Analyzing Favorite Activities in Flickr

Social Media Analysis

Helmiriawan
Delft University of Technology
Student Number: 4623835
helmiriawan@student.tudelft.nl

ABSTRACT

Shared interest or appreciation toward a content in social media can be depicted from the number of likes or favorites. This number can also be used to measure the popularity of the content. It is normal for a user to seek for popularity or attention from other users when using social media. But what is the factor that can affect the number of like or favorite for a content? In this study, the favorite activities in Flickr, one of the most popular social media for photo sharing, was analyzed. Several aspects were investigated from a sample set of photos as the contents. The results revealed that several elements, such as comments, photos, tags, followers, and groups, influence the number of favorites in Flickr.

Keywords

Social media; Flickr; Favorite activity

1. INTRODUCTION

The number of photos which have been shared in social media is increasing over the year. In 2015, there were two million public photos uploaded in Flickr per day in average [1]. It is a common thing these days for people to upload photos in social media. One of the reasons is because they want to get attention from others [6]. Social media like Instagram, Flickr, and Pinterest provide ability for people to share their photos and interact with large number of users.

One type of interaction which is provided in social media is the ability to like or favorite the photos. This might indicate that a person has the same interest with the owner [4] or appreciate the quality of the images. The number of like or favorite can be used to measure the popularity of a photo [8]. However, a research has revealed the fact that popularity is not aligned with quality [6]. High quality photos might be remain unseen if the owner is not having high quality interaction with other users in social media. Maybe other users are not noticed those photos because it is hard to find them through search. A study has mentioned that users can use tags in their photos to gain greater attention from others [9].

Beside personal satisfaction, there are also other reasons why the users seek for more attention in social media. It is because social media nowadays has become one of the powerful tools to gain more exposure for business purposes. Apart from youth or common users, there are also professional photographers who use social media to show their artistry and enlarge their business power. These people use social media as a gallery for exposing their best photos to large number of people. At the same time, they can also interact with their audience quickly and directly, in order to get immediate response from admirer and potential clients.

For big companies, social media can also be one of the widest window for advertising, yet the cost might be lower compared with other media such as TV or magazine. Artists, musicians, or filmmakers can also use this media to announce their latest works or merely to build intimate interaction with their fans. These people can make their own page or ask popular users to advertise their products. With a lot of users in social media, they can attract more people to acknowledge their products and gain more potential buyers.

Studying like or favorite activities in social media can help us to understand more about characteristics of the users, in order to draw more attention from them. With the data that has been recorded in this media, it is easier and cheaper nowadays to analyze the behavior of people toward things that interesting for them. By having more comprehensive understanding about that, one can easily gain more benefit as mentioned previously.

Previous research has analyzed like activities in Instagram [4], to find what factors that might influence the number of likes in a photo. It is mentioned that the findings might be different in other social media because the user characteristics might be different as well. This study is intended to do similar thing in Flickr, as one of the most popular social media that has providing features for photo-sharing [3]. The following research question is the objective of this study: to what extent do user's other activities influence favorite activities in Flickr?

Basically, Flickr is not as popular as Instagram. But it has several advantage points [2], like allowing the user to track the camera details, organize the photos, and have more control over the privacy settings. To understand the factors influencing number of favorites in Flickr, an analysis from 6 thousand photos was conducted. Several activities such as comments, photos, tags, follows, followers, and groups were investigated. The last activity is one of the unique features in Flickr which is not available in Instagram, as can be seen in Figure 1.

The detail methodology of this study can be seen in Section 3. But before move into the detail, the related works were mentioned in Section 2. The result and discussion were explained in Section 4 and Section 5.

[1] https://www.flickr.com/photos/franckmichel/6855169886

[2] https://www.linkedin.com/pulse/flickr-vs-instagram-which-better-small-business-henny-kel

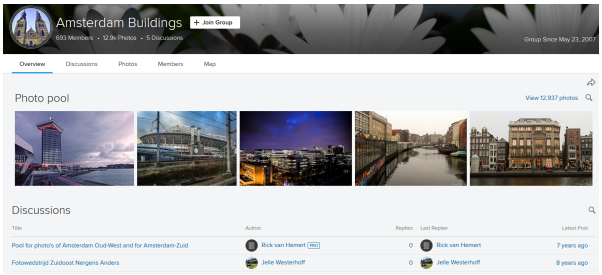

Figure 1: Sample group in Flickr

2. RELATED WORK

Previous study has shown that high quality photos in social media might be remain unseen if the poster is not having high quality interaction with other users [6]. Therefore, having high quality photo does not simply lead to high popularity. Trevisiol et al. [8] have analyzed image rankings by using favorites as the metric. It is mentioned in the paper that an image which is very popular in a group may received many favorites from other users. Meanwhile, another study has proposed an algorithm to predict the popularity of photos (i.e., the number of views, comments, and favorites) in Flickr [9]. By using recommended tags from the proposed algorithm, it is denoted that greater attention from other users can be gained.

In Instagram, a study has shown that photos, comments, tags, and followers are positively related to the number of likes [4]. The paper mentioned that tags has the highest influence among all of these elements. Meanwhile, other study has revealed that photos with faces are having greater chance to receive likes in Instagram [1]. However, the age and the gender of people in the photo do not have significant effect to the number of likes. Jang et al. [5] have denoted that teens are likely to get more likes, add higher number of tags, and have more comments compared to adults in Instagram. The analysis was because teens are more likely to express their emotions and social compared to adults.

Similar studies also conducted in other social media. In Twitter, factors impacting retweet were investigated [7]. The results indicate that among the content features in Twitter, hashtags and URLs are more correlated with retweetability. While in Pinterest, Gilbert et al. [2] have analyzed repins as the object in Pinterest. From this study, it is revealed that likes, comments, and followers influence the number of repins. In addition, it is also mentioned that women are more likely to receive repins compared to men.

3. METHODOLOGY

3.1 Data Collection

Like other social media, Flickr also provides an API that can be used for data collection. To speed up the processing time for analyzing the data, 6 thousand public photos posted from 4 to 6 January 2016 were randomly selected, with 2 thousand photos from each day [3]. Note that the data was the snapshot of 1 week after the photos were posted, because the actual number of each variable can be changed anytime. For future work, it is possible to collect higher number of photos and longer time period.

From the list of photos, several interesting elements were extracted by using the API. But unfortunately, Flickr does not provide API for retrieving number of user's followers. As an alternative, this data was crawled from the Flickr web page. But there is an issue when scraping the data from the web page. When the user has more than 1000 followers, the number will be rounded. Therefore, for higher number of followers, the information is less accurate.

Table 1: Variables Overview

Variable	Description
Photos	Poster's photos
Followers	Poster's followers
Follows	Poster's follows
Comments	Comments added to the photos
Tags	Tags added to the photos
Groups	Groups added to the photos
Favorites	Favorites added to the photos

From the metadata of the photos, the photo ID and user ID of the poster were extracted. By using both of these data, other information, such as photo information (comments, tags, groups, favorites) and poster information (photos, followers, follows), can be collected. Table 1 describes the variables that were considered and tested in this study to see if they influence the number of favorites in Flickr.

3.2 Data Analysis

This study used a statistical approach to investigate the factors that can influence the number of favorites. The prior assumption was a photo might received higher number of favorites if: the poster has many followers, the addition of many tags to the photo, the poster added the photo to many groups, and people favorite the photo while adding comments. For data analysis, a statistical approach called negative binomial regression was used. This approach has been previously used to understand the correlation among variables in other studies [1][2][4]. In this study, number of favorites was the dependent variable, while the factors which were considered in this study were the independent variables. SAS software was used to do the data analysis. For future work, other software such as STATA or R can be used as alternatives.

4. RESULT

Table 2: Descriptive Statistics of Dataset

Variable	Mean	S.D	Maximum
Photos	8481.79	31118.55	542654
Followers	144.26	495.65	8100
Follows	150.25	1187.3	38180
Comments	0.38	2.52	53
Tags	1.87	5.35	55
Groups	2.10	15.03	312
Favorites	2.48	15.14	393

Table 2 shows the descriptive statistics of the dataset which has been used in this study. It can be seen that there is a wide range of variances in all variables, denoted by

[3]Available at https://github.com/helmiriawan/IN4252

higher number of standard deviation compared to the mean. While a lot of photos have not received any favorites from other users, there is a photo which has 393 favorites within a week since it has been posted. There are only small number of photos which have higher number of favorites among all photos that have been investigated.

The result of the negative binomial regression can be seen on table 3. Note that the estimate of dispersion parameter is 5.95. This indicates that the data is over dispersed. Hence, using negative binomial model is more suitable than using Poisson model. It can be concluded that except number of follows, the other variables have significant effect on the dependent variable ($p < 0.0001$). This test revealed that comments has the highest effect to the number of favorites. The estimate value for comments (0.417218) means that for a one unit change in this variable, the number of favorites is expected to increase by 0.417218, given that the other variables remain constant. Commenting is another way for people to appreciate the content. It seems that people who add comment are also favorite the content.

Table 3: Result of Negative Binomial Regression

Variable	Estimate	S.D	p
Comments	0.417218	0.030411	<0.0001
Groups	0.052651	0.005189	<0.0001
Tags	0.041890	0.008049	<0.0001
Followers	0.002443	0.000210	<0.0001
Photos	-0.000032	0.000004	<0.0001
Follows	0.000269	0.000139	0.0524

Similarly, the number of favorites is expected to be increased by 0.052651, 0.041890, and 0.002443 for every one unit increase in groups, tags, and followers respectively, while holding the other variables in the model constant. In Flickr, users with the same interest can create a group to share their photos and hold discussions around a common theme. Because they are having similar interest, then the photos that have been shared in a group are more likely to receive favorites from people in there. As in Instagram and Twitter, tags can also help the users to receive more favorites, since it can make the photo easier to be discovered through search. The last variable indicates that users are more likely to favorite the photos which have been posted by people whom they are following. Thus, having more followers means greater chance to receive favorites.

Regarding photos, this variable shows a negative correlation (-0.000032) to the number of favorites. It means that having more photos may lead to lower number of favorites. This finding can be partly explained by the fact that Flickr gives a lot of free space for the users to upload their photos. It makes Flickr being an alternative for people to save some space on the their personal storage. Therefore, it is make sense if there are many users who have a lot of uninteresting photos with low number of favorites, since they do not want to show off and just want to use Flickr for personal use. The small number indicates that the number of photos posted was significantly higher than the number of favorites activities. Lastly, following other users does not statistically affecting number of favorites received ($p = 0.0524$). This may indicate that increasing or decreasing number of follows does not have any impact to number of favorites received.

5. DISCUSSION

5.1 Analysis of Favorites Activities

Based on literature review, it seems that less studies have been conducted to analyze characteristics of users in Flickr compared to other social media such as Instagram or Twitter. The fact that Flickr was launched before Instagram and Twitter does not make it the most popular social media among them. However, this study has added small insights about users activities in Flickr. This might help people who want to know more about characteristics of the users in this social media.

In summary, the statistical results indicated that number of favorites is statistically related with comments, groups, tags, followers, and photos to different extent. Particularly, comments showed the highest impact toward number of favorites, followed by groups, tags, followers, and photos. Like Instagram, the users in Flickr do not have to follow a user in order to view or favorite the photos. This might explain why groups and tags are more related to number of favorites compared to followers. Surprisingly, groups is statistically having more influence compared to tags. This revealed interesting fact since there is no such feature in Instagram. Previous study [4] has shown that followers are having the highest influence toward number of likes in Instagram, followed by photos and comments, which is kind of different with the results of this study.

5.2 Threats to Validity

In this study, the threats were related to data collection process, variables selection, and validity of the statistical conclusion. The data about user's followers was not accurate. Because in the data source, the number was rounded when a user was having more than 1000 followers. The reason why this study chose 1 week as the threshold for the snapshot was for giving other users to favorite the photos. Thus, this might created bias in other variables, such as number of photos, follows, and followers, since the numbers can be changed anytime. The variables that can be analyzed were not limited to what have been investigated in this study. There were still other interesting variables such as camera brand, gender, or location where the photo was taken. Different statistical approach might also be used in this study, but it does not guarantee that the result might be remain the same.

5.3 Future Work

Other statistical methods can also be used for this study, to see the difference when using different statistical approaches. Data analysis in this study only used quantitative approach. Perhaps, the text which has been used for tags or the type of group that has been be added can present additional insights about favorite activities in Flickr. Hence, applying qualitative approach in this study might be interesting for future work.

Flickr provides a lof of API methods for data collection. The information about the EXIF tags in a photo is also provided. Thus, the brand and specification of the camera can also be analyzed by using this data. Another study can also be conducted such as find out favorite places, favorite events, or favorite scenes in Flickr. However, different statistical method might be needed to analyze these categorical data.

6. CONCLUSION

Flickr is one of the most popular social media for photo-sharing. This paper has investigated several elements that might influence the number of favorites in Flickr. From the dataset of 2 thousand photos, several variables were extracted and analyzed by using negative binomial regression. From the statistical result, it is revealed that comments has the highest impact toward number of favorites, followed by groups, tags, followers, and photos. It can be concluded that the user can use these elements to gain higher number of favorites in Flickr. For future work, using qualitative approach or analyzing other variables might be interesting to do.

7. REFERENCES

[1] S. Bakhshi, D. A. Shamma, and E. Gilbert. Faces engage us: Photos with faces attract more likes and comments on instagram. In *Proceedings of the 32nd annual ACM conference on Human factors in computing systems*, pages 965–974. ACM, 2014.

[2] E. Gilbert, S. Bakhshi, S. Chang, and L. Terveen. I need to try this?: a statistical overview of pinterest. In *Proceedings of the SIGCHI conference on human factors in computing systems*, pages 2427–2436. ACM, 2013.

[3] K. Gopalakrishnan, A. Pandey, and J. Chandra. Social interaction in the flickr social network. In *2016 8th International Conference on Communication Systems and Networks (COMSNETS)*, pages 1–6. IEEE, 2016.

[4] J. Y. Jang, K. Han, and D. Lee. No reciprocity in liking photos: Analyzing like activities in instagram. In *Proceedings of the 26th ACM Conference on Hypertext & Social Media*, pages 273–282. ACM, 2015.

[5] J. Y. Jang, K. Han, D. Lee, H. Jia, and P. C. Shih. Teens engage more with fewer photos: Temporal and comparative analysis on behaviors in instagram. 2016.

[6] R. Schifanella, M. Redi, and L. Aiello. An image is worth more than a thousand favorites: Surfacing the hidden beauty of flickr pictures. *arXiv preprint arXiv:1505.03358*, 2015.

[7] B. Suh, L. Hong, P. Pirolli, and E. H. Chi. Want to be retweeted? large scale analytics on factors impacting retweet in twitter network. In *Social computing (socialcom), 2010 ieee second international conference on*, pages 177–184. IEEE, 2010.

[8] M. Trevisiol, L. Chiarandini, L. M. Aiello, and A. Jaimes. Image ranking based on user browsing behavior. In *Proceedings of the 35th international ACM SIGIR conference on Research and development in information retrieval*, pages 445–454. ACM, 2012.

[9] T. Yamasaki, S. Sano, and K. Aizawa. Social popularity score: Predicting numbers of views, comments, and favorites of social photos using only annotations. In *Proceedings of the First International Workshop on Internet-Scale Multimedia Management*, pages 3–8. ACM, 2014.