Earn More Social Attention: User Popularity Based Tag Recommendation System

Xueting Wang
xt_wang@hal.t.u-tokyo.ac.jp
The University of Tokyo
Tokyo, Japan

Yiwei Zhang
zhangyiwei@hal.t.u-tokyo.ac.jp
The University of Tokyo
Tokyo, Japan

Toshihiko Yamasaki
yamasaki@hal.t.u-tokyo.ac.jp
The University of Tokyo
Tokyo, Japan

ABSTRACT

Enhancing social popularity of a post (i.e., the number of views or likes) on social network services is important for both ordinary users and companies who want to promote themselves. In this paper, we have implemented an online tagging support system to achieve this using an algorithm that recommends appropriate hashtags considering not only content popularity but also user popularity. The effectiveness of this technology has been verified by actually uploading photos with recommended hashtags on a real social network service.

CCS CONCEPTS

• **Information systems** → **Social networking sites**; *Information systems applications*; • **Human-centered computing** → **Collaborative and social computing**.

KEYWORDS

tag recommendation, social popularity, user-aware, tag ranking, social media, tagging system

ACM Reference Format:
Xueting Wang, Yiwei Zhang, and Toshihiko Yamasaki. 2020. Earn More Social Attention: User Popularity Based Tag Recommendation System. In *Companion Proceedings of the Web Conference 2020 (WWW '20 Companion), April 20–24, 2020, Taipei, Taiwan.* ACM, New York, NY, USA, 5 pages. https://doi.org/10.1145/3366424.3383543

1 INTRODUCTION

Online sharing services such as Flickr, Instagram, and Facebook have become the significant part of our modern lives. In these social media services, the number of views, comments, and favorites received after uploading indicate the popularity of the content, which can be referred to as "social popularity" or "social popularity scores" [11]. Although social popularity is critical for many individuals and corporations who want to acquire more attention, there are facts that only a small amount of social content became popular, while the vast majority of content can only attract limited attention. Moreover, there are many studies on predicting social popularity scores of posted content [2, 12]. However, the prediction performance is still limited and there are seldom researches working on how to enhance social popularity. Therefore, social popularity boosting is a very meaningful and still challenging task.

Users can add text tags to their posts in most social media services. These text tags play an important role as media content is often searched by text search engines. Actually, some works have shown that adding proper tags can affect social popularity significantly [9, 10]. However, tagging is a time-consuming process, and sometimes not an easy work for ordinary users who don't know what and how to tag. Therefore, we design this tagging support system to provide appropriate annotation recommendation, and even enhance the possibility of attracting more attention.

Traditional tag ranking and recommendation systems in social media are often designed to recommend semantically relevant or frequently co-occurring tags [3, 4, 11]. However, we focus more on extracting tags that have influence on social popularity. A tag recommendation algorithm for social popularity enhancement, called FolkPopularityRank (FP-Rank) [10], can recommend tags based on the social popularity of the posted content and the co-occurrence with tags. While our tagging support system can recommend appropriate and popularity effective text tags by using an algorithm [9], which is not only based on the information of media content, but also the social connection and popularity of users.

The contribution of this work can be summarized as follows:

- Our algorithm can recommend hashtags that be able to enhance the social popularity of posted contents, and the effectiveness is verified via uploading to a real social network service.
- The proposed online tagging support interface using the recommendation algorithm can help users select the appropriate tags by providing candidate tags that can enhance social popularity of their posts according to the original attached tags.

2 TAG RECOMMENDATION FOR SOCIAL POPULARITY BOOSTING

The proposed method can increase popularity by recommending tags according to the following concepts:

- Tags attached to content with high social popularity are important tags.
- The more tags attached to the content, the less influential to the popularity each tag is.
- Tags that co-occur with important tags are also important.
- Tags used by users with high social popularity are important.
- The more tags used by a user, the less influential each tag is, but tags used more frequently are more important.

WWW '20 Companion, April 20–24, 2020, Taipei, Taiwan

ACM ISBN 978-1-4503-7024-0/20/04.
https://doi.org/10.1145/3366424.3383543

The former three are based on the relationship among tags and posted content, which are inspired from the concepts of PageRank [5], a well-known web page ranking algorithm. As user popularity has not been investigated, we come up with the latter two concepts to include the relationship between tags and users.

To achieve these, the proposed tag recommendation algorithm consists of two steps: (1) a tag ranking step to calculate scores of tags by constructing a weighted adjacency matrix of tags from all posts in the source dataset; (2) a tag recommendation step to recommend new tags based on existing tags for posts in the target dataset.

2.1 Tag Ranking

2.1.1 Tag Scoring. In this step, we calculate a score representing the ability on affecting social popularity of each tag. Then the tags can be ranked based on the importance scores. The vector of scores of all tags $\mathbf{r}_{UFP}$ is calculated using weighted adjacency matrix of tags $\mathbf{A}_{UFP}$ by iterating from initial scores as follows:

$$\mathbf{r}_{UFP} = \alpha \mathbf{A}_{UFP} \mathbf{r}_{UFP} + (1 - \alpha)\mathbf{p}, \quad (1)$$

where α is the damping factor set to 0.85 same with PageRank in this study, $\mathbf{p}$ is a preference vector (random surfer component) representing the importance of each tag. The weighted adjacency matrix of tags $\mathbf{A}_{UFP}$ is a square matrix with the size of $T \times T$, T is the number of unique tags in the dataset.

The initial vector of $\mathbf{r}_{UFP}$ can be set as equal values for all tags with the same sum with the preference vector. $\mathbf{r}_{UFP}$ can converge after approximately 50 iterations, and 10 iterations are sufficient for practical application as discussed in [5]. We confirmed this and set the converge standard as the same max iteration times or when the change is smaller than a threshold.

2.1.2 Adjacency Matrix Construction. In a previous work, FP-Rank, the adjacency matrix of tags is calculated only considering the social popularity of posted content represented as $\mathbf{A}_{FP}$ [10]. In the algorithm, we define an adjacency matrix of tags $\mathbf{A}_{UP}$ weighted by social popularity of users. Here, $\mathbf{A}_{FP}$ and $\mathbf{A}_{UP}$ are the same size of $T \times T$. In this study, we propose the following matrix design called UFP-product-Rank as shown in Equation (2): element-wise multiplication of two matrices thus $\mathbf{A}^{m}_{UFP}$ will be weighted considering social popularity of both content and users, while co-occurrence among tags only exists when they are attached to both the same user and the same posted content.

$$\mathbf{A}^{m}_{UFP} = \mathbf{A}_{FP} \odot \mathbf{A}_{UP}, \quad (2)$$

$$\mathbf{A}_{UP} = \mathbf{U}_p \times \mathbf{U}_t^T. \quad (3)$$

Different to FP-Rank ($\mathbf{A}_{FP}$), $\mathbf{A}_{UP}$ is calculated using the users' social popularity and tag usage frequency according to Equation (3).

$\mathbf{U}_p$ and $\mathbf{U}_t$ are $T \times N$ sized matrices, N is the number of users in the source dataset. The ith row vector of $\mathbf{U}_p$ is a set of users' social popularity scores using the tag i, normalized by the sum of scores in the row. In this paper, the social popularity of a user is the total number of social popularity of his/her posted content. Since the user can use a tag multiple times, the jth column vector of $\mathbf{U}_t$ is the frequency of each tag has been used normalized by the total usage frequency of tags used by the jth user. Consequently, the element a^U_{ij} of $\mathbf{A}_{UP}$ is calculated as follows:

Table 1: Overview of the source and target datasets.

Data Set	Source	Target
Total Number of Images	60,000	1000
Average Number of Views of an Image	13,139.5	-
Average Number of Tags of an Image	37.1	23.1
Total Number of Users	6462	-
Average Number of Images of a User	9.3	-
Average Number of Views of a User	122001	-

$$a^U_{ij} = \sum_{l \in \mathbf{L}} \frac{u^U(l) \times (\text{the usage frequency of tag } j \text{ by user } l)}{(\text{the sum of usage frequency of all tags by user } l)}, \quad (4)$$

$$u^U(l) = \frac{(\text{social popularity of user } l) + k}{\sum (\text{social popularity of users using tag } i)}, \quad (5)$$

where l is the index of the user simultaneously using the tags i and j, $\mathbf{L}$ is the set of users in the source dataset. $u^U(l)$ is the weight of tag i and j calculated by the popularity of user l. k is a parameter to prevent $u^U(l)$ from becoming 0.

2.2 Tag Recommendation

Similar to FolkRank [1], an expanded version of PageRank, we can recommend new tags through the following equation based on the tags already attached to the content:

$$\mathbf{w}_{UFP} = \mathbf{r}^1_{UFP} - \mathbf{r}^0_{UFP}. \quad (6)$$

The difference between $\mathbf{r}^1_{UFP}$ and $\mathbf{r}^0_{UFP}$ is the setting of the preference vector $\mathbf{p}$ in Equation (1) when calculating $\mathbf{r}_{UFP}$ for the target dataset using the constructed matrix $\mathbf{A}_{UFP}$. For example, when we generate $\mathbf{r}^1_{UFP}$, the tags already attached to the post are weighted as 1, and the others have a weight of 0 in the preference vector $\mathbf{p}$. For $\mathbf{r}^0_{UFP}$, we give equal weights to all tags in the preference vector and the sum of weights is the same as in the preference vector of $\mathbf{r}^1_{UFP}$. Therefore, the final ranking score vector of all tags $\mathbf{w}_{UFP}$ can represent the co-occurrence with existing tags that the ones co-occur with tags already attached will be ranked higher. We then can recommend the top tags according to the calculated ranking scores. Namely, our UFP-product-Rank does not simply recommend generally popular hashtags, but recommend influential hashtags to the popularity that are related to already annotated hashtags.

Both $\mathbf{r}^1_{UFP}$ and $\mathbf{r}^0_{UFP}$ are iterated until convergence. The resulting scores of $\mathbf{w}_{UFP}$ reflect the co-occurrence with the tags already attached, and their influence on the social popularity. Tags are ranked according to these scores and the top tags are recommended as new tags.

3 ONLINE EVALUATION

We conducted online evaluation on Flickr. We added ten recommended tags to the original tag sets for each method and randomly uploaded all the proposed and comparative methods one by one to avoid multiple identical images with tags recommended by different methods to be seen at the same. For each recommendation method,

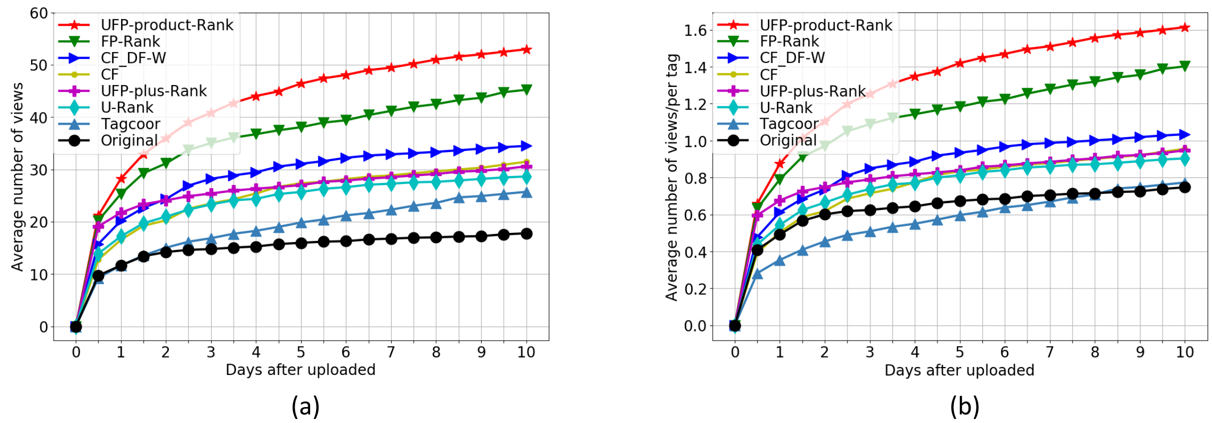

Figure 1: (a) Popularity change on the average number of views of each image with different methods lasting for ten days after uploading; (b) Popularity change on the average number of views of each image and each tag with different methods.

we record the number of views of all the images every 12 hour in ten days after uploaded independently.

3.1 Experimental Setting

In this study, the number of views is used as the measure of social popularity. For the source dataset to train the adjacency matrix of tags, we randomly select 60,000 images (uploaded by 6462 users) with over 20 tags and over 5000 views from Flickr's public dataset YFCC100M [8]. Consequently, there are over 254,000 unique tags included in the dataset, which is a broad resource for constructing a generalized matrix of tags. More details can be found in Table 1.

For testing, in the target dataset, content with annotated initial tags is needed for new tags recommendation. Regarding the cold-start problem in practical recommendation, we created a target dataset including 1000 images randomly selected from Wikimedia Commons[1] for testing. And then corresponding initial tags were generated according to the image content automatically by a computer vision API provided by the Microsoft Cognitive Services[2] (MCS). For each recommendation method, we selected the top ten ranked tags as recommendation added to the initial tags for uploading test. The detail of the dataset can be found here [3].

3.2 Comparative Methods

In the experiment, we compared the proposed methods with five other recommendation methods: (1) Original (MCS), the tags generated by an off-the-shelf computer vision API; (2) Tagcoor [6]; (3) Collaborative Filtering (CF) [3, 7]; (4) CF with DF-W (CF_DF-W) [11], and (5) FP-Rank ($\mathbf{A}_{FP}$) [10], considering only the relation between content and tags. (6) U-Rank ($\mathbf{A}^u_{UFP}$), baseline method considering only the relation between users and tags as shown in Equation (7); (7) UFP-plus-Rank ($\mathbf{A}^p_{UFP}$), baseline method considering social popularity of both content and users while co-occurrence among tags exists when they are used by the same user by using element-wise addition as shown in Equation (8); All the algorithms were implemented by ourselves.

$$\mathbf{A}^u_{UFP} = \mathbf{A}_{UP}, \tag{7}$$

$$\mathbf{A}^p_{UFP} = \mathbf{A}_{FP} + \mathbf{A}_{UP}. \tag{8}$$

3.3 Results Analysis

Figure 1(a) [9] shows the average number of views per image with each method for the test data. We can see that the UFP-product-Rank achieved the highest number of views when uploaded after ten days. The number of views of UFP-product-Rank is almost 2.8 times larger than that of the initial tags generated by a computer vision API provided by the Microsoft Cognitive Services (Original) and 1.2 times larger than (significantly higher by paired T-test, $p < 0.01$) that of the results of FP-Rank, which was the state-of-the-art method. In addition, all the user-aware proposed methods (UFP-product-Rank, UFP-plus-Rank, U-Rank) can improve social popularity from just using initial tags. These results verified the effectiveness of using users' social popularity for popularity enhancement in social networks.

By contrast, the recommendation only using users' social popularity (U-Rank) and weak co-occurrence among tags as long as they are used by the same user (UFP-plus-Rank), improved less than the content-aware FP-Rank and co-occurrence based CF_DF-W. It can be considered that the co-occurrence among tags in the U-Rank and UFP-plus-Rank is weakened from "attached to the same image" to "can be attached to different images as long as posted by the same user." Thus, some tags with less positive or even negative effects on popularity may be involved comparing to the FP-rank and CF_DF-W. Consequently, it can be inferred that the appropriate combination of popularity of users and content with strong co-occurrence among tags is important for social popularity enhancement.

To avoid the influence of the number of tags (ten more than the number of initial tags for each recommendation method), we divided the average number of views of each image by the number of tags with each method. The result is shown in Figure 1(b) [9]. We can find that the UFP-product-Rank still achieves the highest number of views over the other tag recommendation methods.

[1] https://commons.wikimedia.org/wiki/Main_Page
[2] https://azure.microsoft.com/en-us/services/cognitive-services/computer-vision/
[3] https://github.com/xueting-wang/UFP-Rank

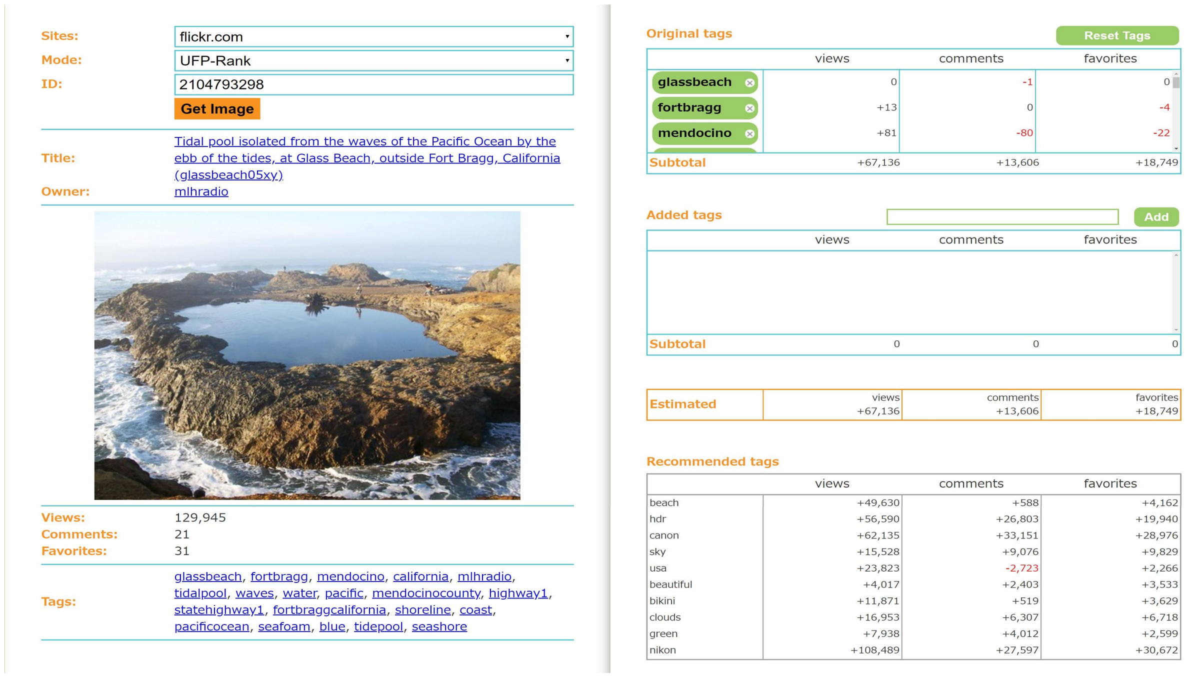

Figure 2: The overview of the demo system. The photo used is with Creative Commons license.

Moreover, our algorithm is able to be applied to other domains such as other social network services if the weighted adjacency matrix of tags can be trained using the data from that domain.

4 TAGGING SUPPORT SYSTEM

We developed a tagging support system to help users in the text tagging process in the social network. The system can predict social popularity. Furthermore, the most important, it can help people gain higher popularity by using UFP-product-Rank tag recommendation technology introduced above. Note that the hashtags are just recommended; they are not automatically annotated to the content to let users select the hashtags that they really want. This would also avoid converging to the same tag sets for all users.

4.1 Interface

Figure 2 shows a demo interface of the system. We first gather enough text tags to build the model. The weight vectors correlated to views, comments, and favorites are calculated according to a previous work [11], thus the predicted social popularity of each tag could be shown on the system. When indicating the image ID, which can be searched from the Flickr, the system can acquire the image along with some metadata including original tags, the number of views, comments, and favorites if it has. The system would offer ten recommended tags generated by UFP-product-Rank algorithm according to the original tags of this image. These recommendation and the predicted number of views, comments, and favorites are shown on the bottom. Users can modify tags to gain more attention based on the recommendation and predicted social popularity scores.

4.2 Requirements

Since this demo system is shown in an online web page, we can use our note PCs for demonstration so that only a desk and power supply are needed. Network environment is also preferred.

5 CONCLUSION

In this paper, we proposed a tagging support system that can recommend proper and influential hashtags to help users enhance social popularity of their social media content. In the future, we will also further consider how to support the creation of posting content with higher social popularity.

ACKNOWLEDGMENTS

This is a joint project between CyberBuzz, Inc. and Yamasaki lab of The University of Tokyo. This work was partially financially supported by the Grants-in-Aid for Scientific Research Numbers JP19K20289, JP19J22939 and JP18H03339 from JSPS.

REFERENCES

[1] Andreas Hotho, Robert Jäschke, Christoph Schmitz, and Gerd Stumme. 2006. Folkrank: A ranking algorithm for folksonomies. In *Proceedings of the LWA*, Vol. 1. 111–114.

[2] Feitao Huang, Junhong Chen, Zehang Lin, Peipei Kang, and Zhenguo Yang. 2018. Random forest exploiting post-related and user-related features for social media popularity prediction. In *Proceedings of the ACM International Conference on Multimedia (ACMMM)*. 2013–2017.

[3] Pasquale Lops, Marco de Gemmis, Giovanni Semeraro, Cataldo Musto, and Fedelucio Narducci. 2013. Content-based and collaborative techniques for tag recommendation: an empirical evaluation. *Journal of Intelligent Information Systems* 40, 1 (2013), 41–61.

[4] Hanh T. H. Nguyen, Martin Wistuba, Josif Grabocka, Lucas Rego Drumond, and Lars Schmidt-Thieme. 2017. Personalized Deep Learning for Tag Recommendation. In *Proceedings of the Pacific-Asia Conference on knowledge discovery and data mining*. 186–197.

[5] Lawrence Page, Sergey Brin, Rajeev Motwani, and Terry Winograd. 1999. *The PageRank Citation Ranking: Bringing Order to the Web.* Technical Report 1999-66.
[6] B. Sigurbjörnsson and R. van Zwol. 2008. Flickr tag recommendation based on collective knowledge. In *Proceedings of the World Wide Web Conference on World Wide Web (WWW).* 327–336.
[7] Xiaoyuan Su and Taghi M. Khoshgoftaar. 2009. A Survey of Collaborative Filtering Techniques. *Advances in Artificial Intelligence* 2009 (2009).
[8] Bart Thomee, David A. Shamma, Gerald Friedland, Benjamin Elizalde, Karl Ni, Douglas Poland, Damian Borth, and Li-Jia Li. 2016. YFCC100M: The New Data in Multimedia Research. *Commun. ACM* 59, 2 (2016), 64–73.
[9] Xueting Wang, Yiwei Zhang, and Toshihiko Yamasaki. 2019. User-Aware Folk Popularity Rank: User-Popularity-Based Tag Recommendation That Can Enhance Social Popularity. In *Proceedings of the 27th ACM International Conference on Multimedia (ACMMM).* 1970–1978.
[10] Toshihiko Yamasaki, Jiani Hu, Shumpei Sano, and Kiyoharu Aizawa. 2017. Folkpopularityrank: Tag recommendation for enhancing social popularity using text tags in content sharing services. In *Proceedings of the 26th International Joint Conference on Artificial Intelligence (IJCAI).* 3231–3237.
[11] Toshihiko Yamasaki, Shumpei Sano, and Kiyoharu Aizawa. 2014. Social popularity score: Predicting numbers of views, comments, and favorites of social photos using only annotations. In *Proceedings of the International Workshop on Internet-Scale Multimedia Management (WISMM).* 3–8.
[12] Wei Zhang, Wen Wang, Jun Wang, and Hongyuan Zha. 2018. User-guided hierarchical attention network for multi-modal social image popularity prediction. In *Proceedings of the International Conference on World Wide Web (WWW).* 1277–1286.