstarts with discussions on how , can be formulated and calculated by the proposed system based on themeasurements followed by how recommendations can be madefrom the measurement. By Bayes' theorem,

IV.EXPERIMENTAL RESULTS

This section introduces how the experiment is conducted, followed by the experimental results with two showcases. A discussionof the results concludes this section.

Setup

Based on the observation that user pairs with a higher BoFT similarity are more likely to be follower/followee, discovered connections can be evaluated as a follower/followee recommendation system using . Fig. 12 shows the experiment setup for the evaluation with user shared images from Skyrock and 163 Weibo. As in step 1 of Fig. 12(a), the user shared images are analyzed using BoFT, as in Fig. 3(a), and users are represented by user profiles and the distribution of BoFT labels that the user

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 16, March 2016

has, as in Fig. 3(b). Connections are discovered by computing the pairwise by (2) with the user profiles. The list of users to be recommended to user , are ranked by in the discoveredconnections. By (16), the set of users are most likely to be follower/followees of user if they are users with the highest similarities. As a result, the set of users with the highest similaritiesare recommended to user , and the results are evaluated by two common metrics of prediction performance, recall rate and precision rate , for users with the highest similarities, as instep 2 of Fig. 12. The precision rate, , measures the percentage of discovered follower/followee relationships that exist in the ground truth, while is the percentage of existing follower/followee relationships in the ground truth that are recommended. A better discovery and recommendation method should give ahigher value of and . In order to evaluate the effectiveness of the proposed system,three connection discovery methods for follower/followee recommendation are implemented for comparison. The firstmethod is FoF, which is an achievable upper bound when difficult and limited access SGs are available. The recommendationis from the similarity of the SGs

RESULTS

The number of recommendations is set to be 5 to 10, to simulate a normal recommendation system; however, the same trendcan be found even when a smaller or a bigger number of recommendations is used. As the selected users are scraped randomly, the network densities are low, with 1.41% and 0.71% on Skyrock and 163 Weibo, respectively. The low network densities make the recommendation challenging, with Rand only able to achieve 1.41% and 0.71%. Fig. 13(a) and (b) show the of different methods against the number of recommendations,
on Skyrock and 163 Weibo, respectively. It is observed that the proposed method is at least 4 times better than Rand, and achieves 25% of the performance of FoF, the achievable bound of the discovery. Fig. 13(c) and (d) show the of different methods against the number of recommendations from 5 to 10on Skyrock and 163 Weibo, respectively. The same observation is found for . BoFT achieves 20% and 41% of the value of

| (a) | | |
|---|---|---|
| Process | Time (in sec) | |
| | Facebook | Twitter |
| Feature Extraction | 215,662 | 130,051 |
| Codebook Generation | 43,772 | 42,351 |
| Feature Coding & Pooling | 606 | 398 |
| (b) | | |
| Clustering | 71,723 | 61,586 |
| Profile Learning | 1,038 | 854 |
| Similarity Calculation | 788 | 1,089 |
| Total | 333,589 | 236,329 |

which it takes 583s and 265s for each user on 163 Weibo andSkyrock, respectively. A longer runtime is expected when more
users and user shared images are involved, and a big data systemis needed to handle the data.

*Showcase 1: Annotated tags by user*

Tagging is a popular feature on many social networks today, such as Flickr. Flickr is an image oriented social network, thatfocuses on the sharing of images, and all content shared involves at least one image. Follower/followee relationships can be predictedwith the discovered connections from the similarity of the user annotated tags on the shared images between two users. It isinteresting to compare the effectiveness of the proposed method on an image oriented social network, using user annotated tagsto calculate the user similarity, UserT. A comparison between UserT and the proposed method on Flickr is evaluated with over201,006 user shared images from 562 users with 902 relationships, as used in [23], and is shown in Fig. 14.

*Showcase 2: Discovering Gender*

Profiles on social media are also important for applications but are not always available. Among the information in profiles, gender is interesting, as it is useful for recommendation.Another showcase using the same Flickr dataset is condutedshow how gender can be identified with discovered connections. 445 out of the 562 users provide their gender and of these there are 79 females and 366 males. In each trial, 50 females and 50 males are selected randomly and is calculated. The experiment uses a 5-NN approach for identification

CONCLUSION

This work has proposed a connection discovery method and system for follower/followee recommendation based on user shared images. A practical method, BoFT, is discussed to label user shared images with BoFT labels on over 360,000 user shared images. The characteristics of user shared

ISSN 2394-3777 (Print)
ISSN 2394-3785 (Online)
Available online at www.ijartet.com
International Journal of Advanced Research Trends in Engineering and Technology (IJARTET)
Vol. 3, Special Issue 16, March 2016

images are then investigated and modeled as exponential distributions based on the analysis of 3 million follower/followee relationshipsfrom two social networks with different origins, Skyrock and 163 Weibo, for which similar observations are found. Basedon the observations, a practical follower/followee recommendation system is proposed and formulated with the discovered connections, which are extensively verified with ground truth. It is concluded that follower/followee recommendation usingdiscovered connections by user shared images is possible, and the recommendation is 60% better than UserT and achieves

## REFERENCES

[1] J. Chen, W. Geyer, C. Dugan, M. Muller, and I. Guy, “Makenew friends,butkeeptheold: Recommending people on social networking sites,in Proc. SIGCHI Conf. Human Factors Comput.Syst.,2009,pp. 201–210

. [2] V. Leroy, B. B. Cambazoglu, and F. Bonchi, “Cold start link predic- tion,” in Proc. 16th ACM SIGKDD Int. Conf. Knowl. Discovery Data Mining, 2010, pp. 393–402.

[3] I. Kahanda and J.Neville,“Using transactionalinformation to predict link strength in online social networks,” in Proc. 3rd Int. AAAI Conf. Weblogs Social Media, 2009, pp. 74–81.

[4] Christo Ananth, A.S.Senthilkani, Praghash.K, Chakka Raja.M., Jerrin John, I.Annadurai, “Overlap Wavelet Transform for Image Segmentation”, International Journal of Electronics Communication and Computer Technology (IJECCT), Volume 4, Issue 3 (May 2014), pp-656-658

. [5] D. Krackhardt and M. Kilduff, “Whether close or far: Social distance effects on perceived balance in friendship networks,” J. Personality SocialPsychol.,vol.76,no.5,pp.770–782,May1999.