修　士　論　文

Tag Analysis and Recommendations in Social Media for Popularity Prediction and Boosting

ソーシャルメディアにおけるコンテンツの人気度の予測・向上に向けたタグ解析と推薦

指導教員　**山﨑 俊彦　准教授**

東京大学大学院
情報理工学系研究科
電子情報学専攻

氏　名　48-156447　**胡 佳儤**

提 出 日　2017 年 2 月

Abstract

Social media has been a huge part of our daily lives and it is vital to understand what makes online content popular and evaluate popularity growth. For better categorize and retrieve online content, most social media allow users to add open-ended tags to annotate online content. Tags become important user defined data and have been proved to be good predictors for social popularity. However tags used in social sites differ greatly from person to person, and the quality of tags varies widely. Apart from that, tagging is time consuming and the majority of ordinary users usually do not add many tags. Thus our work aims at supporting users during the tagging process with the purpose of gaining more attention.

In this thesis we propose two tag ranking algorithms, (Document Frequency-Weights from regression) and FolkPopularityRank, which can extract tags greatly influencing popularity. We then present three applications of the proposed ranking methods. Firstly, we investigate the spatial and temporal changes of influential tags and explore the evolution of community focus and user interests. Secondly, we apply the influential tags to social popularity prediction and show the efficiency of our proposed ranking methods. Thirdly, to support users in the tagging process, our proposed ranking methods can also be used for tag recommendations, which recommend tags that not only for appropriate annotations but also for popularity boosting.

Contents

List of Figures

List of Tables

Chapter 1

Introduction

1.1. Social Media

Social media is a group of Internet-based applications that build on the ideological and technological foundations of Web 2.0, and that allow the creation and exchange of user-generated content [1]. Users are free to interact and share ideas, pictures, posts, activities, events, and interests in virtual communities and networks. Each user is a contributor, instead of a passive viewer of web contents.

Social media has changed the world greatly and becomes the significant part of our modern civilization. Since 2013, over 500 million tweets have been sent per day in Twitter[1][2]. Around one million photos were uploaded to Flickr[2] every day in 2014 [3]. According to YouTube's[3] latest statistics of 2015, the site has over a billion users – almost one-third of all users on the Internet[4].

One key characteristic of social media is the fast content production rate, strikingly different compared to other non-user generated services [5]. It requires less production efforts and every registered user can publish their own production. Statistics of several major social media was summarized in Table 1.1, which clearly shows the significant popularity among all the Internet users and explosive increase in size.

Another characteristic of social media is the inequality of popularity. As we can see from Figure 1.1, 10% of the top popular Youtube videos account for nearly 80% of views, while the rest 90% of the videos account for a very small number of requests [5]. It also occurs to other social media, that a small amount of social contents become popular, while the vast majority of contents can only attract limited attention, which is the so-called long tail distribution. Social sharing sites often rank and categorize content based

[1] https://twitter.com/
[2] https://www.flickr.com/
[3] https://www.youtube.com/

TABLE 1.1 – Statistics of Major Social Media published in 2015.

Social Media	Founded	Content	Monthly Activate Users	Content Production Rate
Flickr[3]	2004	Photos	112 million	1 million/day
Youtube[4]	2005	Videos	1 billion	300 hours/minute
Twitter[2]	2006	Microblogs	320 million	500 million/day
Instagram[6]	2010	Photos	400 million	80 million/day

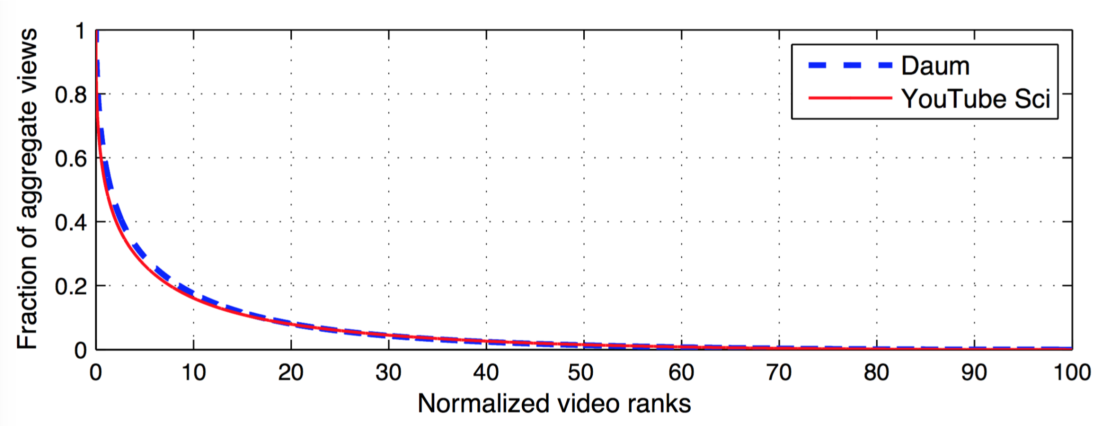


FIGURE 1.1 – Skewness of user interests across Youtube videos.

on past popularity and user appeal, to select published content favored by as many visitors as possible. For example, every day Flickr choose 500 most "interesting"[4] of the newly uploaded images to feature on the Explore page. This placement results in a positive feedback mechanism leading to rich-get-richer attention accrual for the very popular items, though the pattern pertains to only a small fraction of the submissions that rise to the top. Such dynamics possibly confine some valuable content in the very tail of the popularity distribution [7].

1.2. Tag and Folksonomy

The explosive increase in size of online social networks and the availability of large amounts of shared data have made it a big challenge to search for online resources. To better categorize and retrieve online contents, especially images and videos, most social sharing media allow users to add freely chosen keywords (tags) that match their real needs, tastes, or language to online contents. The resulting assemblage of tags form a "folksonomy", a conflation of the words 'folk' and 'taxonomy' [8]. Folksonomy

[4]https://www.flickr.com/explore/interesting/

has proven to be a practical solution to large-scale retrieval systems, in contrast to a top-down taxonomic classification, which lack flexibility and are generally expensive to create and maintain. In addition, tagging is easy to understand and apply, even without training and previous knowledge in classification or indexing.

Folksonomies come in two forms, depending on the underlying tagging rights [9]. Broad Folksonomies, such as Delicious, where items can be tagged by the entire community. Users of broad folksonomies usually assign tags, by their own vocabulary, for personal organization and retrieval, as a result of which the whole community can benefit from the richness of broad folksonomies. On the contrary, narrow folksonomies restrict the tagging of online contents to only a limited number of users. A common example of this type is Flickr, where photos are annotated by their uploader. Recent researches reveal that in such community, users annotate their photos with the motivation more to make their contents better accessible to the general public and gain reputation in the community, instead of personal organization and retrieval [10]. Many expert Flickr users pay considerable effort on assigning tags to their photos, sometimes as many as 70 [11].

1.3. Study of Social Popularity

It is widely perceived that understanding the popularity characteristics is important, since it can help to make good recommendations for users to reduce the clutter and find the most valuable contents. Companies can also rely on popularity analysis results to understand user behavior, which greatly affects the strategies for marketing and target advertising. Moreover, content-distribution networks can proactively allocate resources according to the future users' demand. On the other hand, one of the common use of social media is undoubtedly to accrue social capital by attracting as much attention as possible. However, only a few expert users and the contents they produce become popular, while the great mass of ordinary users can only gain limited influence. Popularity prediction results of a web content before published, can assist users to make wise choices when uploading contents.

Studying the popularity of web content, however, is a challenging task. Unlike traditional professionally generated content (e.g., TV programs, movies), the explosive increase of shared data has made it a really difficult task to rank and classify attractiveness of user generated contents in social media. Besides, different factors known to influence

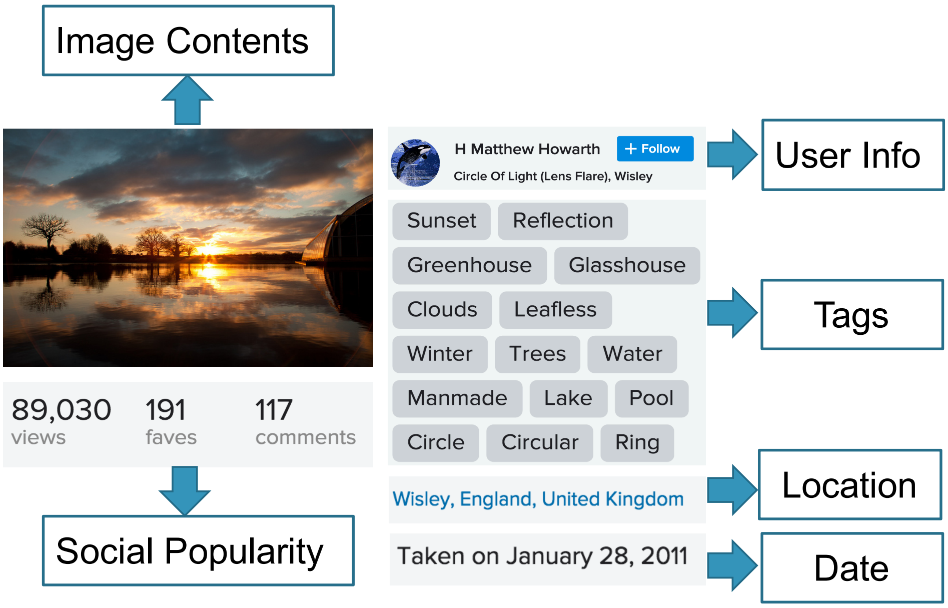


FIGURE 1.2 – An example of Flickr.

content popularity, such as content quality, user interactions in social media, social tagging and other context information, are difficult to measure.

In recent years, significant effort has been expended in evaluating and predicting social popularity, using various features. Figure 1.2 shows an example image of Flickr, there are not only the visual image content and popularity scores, but also other metadata in the social sites. social connections [12, 13, 14], textual information [13, 14, 15] and image content [13, 14, 15]. Image visual features, like color patches and coarseness, can be views as a direct measure of aesthetic, but are significantly hard to extract and require high computational cost. It also has been pointed that popularity is distinct from intrinsic quality[7]. Many high-quality content are confined in the very tail of the popularity distribution[16]. Similarly, explicit and implicit social connections is difficult to model, and new users usually do not have many contacts which leads to a cold-start problem. On the contrary, tags share a close relationship with popularity since text search is one of the most common ways to retrieve web objects. It makes search engine easier to index and find your content by adding more tags. Therefore in this thesis, we focus on research of tag and content popularity.

1.4. Purposes and Contributions

Social tags not only provide meaningful descriptors of the uploaded photos and allow efficient index of online content, but also reflect user interests and characteristics, so they have become important data for the study of social media and user behavior. Different from other text information in ordinary documents and web pages, tags usually have the spatial and temporal dimension. Both user interests and tag frequency shift from time to time and place to place. By analyzing users' behavior in social tagging over time and location, we can explore evolution of trends in interests and communities.

Tags are closely related to popularity of social contents, both number and quality of tags can be good predictors for future popularity. Tags are easy to analyze and computationally efficient in contrast to visual features and social connections. So we try to use tags for popularity prediction before upload, which can assist users for appropriate annotation and foster the possibility to attract more attention.

There is no doubt that most users are eager to post interesting contents to win as much attention as possible. However, only a few expert users and the content they produce become popular, while the vast majority of ordinary users can only gain limited influence. Some expert users have realized the importance of tagging, though in most cases ordinary users add very few tags or even none at all. In Flickr, about 20% of public photos have no tag at all and those with 1-3 tags take 64% of the images with any tags [17]. One of the main reasons is that users are often reluctant to enter many useful tags or indeed any at all. Tagging an image takes considerably more time than just selecting an image for upload. One the other hand, some photo sharing services like Instagram, only 30 or less hash tags can be assigned for annotation, so tags should be carefully selected. It is doubted whether ordinary users have sufficient knowledge to use precise words to describe their photos, let alone to choose tags that are capable of attracting attention and boosting popularity. Thus our work aims at supporting users during the tagging process with the purpose of gaining more attention.

We briefly summarize the main contributions of our work in the following:

- Our work is one of the first to study popularity on image/video sharing sites by analyzing text tags. We propose two tag ranking methods, DF-W and FolkPopularityRank, which can extract tags that highly influence content popularity. We apply the two method to a real-world dataset of Flickr, and compare the top ranked tags by different ranking methods.

- We investigate the spatial and temporal changes of tags that significantly affect popularity, and explore the evolution of community focus from time to time, place to place. It allows us to look into difference of online user behavior and increase the effectiveness of popularity study. Besides, we analyze top ranking tags over various image/video sharing media and reveal main differences among mainstream media.

- We present an approach to predict popularity of web content before uploading them to social services based on our proposed tag ranking methods. We contrast our methods with simpler tag rankings and experiment results show the efficiency of our methods in popularity prediction.

- To improve accuracy, we apply a multimodal learning approach for popularity prediction, which attempts to combine both of tag and visual features. And our work reveals that tag feature is much powerful than visual feature in popularity prediction.

- We present two new tag recommendation strategies based on the two proposed tag ranking methods, which focus more on popularity boosting rather than merely semantics and descriptive annotations. A series of experiments are designed to compare the performance of various tag recommendation methods in popularity boosting.

- We develop a popularity estimation and tag recommendation system, which can assist users to estimate how popular the uploaded images or videos will be and recommend tags that can help to gain more social attention.

1.5. Organization of This Thesis

The rest of the thesis is organized as follows. We begin by introducing related works on social popularity, popularity prediction, and tag recommendations in Chapter 2. Chapter 3 describes the two proposed tag ranking algorithms, DF-W and FolkPopulairtyRank, in details. The next four chapters present the experiments we conducted, consisting of tag ranking by different methods on Flickr in Chapter 4, spatial and temporal analysis of influential tags in Chapter 5, popularity prediction based on top ranked tags in Chapter 6, and tag recommendations for popularity boosting in Chapter 7. Chapter 8 concludes this thesis and shows directions of our future work.

Chapter 2

Related Works

2.1. Social Content and Its Popularity

There are multiple social sites sharing a variety of types of content, and different contents have been collected for recent researches, including online video, bookmark, photo, and microblog. All social sites are unique in their own way, so in the this sections, we briefly introduce several major social media with different types social content.

Online videos, accounting for a significant amount of Internet traffic, have been one of the main focus of the existing studies [5, 18, 19]. YouTube, the world's largest video sharing platform, allows registered users to upload videos, watch videos, subscribe others' channels, and leave comments. It has reportedly severed over one billion users in 2015, and every day people watch hundreds of millions of hours on YouTube and generate billions of views [4]. Researches have been done to analyze the popularity growth patterns, and it is said that videos tend to get most of their views much earlier in their lifetimes [5, 18]. If a video did not get enough attention during its first days, then it is unlikely that it will get many requests in the future [5].

Social bookmarking sites such as Digg[1], Delicious[2], Reddit[3], are developed for storing, sharing, and discovering web bookmarks. Unlike file sharing, these sites does not save the resources themselves, but merely bookmarks that reference them, i.e. a link to the bookmarked page. Users are allowed to add descriptions and tags to these bookmarks, so other users may understand the content of the resource without first needing to download it for themselves. It is observed that contents published on Digg experience an even greater rate of change with stories reaching their attention peak in the first six hours after publication and being completely saturated within one day, which results

[1]http://digg.com/
[2]https://delicious.com/
[3]https://www.reddit.com/

from the fact that most of all bookmarks were links to breaking news, fleeting Internet fads or technology-related themes with a naturally limited time for user appeal [18].

Another major type of social media is the photo-sharing site such as Flickr and Instagram[4]. These applications allow users to upload photos, view and comment photos created by others, etc. As is common to other social media, photo-sharing sites also allow users to designate others as "contacts" and to track their activities in real time. The contacts interface on Flickr enables users to see latest images submitted by their friends. Unlike Flickr, which offers more professional-oriented features, Instagram, being designed for mobile users, resembles an amateur photo-blog [20]. It incorporates features to quickly take photos and apply visual effects.

Microblogs, such as Twitter and Weibo[5], are a specific type of social media that have been extensively studied [21]. Registered users create and share information in the form of short messages, called tweets, which contain up to 140 characters. When a user posts a tweet it becomes visible to all its followers. As followers can further share the message to their own list of followers by retweeting. Twitter tracks phrases, words, and hashtags (i.e., a word with a '#' character) that are most often mentioned and posts them under the title of "trending topics" regularly. Similarly, a study on Twitter revealed that most tweets receive half of their retweets within the first hour after publication [22].

Besides, different sites define the popularity of social media content differently, such as the number of "views" on Yoube [5, 18], the number of "diggs" on Digg [18] or the number of "retweets" on Twitter [23]. It is difficult to precisely pick any singe one as the true notion of popularity. Different measures capture different levels of user engagement and provide valuable information: views imply the total requests of social contents, likes improves content quality, comments increase the time spent on a application, and sharing gives contents a greater notoriety [24]. In general, it has been pointed out that there is a moderate correlation between the different popularity measures [7, 25], as they probably capture different types of interests on the Internet.

2.2. What Makes Social Content Popular?

A huge amount of social contents are uploaded to the Internet every day. While some will have the widespread impacts on opinions, thoughts, and cultures, others are

[4] https://www.instagram.com/
[5] http://www.weibo.com/

completely ignored. Not all social content will reach the same popularity and have the same impact.

The magic formula of what makes social content popular is still unknown but some of the ingredients have been discovered. The social content features play a major role in its future success. For example, tweets contain URLs and hashtags have strong relationships with retweetability in Twitter [26]. It has also been pointed out that content that generates high-arousal emotions (e.g, awe, anxiety) disseminates faster on the Internet and captures a larger amount of users' interest [27]. Similarly, emotionally charged twitter messages tend to be retweeted more often and more quickly compared to neutral ones [28]. In addition, the inherent quality of the content [7, 29] and topics [30] are positively correlated with content popularity.

On the other hand, there are also several content-agnostic factors (such as dissemination mechanisms, social influence) that have a strong impact on the popularity growth [31]. In social media, there are basically three ways for users to find and view new social contents: (1) latest publication by their contacts or subscribed special interest groups (or so-called channels), (2) internal search engine by keywords, (3) recommendation system.

Users can view latest publication by their contacts or or subscribed groups in the timelines and anyone who finds the social content interesting can share it with friends. It is not difficult to understand that the social influence of content publisher make a difference to social popularity. The greater a social network of the publisher, which means that more users will have chance to see the item, the greater the increase in social popularity in early stages after publication. It is said that in Youtube uploaders of top-ranked videos have large social networks [31].

In Youtube, the internal search engine accounts for most of the views, followed by the recommendation system [19]. Thus these services also play an important role in attracting popularity. To better categorize and retrieve contents, most social media allow users to add annotations (e.g, title, description, open-ended tags) to the item. Thus, the well annotated contents are more likely to become popular. Lerman et al. studied users browsing behaviors in Flickr, and found that expert users shared images through over 100 groups and use many tags to increase the visibility of an image [11]. Besides, recommendation system also depends on these annotations to recommend similar contents to users. It will help for new social contents to gain more popularity by choosing relevant title and description with similar popular contents [32].

2.3. Predicting Popularity of Online Content

Much effort has been devoted to predict popularity of online content in the last decade. Tatar et al. propose a classification that groups the methods according to the type and the granularity of information used in the prediction process [24].

After publication: a groups of prediction models include data about the attention that one item receives after its publication. There are two main kinds of models: rich-get-richer model, and user behavior based model.

Before publication: predicting popularity before the publication of an item is one of the most challenging problems. It relies only on content metadata or the online social connections of the uploader.

An important observation of early studies was that there is a strong positive correlation between the popularity of a submission at different stages during its lifetime [18]. It implies a classical rich-get-richer phenomenon [33], which suggests that a submission will attract new attention at a rate proportional to the amount of attention already acquired. Many prediction methods, thus, use the amount of attention that a submission generates early as a predictor of future popularity.

Szabo and Huberman [18] propose a logarithmically transformed linear regression model (we call it log-linear model) to study the *cumulative growth* of attention, i.e., the amount of attention that a submission receives from the moment it was published until the prediction moment.

The log-linear model is simple and shows good predictive performance on Digg stories and Youtube videos [18]. However, even online contents that have very similar popularity in early time, will end up with very different total popularities. Different contents may experience very different popularity evolution patterns [34]. Thus Pinto et al. propose a multivariate linear regression model (ML model) that, building on the log-linear model, incorporates *temporal analysis* of how social popularity evolves over time [34]. They further propose a Multivariate Radial Basis Function model (MRBF model) [34] that aims to exploit the different *popularity evolution patterns* a submission can follow in a more explicit way, by measuring the similarity of a video and known examples from the training set, and changing the prediction based on this information. Pinto et al. evaluate the log-linear model, ML model and MRBF model on two datasets of Youtube videos, top set (videos in the top list) and random set (a random sample of Youtube videos).

Instead of treating each user's reaction equally in the prediction process, Lerman and Hogg describe a stochastic model of user behavior during a browsing session allows predicting social popularity based on early user reaction to new content on Digg [35]. The popularity of a story on Digg depends on the combination of its visibility and interestingness, with visibility coming from different parts of the Digg user interface: the friends page, upcoming page and front page list, and the position within each list. The authors validate the model on a small sample of Digg stories. By using stochastic model, the authors reveal that they can predict in 95% of the cases which stories will become popular enough to reach Digg's front page.

The after publication prediction models achieve good performance, though, it is most desirable to predict the popularity of social contents prior to their publication, fostering the possibility of appropriate decision making to modify an item and the manner of its publication.

Previous researches have tried to use various features to predict social popularity in different social media. Pedro et al. [15] exploit image features that affect visual quality (sharpness, colorfulness, saturation, etc) as well as textual meta data to categorize and rank photos according to their attractiveness. Khosla et al. [13] combine content features (such as GIST, color histogram, texture, color patches, etc) and social cues (mean views, photo count, groups, etc), and use linear support vector machines (SVMs) to predict popularity of Flickr photos before they are published. Prediction and evaluation were preformed on 3 different datasets namely one-per-user, user-mix, and user-specific, which simulate different user scenarios.

User characteristics are also important factors influencing popularity. Number of contacts and amount of contents [27, 36] has been uploaded are proved to be strong predictor of the popularity of tweets and Flickr images. In addition, uploader behavior and other characteristics of the user may also be useful predictors.

Furthermore, it provides a variety of context features in social media, which can be good predictor in social popularity prediction. Social tagging has become a popular means to annotate various social contents, and it provides meaningful descriptors of the objects, and allows search engine to organize and index social content. [37] tried to include tag features in the prediction models and the performance is promising.

2.4. Tag Recommendation

In the past few decades, recommender systems have been popular both commercially and in the research community, where many approaches have been suggested for providing recommendations. The well known algorithm, Collaborative Filtering [38, 39], has achieved widespread success and been one of the most dominant methods used in recommender systems, due to its simplicity and promising results. The main idea is to collect information about preferences or past behavior of an existing user community (collaborating) for make automatic prediction (filtering) about the interests of a specific user.

As tagging become a popular means to annotate various web resources, there have been many efforts to automatically recommend tags. One of the most frequently cited studies among folksonomy-based algorithms is FolkRank, which adapts the PageRank for a folksonomy space. The key idea of FolkRank is that a resource which is tagged with important tags by important users becomes important itself [40]. The same holds symmetrically for tags and users. A weight-spread ranking scheme is then employed, and from which the top n tags are selected for recommendation.

It is worth noting is that folksonomies come in two forms, depending on the underlying tagging rights [8]. FolkRank performs well in broad Folksonomies, such as del.icio.us, where items can be tagged by the entire community. In this paper, however, we put our focus on narrow folksonomies, which restrict the tagging of online contents to only a limited number of users. In case of Flickr, any particular object (an image) is only tagged by a single user (the owner). This has to be contrasted with the setting for broad folksonomies, as a result of which, the recommendation methods mentioned above cannot be applied to social media like Flickr. Thus Sigurbjörnsson et al. proposed a recommendation method based on tag co-occurance (we call it Tagcoor for simplicity), which can be applied to Flickr and is proved to be efficient in recommending relevant tags for photos with different levels of exhaustiveness of original tagging [17].

The existing recommender systems perform well in suggesting semantically proper tags and enriching the folksonomies, however, users may prefer tags that can help to greatly extend influence than those that only describe the photos. Therefore our work focus on making better recommendations with a higher level of influence on popularity boosting.

Chapter 3

Proposed Methods

3.1. The DF-W Algorithm

In previous work, we proposed a tag ranking method, called DF-W, which extracts tags that significantly affect content popularity and evaluating how these tags contribute to popularity. We would like to summarize it in this section. The DF-W algorithm is inspired by the term frequency-inverse document frequency (TF-IDF) algorithm, which is frequently used for calculating the importance of words in documents [41]. TF reflects how often the term (word) is used and IDF corresponds to the rareness of the word across all documents. TF-IDF is very powerful, but cannot be directly applied to tag analysis. Firstly, a tag appears only once at most in each content, thus TF is always one. Secondly, tags are sparse, making IDF meaningless.

In [42], therefore, document frequency-weights from regression (DF-W) is proposed. Assume we have T kinds of tags, $\{tag_1, tag_2, \cdots, tag_T\}$. First, document frequency (DF) is counted, which represent how popular each tag is:

$$DF_i = \text{The number of counts of the } i \text{ th tag in dataset}, \tag{3.1}$$

Then, a linear SVR model is trained by using the feature vector defined as follows:

$$FV_i^{tag} = \{f_{i1}, f_{i2}, \cdots, f_{ij}, \cdots, f_{iT}\}, \tag{3.2}$$

where FV_i^{tag} is the feature vector for the ith image and f_{ij} represents whether ith image has the jth most frequently appeared tag. The target value is the social popularity scores such as the number of views, comments, of favorites. After the training, we can obtain the weight vector W, which represents the normal vector of the hyper-plain of the T

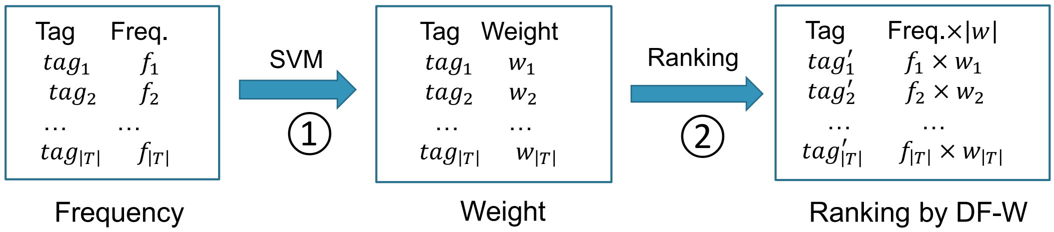


FIGURE 3.1 – The flow of DF-W.

dimension:

$$W = (w_1, w_2, \cdots, w_T). \tag{3.3}$$

It is natural to think that W can be used as a measure as importance score of tags and we can sort the tags in the decreasing order of weights. However, we find that a notable number of tags in such a top list appear only in a few images. This means that top tags obtained by this method are not appropriate for predicting the social popularity, because they are only useful for a limited number of images. Therefore, we define the importance score of tags as follows:

$$DF - W_i = DF_i \times |w_i| \tag{3.4}$$

We sort out all the tags by the importance score, and obtain the top N influential tags to predict social popularity. In Eq. 3.4, the absolute value of w_i is employed in order to include both positively influential tags and negatively influential ones. If users are interested only in positively influential tags, the raw value of w_i can be used. The advantage of the this algorithm is its simplicity and the tag scores and tag recommendations can be updated with negligibly small cost. The flow of this method is summarized in Figure 3.1.

3.2. FolkPopularityRank

Our algorithm is inspired by the PageRank and FolkRank algorithms. The main idea consists of two assumptions: (1) tags used for popular content are important, (2) the tags co-occurring with such important tags is also important. In contrast to FolkRank, we only consider the relation between content and tags. For a set of content $C = \{c_1, c_2, \ldots, c_{|C|}\}$ and a set of unique tags $T = \{t_1, t_2, \ldots, t_{|T|}\}$, a folksonomy is represented as a bipartite graph $H = (V, E)$, where $V = C \cup T$ is the set of nodes and $E \subset C \times T$ is the set of undirected edges.

We have thus a graph of vertices which are mutually reinforcing each other by spreading their FolkPopularityRank score, which indicates popularity influence of a particular tag. The final FolkPopularityRank score, $\mathbf{s}$ is computed as follows:

$$\mathbf{r}^1 = d\widetilde{\mathbf{A}}_{pop}\mathbf{r}^1 + (1-d)\mathbf{p} \tag{3.5}$$

$$\mathbf{r}^0 = d\widetilde{\mathbf{A}}_{tag}\mathbf{r}^0 + (1-d)\mathbf{p} \tag{3.6}$$

$$\mathbf{s} = \mathbf{r}^1 - \mathbf{r}^0 \tag{3.7}$$

where $\mathbf{r}$ is a weight vector with one entry for each tag, $\widetilde{\mathbf{A}}_{pop}$ is a column stochastic matrix of the $\mid T \mid \times \mid T \mid$ adjacency matrix $\mathbf{A}_{pop}$, and so is $\widetilde{\mathbf{A}}_{tag}$ of $\mathbf{A}_{tag}$. The entries of $\mathbf{A}_{pop}$ and $\mathbf{A}_{tag}$, represented as $a_{i,j}$ and $a'_{i,j}$ repectively, are defined as

$$a_{i,j} = \sum_{t_i, t_j \in c_k} \frac{Popularity\,(c_k) + 1}{number\ of\ tags\,(c_k)} \tag{3.8}$$

$$a^{'}_{i,j} = \sum_{t_i, t_j \in c_k} \frac{1}{number\ of\ tags\,(c_k)} \tag{3.9}$$

In addition, $d \in (0, 1)$ is a damping factor and $\mathbf{p}$ is a random surfer component.

For better understanding, the matrix $\mathbf{A}_{pop}$ can also be interpreted as a combination of two matrices by introducing the tag-content matrices:

$$\mathbf{A}_{pop} = \mathbf{B}_w \times \mathbf{B}_t^T. \tag{3.10}$$

where $\mathbf{B}_w$ and $\mathbf{B}_t$ are $\mid T \mid \times \mid C \mid$ row-stochastic and column-stochastic matrices respectively. The ith row vector in $\mathbf{B}_w$ is a set of normalized social popularity scores for the content tagged with the tag t_i. The jth column vector in B_t is a set of normalized tag assignments to the content c_j.

The concept of the procedures in eqs. (3.5) is visually explained by a toy example in Figure 3.2 and Figure 3.3. Let us assume that there are only three images and only three unique tags in the service. The initial importance scores for the tags are all set to 1. Image #1 is annotated with *bird* and *owl*, image #2 with *bird*, *owl*, and *fly*, and image #3 with *bird* and *fly*The images' social popularity scores (which include the numbers of views, comments, or favorites) are 40, 7, and 3, respectively. The tag scores are

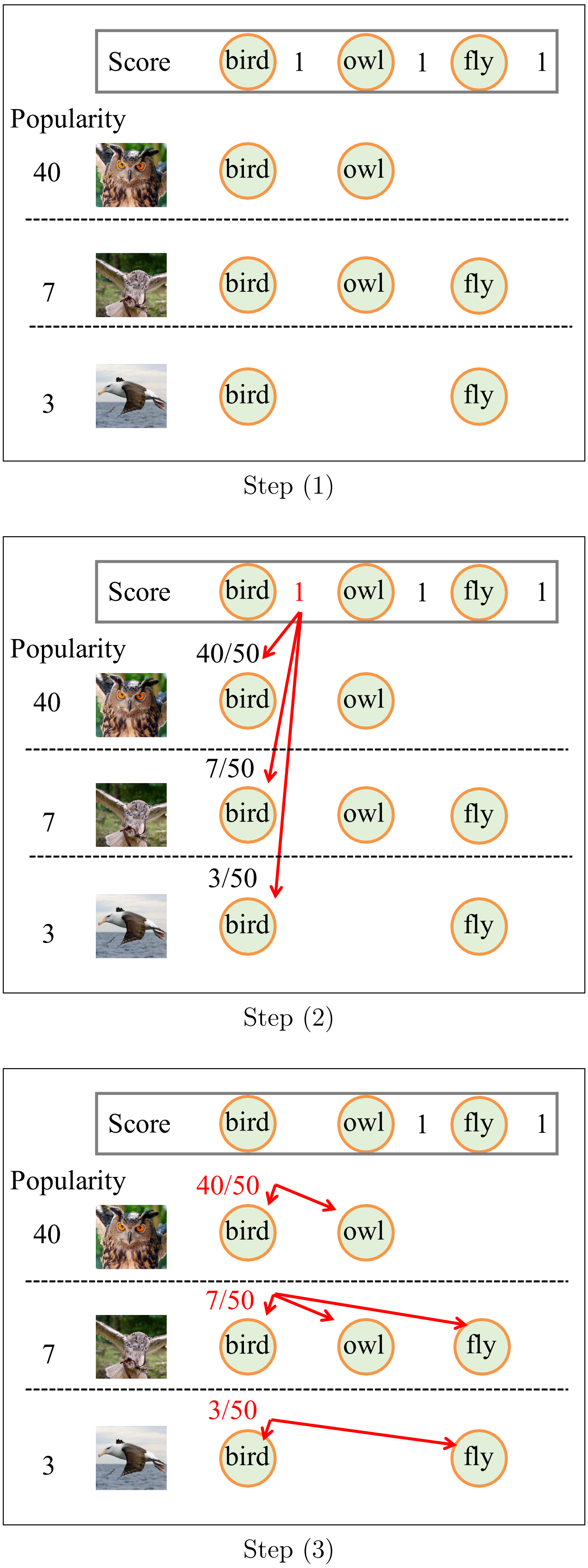
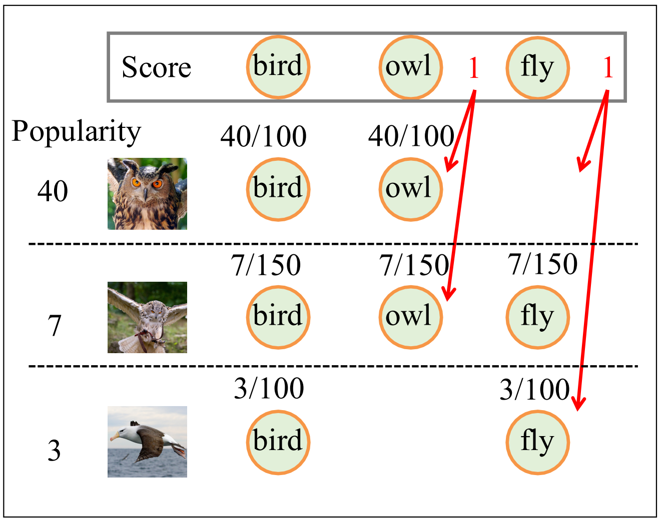
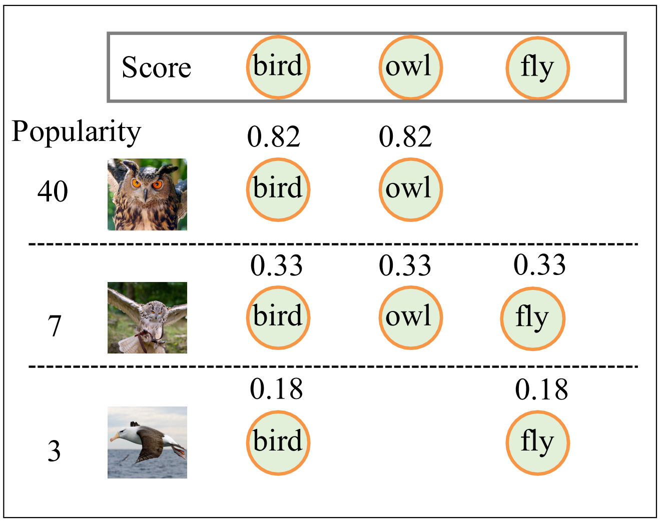


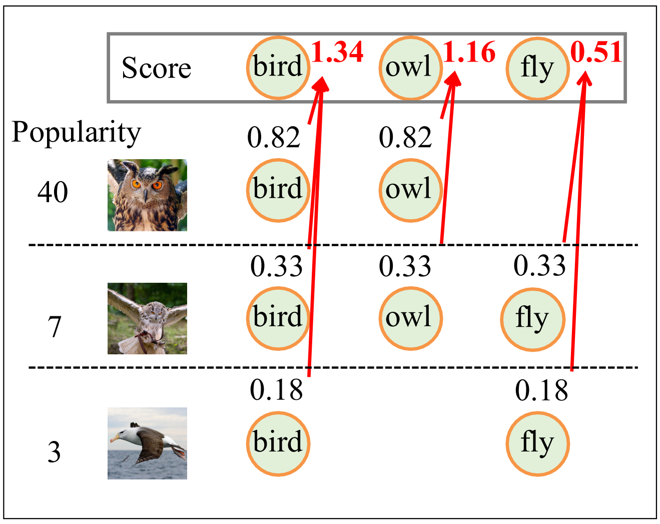
FIGURE 3.2 – A toy example to show how FolkPopularityRank redistribute scores for the tags (for the first iteration): Part I.



Step (4)



Step (5)



Step (6)

FIGURE 3.3 – A toy example to show how FolkPopularityRank redistribute scores for the tags (for the first iteration): Part II.

distributed to the images by considering the social popularity scores. They are collected back to the tags by considering the number of co-occurring tags. Then, the tag scores are updated by summing the scores of the tags. In this case, B_w and B_t are defined as follows.

$$B_w = \begin{pmatrix} \frac{40}{50} & \frac{7}{50} & \frac{3}{50} \\ \frac{40}{47} & \frac{7}{47} & 0 \\ 0 & \frac{7}{10} & \frac{3}{10} \end{pmatrix}, \; B_t = \begin{pmatrix} \frac{1}{2} & \frac{1}{3} & \frac{1}{2} \\ \frac{1}{2} & \frac{1}{3} & 0 \\ 0 & \frac{1}{3} & \frac{1}{2} \end{pmatrix}. \tag{3.11}$$

The iteration in eq. (3.5) is repeated until it converges. The related literature [43] points out that pagerank converges after approximately 50 iterations, although 10 iterations are sufficient for practical ranking systems. We confirmed that this is also the case with our FolkPopularityRank algorithm.

For recommendation, the random surfer component $\mathbf{p}$ for the already existing tags is set to 1, and the others are set to $0 \sim 1$. Setting the random surfer component in this manner causes tags co-occurring with the already existing tags to be extracted. If the random surfer components for the unused tags are 0, there is no chance of selecting random tags. The eq.(3.5) and eq.(3.6) are iterated until convergence. Thus, the final tag scores correspond to the influence of the tags' social popularity scores.

Chapter 4

Tag Ranking on Flickr

In order to evaluate the two proposed tag ranking algorithms, we performed experiments on a real-world Flickr dataset in the following chapters. In this chapter, we first describe the details of the dataset we used throughout this thesis in section 4.1, and how we defined popularity of web content in section 4.2. Section 4.3 contrasts the top-ranked tags extracted by our methods with different tag ranking methods.

4.1. Data Collection

We begin with a brief introduction of data collection we used. All the experiments are mainly conducted on the largest public multimedia collection of Yahoo! Webscope dataset YFCC-100M [44], which contains 100 million public Flickr photos (99.2%) and videos (0.8%). Each object in the dataset carries a Creative Commons license, and is represented by several pieces of metadata, e.g. Flickr identifier, title, tags, geo, date taken. The dataset provides a comprehensive snapshot of how photos and videos were taken, described, and shared over the years, from the inception of Flickr in 2004 until early 2014. Figure 4.1 shows the total number of objects in the YFCC-100M dataset uploaded to Flickr per month.

Since the dataset do not directly include the social popularity scores, we collected the numbers of views, comments, and favorites from Flickr API [45] during a period from December 1st to December 19th, 2014. And in our following experiments, we only consider photos or videos that have been annotated with user tags and still available in the public domain during that period. A summary of the dataset is shown in Table 4.1.

To better understand the dataset, first we investigate how tags are used in Flickr and how many tags, views, comments and favorites each object received. Table 4.2 shows the statistics of the number of tags, views, comments and favorites of each object in the

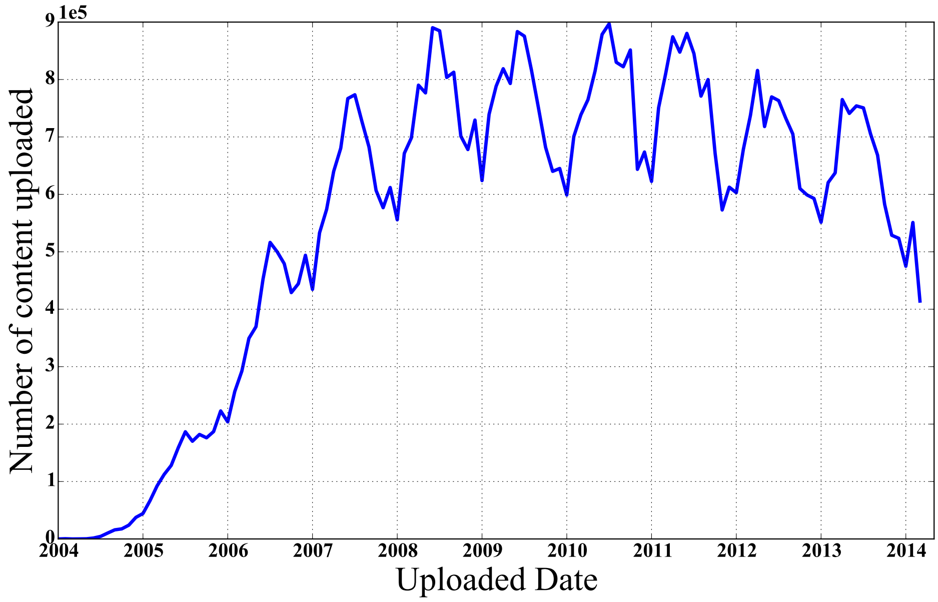
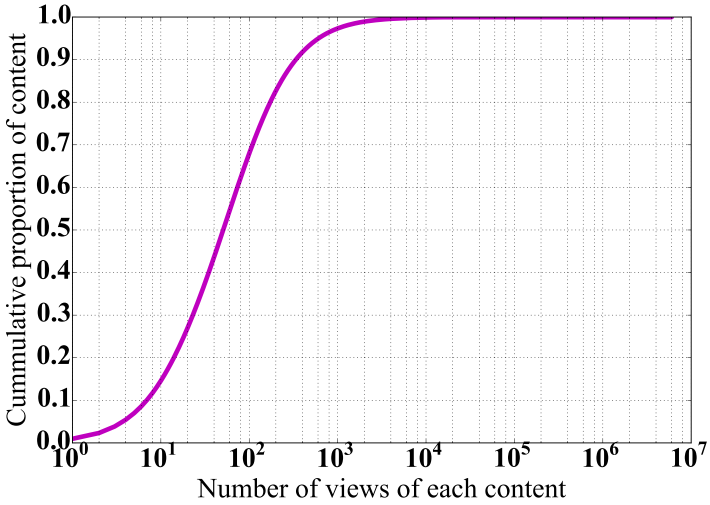


FIGURE 4.1 – Total number of content uploaded per month.

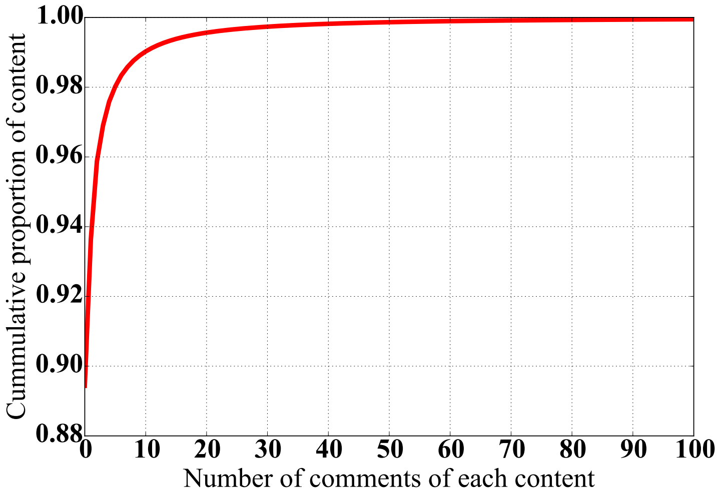
dataset and the distributions of them are shown in Figure 4.3 and Figure 4.2. As we can see, objects of the dataset annotated with at least one tag and at most 387 tags. Objects with more than 10 tags make it to the top 20%, while those with less than 2 tags take up the bottom 20% of the whole dataset. There are over 7.8M unique tags in the dataset, and some tags are frequently used while some are used only by a small group of users. The number of times that objects have been viewed differ greatly, from 1 time to 6, 027, 105 times. Each of the top 20% attracts more than 200 viewers and those of the bottom 20% have only been viewed less than 11 times. We can also observe about 90% of all the objects have received no comments or favorites, whereas the most number of comments and favorites of some object can be 15, 837 and 14, 060 respectively.

TABLE 4.1 – Summary of Dataset Used in This Thesis.

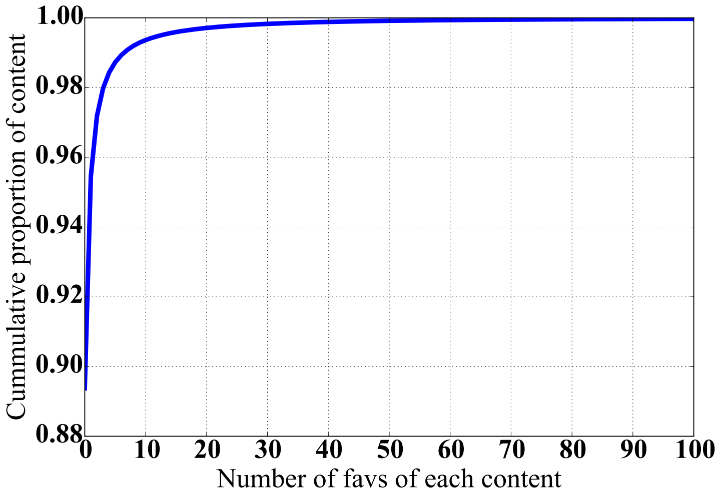
Social Media	Flickr
Total Size	68, 152, 028
Unique Number of Tags	7, 871, 020
Collect Period	Dec 1st to Dec 19th, 2014



(a) Views



(b) Comments



(c) Favorites

FIGURE 4.2 – The distributions of the number of views, comments and favorites of each object.

TABLE 4.2 – Statistics of the number of tags, views, comments and favorites of each object.

Type of Data	Tags	Views	Comments	Favorites
Minimum	1	1	0	0
Maximum	387	6,027,105	15,837	14,060
Median	5	51	0	0
Mean	7.05	186.60	0.58	0.42
Standard Deviation	7.34	1693.01	7.26	5.77

4.2. Measure of Popularity

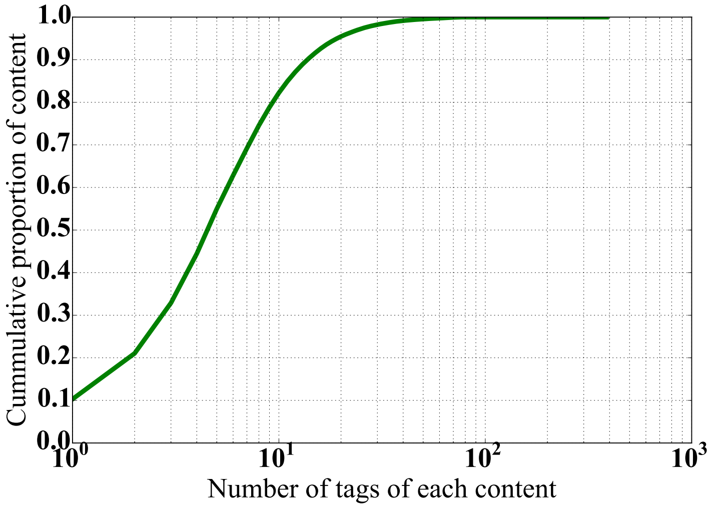
In Flickr, the number of views, comments or favorites can be measure of popularity, however, it is difficult to precisely pick any singe one as the true notion of popularity. As we know, different factors capture different levels of user engagement and provide valuable information: views imply the total requests of online contents, favorites improves content quality, and comments increase the time spent on the social service [24].

In section 4.1 we notice that the majority of the whole dataset have received no comments or favorites, while the number of views of each object range widely. So in this work we focus on the number of views as the measure of popularity. Figure 4.4 (a) shows the distribution of images/videos with different absolute number of views, and we can observe that there is large variation in the number of views of different images/videos. If we apply the absolute view counts to our algorithms, it will cause remarkable error. After applying the log function, the distribution is shown in Figure 4.4 (b). Another problem worths mentioning is that visual content tends to receive views over a period of time (10 days or longer), so the number of views usually vary greatly during this period. To eliminate the effects, we adopted a log-normalized method described in [13], defined as follows:

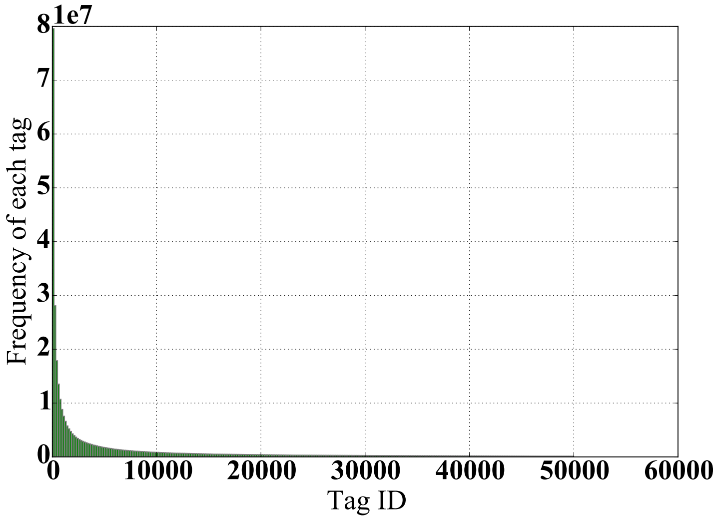
$$\text{Popularity Score} = log\left(\frac{\text{The number of views} + 1}{\text{Days since upload}}\right), \tag{4.1}$$

Figure 4.4 (c) shows the normalized results and we can see the histogram resembles a Gaussian distribution in contrast to Figure 4.4 (b) where only the log function is applied. Therefore, throughout the rest of this thesis, popularity score refers to this log-normalized number of views.

[1] Note that the total number of unique tags is 7.8M, but we truncate the graph on the left to amplify the remaining part.

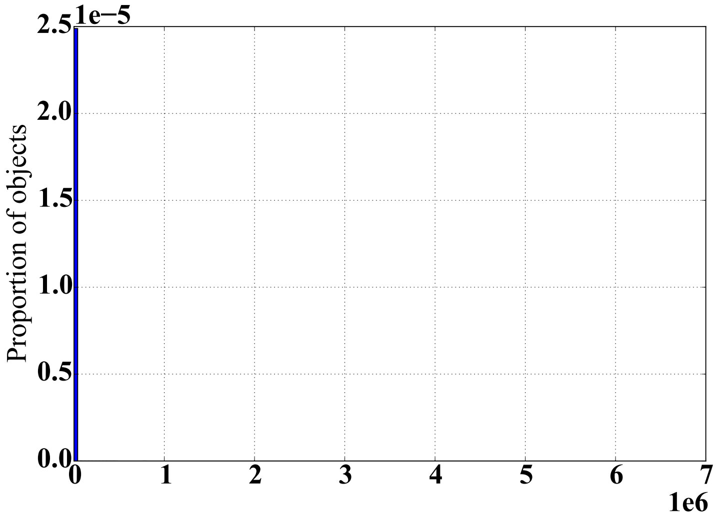


(a) The distribution of the number of tags of each content.

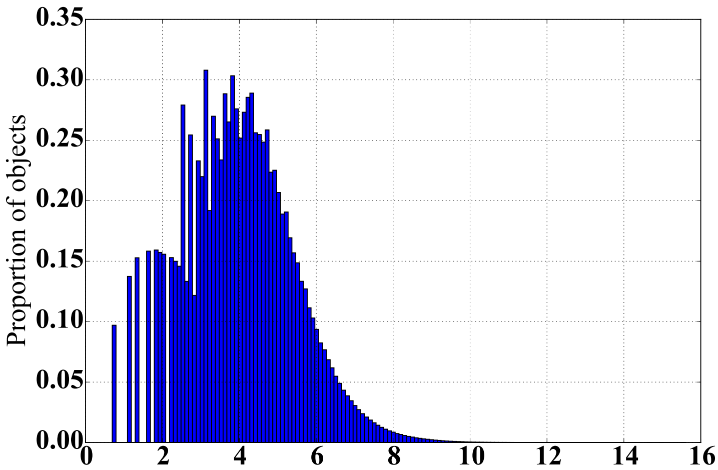


(b) Frequency of each unique tag[1].

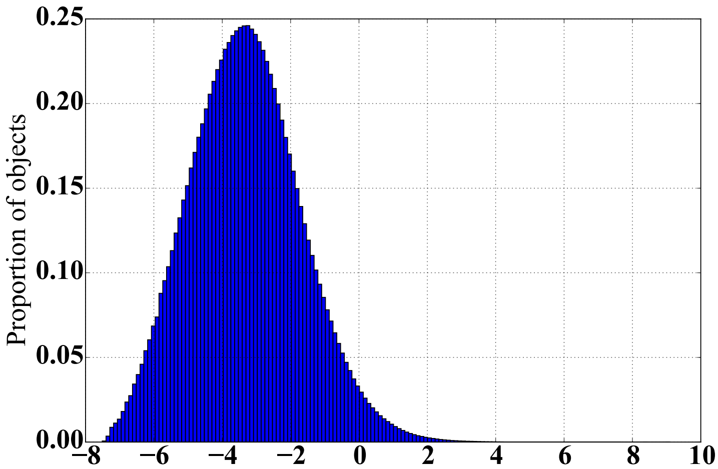
FIGURE 4.3 – Distributions of the number of tags of each content and frequency of each unique tag.



(a) Absolute number of views



(b) log(Number of views + 1)



(c) log((Number of views + 1) / days after upload)

FIGURE 4.4 – Histogram of different processing methods.

4.3. Comparison of Different Ranking Methods

This section compares different tag ranking methods and presents the ranking results by frequency, DF-W, FolkRank, and FolkPopularityRank (FPRank for short) on the dataset described previous. The top 50 tags are listed in Table 4.3, Table 4.4, Table 4.5 and Table 4.6, respectively. As we can see, tags of year number (e.g., *2009*, *2010*, etc), location names (e.g. *california*, *usa*, *london*, etc) and camera (e.g., *nikon*, *canon*, *iphoneography*)come first in all the ranking lists, though, side key differs somewhat. Tags related to instagram (e.g., *square*, *square format*, *instagram app*) are popular and quite frequently used in Flickr. The DF-W method extracts tags that are related to photography (*eye-fi*, *raw*, *hdr*), specific tags of some popular groups (*10 million photos*) or professional photographers (*www.josemariamorenogarcia.es*) in Flickr. Tags of activities and events are highly ranked by FolkRank. FPRank gives higher scores to tags that describe the visual content of photos, such as colors (*blue*, *white*, *black*), people (*girl*, *woman*, *family*), adjectives (*beautiful*, *sexy*, *cute*).

It may be confusing to find the differences among different ranking lists. So we present three applications of the tag ranking algorithms in the following chapters to compare the their efficiency in various aspects. Firstly, we analyzed the spatial and temporal change of the top-ranked tags that greatly influence content popularity in Chapter 5. Secondly, in Chapter 6 we evaluated the efficiency of various tag ranking methods in popularity prediction. Furthermore, to improve prediction accuracy, we did prediction experiments in spatial and temporal dimension, and we also tried to combine tag features and image visual features for prediction. Thirdly, we demonstrate a series of experiments on tag recommendations based on our proposed methods in Chapter 7.

Table 4.3 – Top 50 popular tags by frequency.

Rank	Tag	Frequency	Rank	Tag	Frequency
1	square	1,410,889	26	england	734,792
2	iphoneography	1,352,341	27	wedding	714,903
3	square format	1,305,508	28	italy	682,348
4	instagram app	1,297,563	29	new york	679,249
5	california	1,214,649	30	vacation	673,274
6	travel	1,180,771	31	city	673,081
7	usa	1,176,968	32	germany	668,016
8	nikon	1,175,526	33	canada	667,914
9	2010	1,099,707	34	party	649,522
10	canon	1,087,456	35	park	643,824
11	2011	1,055,605	36	water	633,839
12	2012	1,044,360	37	people	633,629
13	2009	1,020,391	38	uk	632,586
14	london	988,353	39	spain	609,055
15	2008	943,943	40	architecture	603,195
16	japan	925,192	41	festival	594,801
17	france	910,396	42	summer	593,879
18	nature	863,458	43	nyc	592,297
19	2007	847,988	44	taiwan	580,207
20	art	839,824	45	paris	576,946
21	2013	837,096	46	2006	561,608
22	music	818,397	47	san francisco	558,311
23	europe	776,893	48	australia	557,600
24	beach	750,170	49	winter	551,217
25	united states	736,527	50	snow	540,800

Table 4.4 – Top 50 influential tags ranked by DF-W.

Rank	Tag	Score	Rank	Tag	Score
1	2013	1,112,064.30	26	10 million photos	148,605.51
2	instagram app	615,832.04	27	travel	145,118.62
3	2007	595,368.31	28	food	144,399.29
4	2008	506,248.80	29	www.josemaria-morenogarcia.es	134,995.83
5	2012	494,362.75	30	square format	131,527.19
6	2009	401,961.30	31	2010	123,038.54
7	2006	399,619.24	32	2011	114,911.78
8	nikon	336,058.47	33	girl	114,469.57
9	2014	334,683.07	34	seattle	113,487.64
10	square	319,058.24	35	taiwan	113,230.29
11	united states	272,292.72	36	beach	113,044.30
12	iphoneography	262,126.06	37	winter	111,182.73
13	canon	240,141.83	38	graffiti	110,512.34
14	wedding	234,159.13	39	san francisco	108,598.73
15	night	202,275.76	40	chicago	106,813.18
16	portrait	195,103.41	41	cosplay	105,569.83
17	2005	191,549.12	42	people	105,357.00
18	eye-fi	186,142.68	43	europe	96,842.94
19	raw	180,733.98	44	car	96,161.93
20	vacation	178,436.02	45	dc	95,763.62
21	hdr	162,959.12	46	family	94,303.29
22	street	162,092.59	47	water	93,419.84
23	architecture	157,545.28	48	germany	92,464.52
24	london	156,937.87	49	2004	91,487.86
25	landscape	152,857.79	50	nature	89,250.27

TABLE 4.5 – Top 50 tags ranked by FolkRank.

Rank	Tag	Score	Rank	Tag	Score
1	california	14,403.00	26	wedding	7,879.53
2	usa	13,913.40	27	uk	7,835.73
3	nikon	13,633.60	28	water	7,740.91
4	london	13,170.30	29	city	7,658.59
5	travel	12,893.60	30	australia	7,624.51
6	canon	12,887.90	31	party	7,562.60
7	2010	12,842.50	32	architecture	7,502.41
8	2011	12,173.40	33	new york	7,496.53
9	2009	12,160.60	34	people	7,448.81
10	2012	11,756.60	35	park	7,376.45
11	japan	11,173.10	36	food	7,369.11
12	2008	11,133.00	37	vacation	7,276.36
13	france	11,101.20	38	nyc	7,260.62
14	art	10,580.00	39	spain	7,092.31
15	nature	10,466.20	40	paris	7,086.83
16	music	10,049.30	41	square	7,043.05
17	2007	9,944.71	42	summer	6,910.93
18	2013	9,588.51	43	festival	6,821.66
19	england	9,587.61	44	winter	6,761.35
20	europe	8,840.33	45	snow	6,749.83
21	beach	8,690.50	46	san francisco	6,699.32
22	united states	8,401.43	47	2006	6,607.13
23	germany	8,256.83	48	sky	6,580.21
24	canada	8,071.94	49	concert	6,572.09
25	italy	8,050.08	50	night	6,475.78

Table 4.6 – Top 50 influential tags ranked by FPRank

Rank	Tag	Score	Rank	Tag	Score
1	2009	5,723.99	26	vacation	2,712.00
2	2010	5,573.00	27	photography	2,707.91
3	girl	5,250.56	28	taiwan	2,650.44
4	2011	5,067.97	29	explore	2,620.91
5	2008	4,832.77	30	usa	2,604.00
6	portrait	4,772.44	31	italy	2,492.94
7	2012	4,572.33	32	united states	2,440.68
8	2007	4,569.04	33	black	2,355.19
9	wedding	4,069.12	34	party	2,277.77
10	japan	3,946.44	35	australia	2,232.59
11	woman	3,903.36	36	square	2,228.75
12	hdr	3,569.84	37	2013	2,195.75
13	london	3,492.55	38	photo	2,189.28
14	blue	3,416.97	39	san francisco	2,187.54
15	california	3,341.80	40	sexy	2,175.97
16	instagram app	3,140.45	41	paris	2,145.27
17	travel	3,114.32	42	canon	2,143.70
18	square format	3,072.03	43	model	2,118.67
19	beautiful	3,071.46	44	china	2,099.68
20	2006	2,961.53	45	germany	2,039.24
21	white	2,855.51	46	seattle	2,003.47
22	france	2,788.27	47	new york	1,989.96
23	iphoneography	2,760.43	48	family	1,979.72
24	light	2,750.38	49	live	1,978.31
25	music	2,722.59	50	cute	1,932.24

Chapter 5

Analysis of Influential Tags

Social media rarely stays the same for long. The social trends always change rapidly and so as to user tagging. Social sharing sites like Flickr allow users to annotate any words to their visual content, and in consequence of this, the vocabulary of tags used by people usually differ greatly from time to time and place to place. By analyzing the influential tags in the social tagging over location and time, we can study online user behavior and explore the evolution of community focus, which will help us to increase the effectiveness of popularity prediction and tag recommendation as well. In this section, we investigate the change of influential tags on social media over location and time and analyze the spatial difference and temporal evolution of tags.

5.1. Spatial Analysis

First, we selected all the geo-tagged photos or videos and the total number is $39,915,621$. We then divided them into nine clusters by using mean-shift clustering [46] on the geographic locations. The nine clusters are approximately in these regions: Europe, North America, Asia, South America, Australia, Africa, North Pacific, South Pacific, and Indian Ocean, as shown in Figure 5.1 and Table 5.1 shows a summary of each region. We can see that over 80% are concentrated on Europe and North America. Each image/video of each area is annotated with $4-7$ tags on average. These geo-tagged content are relatively less popular than the whole dataset as described in section 4.1.

We applied different tag rankings to each region. Table 5.2 shows the top 10 most frequently used tags of each area, and we can observe that most of them are names of countries and cities (e.g. *france* and *london*). Such tags are so common that neither are they closely related to social popularity nor do they contribute to increasing the social popularity score.

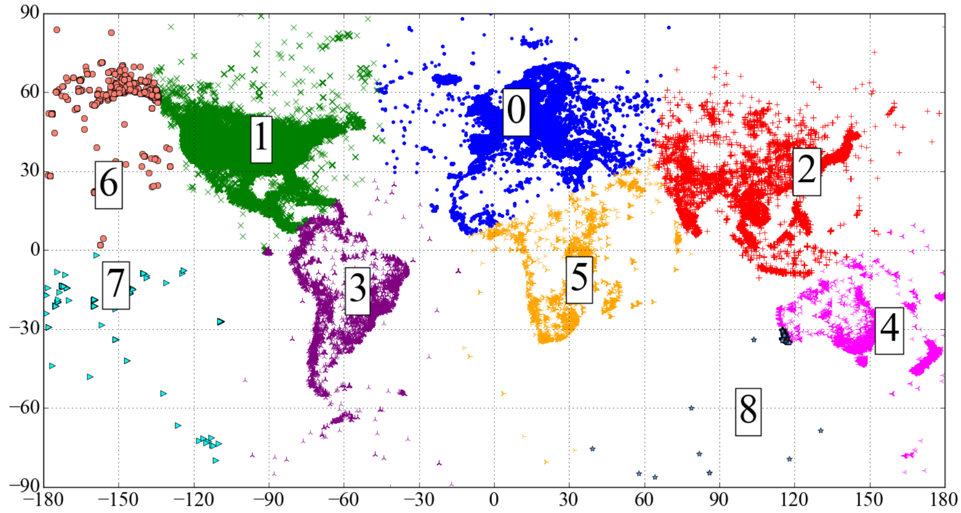


FIGURE 5.1 – Nine region clusters.

TABLE 5.1 – Summary of each region clusters.

Area	Approx. region	Total number of data	Unique tags	Average number of tags	Average number of views
0	Europe & North Africa	16,594,614	2,361,600	6.53	0.59
1	North America	15,879,777	1,860,800	5.72	0.43
2	Asia	4,095,013	554,794	5.12	0.41
3	South America	1,327,732	250,753	7.12	0.65
4	Australia	1,219,873	245,637	5.32	0.46
5	Africa	490,565	111,804	5.37	0.61
6	North Pacific	282,414	6,5721	5.29	0.42
7	South Pacific	25,043	7,159	5.29	0.95
8	Indian Ocean	590	1,433	4.31	3.10

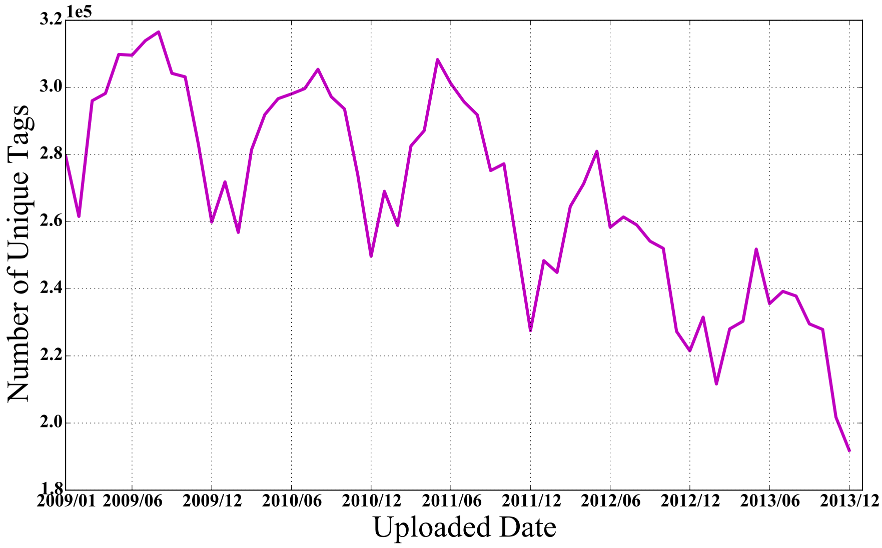


FIGURE 5.2 – Total number of unique tags of every month.

In Table 5.3, the top 10 most influential tags of each area, ranked by the DF-W method, are summarized and some regional difference can be observed from this table. For instance, photos related to art are popular in Europe area (*abode of chaos*, *dadaisme*). In North America and Asia area, people care more about famous photographer (*mlhradio*, *061028choshi*), photography technology (*hdr*, *square format*) and camera types (*nikon*, *range finder*). Photos related to travel are popular in Asia (*world travels*, *beat*) and South America area (*landscape*, *nature*, *explore*, etc.)

5.2. Temporal Analysis

To understand the change of influential tags over time, we investigate the change of influential tags over time. As we can see from Figure 4.1, the amount of content uploaded to Flickr is relatively stable from 2009 to 2013, which contains $43,692,766$ images/videos, so we select this part of data to perform experiments. The total number of unique tags of every month is shown in Figure 5.2. The number changes periodically, with an obvious peak around June and a little one in December of every year, which results from increase of activities in summer and winter holidays.

The most frequently used tags of each year are listed in Table 5.4 (a). We can observe that the top 10 tags of each year do not change a lot. Tags of the year number are most frequently used in each year and tags such as *canon*, *usa*, *california* are always among

Table 5.2 – Top 10 most frequently used tags of each area.

	Europe	North America	Asia	South America	Australia
1	france	california	japan	brazil	australia
2	london	usa	taiwan	brasil	new zealand
3	europe	united states	china	argentina	sydney
4	england	canada	tokyo	chile	nsw
5	germany	new york	travel	peru	melbourne
6	italy	square	india	de	victoria
7	uk	iphoneo-graphy	asia	south america	queensland
8	spain	san francisco	thailand	buenos aires	western australia
9	paris	square format	台灣 (Taiwan)	square	new south wales
10	nikon	instagram app	square	iphone-ography	beach

	Africa	North Pacific	South Pacific	Indian Ocean
1	africa	hawaii	french polynesia	antarctica
2	south africa	alaska	tahiti	bandai
3	tanzania	oahu	easter island	gunpla
4	kenya	honolulu	travel	ice
5	safari	maui	island	perfect grade
6	travel	vacation	chile	strike freedom
7	cape town	kauai	south pacific	gundam seed
8	dubai	usa	honeymoon	fragments
9	namibia	big island	moorea	snow
10	uae	waikiki	isla de pascua	gundam

Table 5.3 – Top 10 tags that ranked by the DF-W method of each region.

	Europe	North America	Asia	South America	Australia
1	thierry ehrmann	mlhradio	free	naturaleza	field hockey
2	abode of chaos	90095	attractive	ландщафт (landscape)	australia
3	raw art	square format	world travels	τοπίο (scenery)	new
4	salamander spirit	nikon	iso 200	παραλία (beach)	oceania
5	dadaisme	creative commons	beautiful	private	nsw
6	paul virilio	phenomenal	range finder	planète (planet)	nz
7	sculpture moderne	bild	canon	de	queensland
8	groupe serveur	2013	canon a35 datelux	america meridionale	beach
9	stockphoto	hdr	flickr meetup	nature	viaje
10	picture	wet lifestyles	061028choshi	explore	panties

	Africa	North Pacific	South Pacific	Indian Ocean
1	united nations	sexy	kap	ice
2	campaign	women	kite aerial photography	antarctica
3	nasa ames research center	boobs	autokap	fragments
4	naciones unidas	gorgeous	polynésie	nasa
5	nasa	royal caribbean	2013	snow
6	square	booty	tahiti	sanae
7	viewing	midway island	pierre lesage	south african national antarctic programmearc-ticantarctic
8	conflict	hawaii	intercontinen-tal	perfect grade
9	ecosystem	serenade of the seas	mururoa	strike freedom
10	plains	disney	blue	dumpr

the most popular tags. And tags corresponding to Instagram (*square*, *square format*, *instagram app*) and iPhone (*iphoneography*) become most popular after 2011.

While Table 5.4 (b) shows the top 10 ranking tags by applying the DF-W method from 2009 to 2013. The most influential tags change greatly over time, in the aspects of arts (*abode of chaos*, *borderline biennale*, *post-apocalyptique*), artists (*hierry ehrmann*, *randomok*, *jmmg*), politics and philosophy (*zionist*, *fatah*, *salamander spirit*), photography technology (*hi-res*, *zoom lens*, *square format*), and famous places (*90095*, *southeast asia*, *atlantic avenue*). And we noticed that the tags of the time of the year (e.g. *2013*) always appeared in the list as well. We see from the table how the community focus and popularity tendency changes over time.

Then we select six typical influential tags and investigate how they change over years. The ranking changes of these tags are plotted in Figure 5.3 and Figure 5.4. For tags such as *instagram app* and *square format*, they started to be popular from 2010 and quickly became one of the most important tags, along with appearance of the online mobile social media Instagram. On the other hand, tags such as *high resolution*, were in one of the most important tags but gradually lost its position after 2011, which result from the fact that high resolution cameras become common and it is not the key point of good photos any more. Strong temporal changes are seen for seasonal tags such as *winter* and *snow*. Peaks are observed around June and December because of the opposite seasons in the northern and southern hemispheres. In addition, tags related to season cycle also show temporal changes. For example, the tag *explore* varied periodically, roughly with two peaks in a year, which corresponds to, as pointed above, the increase of outdoor activities during summer and winter holidays.

5.3. Analysis on Various Social Services

Additionally, there are a variety of social services sharing different types of content. Each social service owns its characteristics. So we perform tag ranking on other social service as well, i.e., Vine[1], Instagram[2], niconico[3], Photohito[4], as summarized in Table 5.5. We use open datasets Instagram[47] and niconico[48], while the datasets of Vine and Photohito are crawled from their sites.

[1]https://vine.co/
[2]https://instagram.com/
[3]http://www.nicovideo.jp/
[4]http://photohito.com/

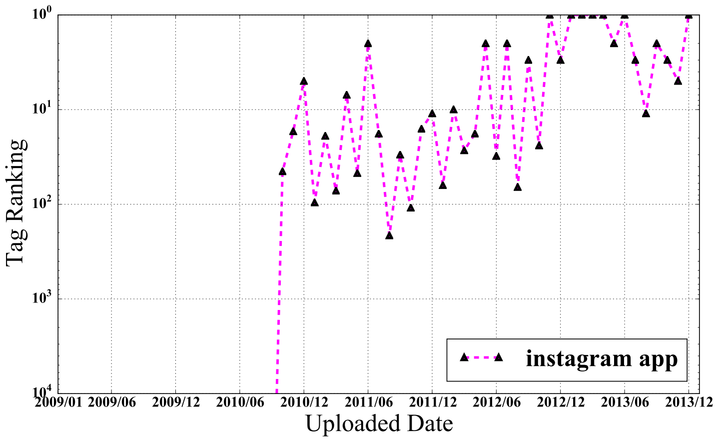
Table 5.4 – Top 10 tags of the year from 2009 to 2013.

(a) Ranking by frequency.

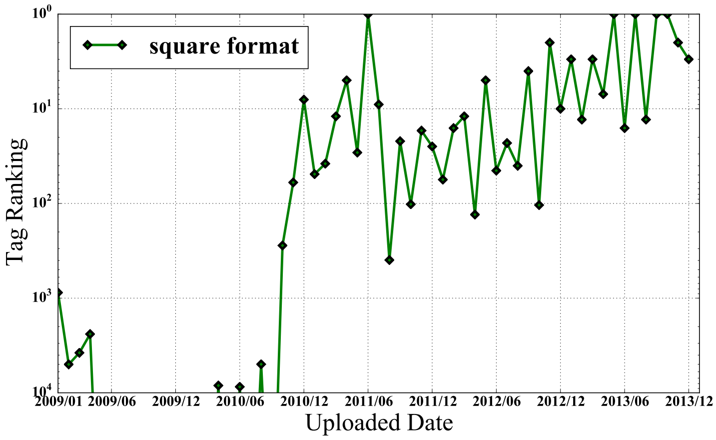
	2009	2010	2011	2012	2013
1	2009	2010	2011	2012	2013
2	nikon	nikon	square	square	square
3	california	canon	iphoneography	iphoneography	iphoneography
4	canon	california	square format	square format	square format
5	travel	usa	instagram app	instagram app	instagram app
6	usa	travel	nikon	usa	united states
7	london	london	usa	canon	usa
8	art	music	canon	nikon	travel
9	music	art	california	travel	california
10	france	nature	travel	london	nikon

(b) Ranking by DF-W.

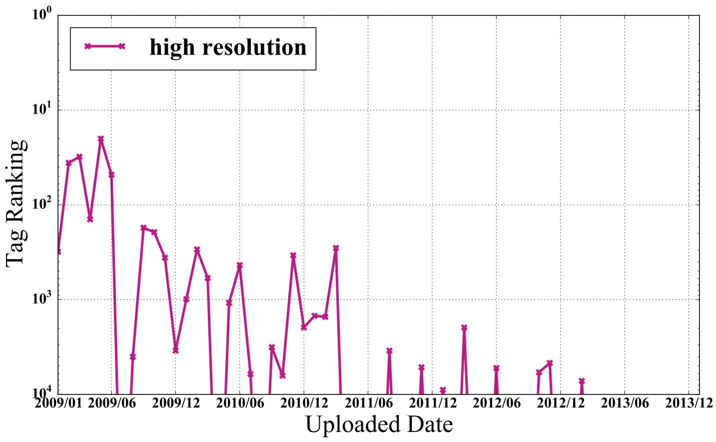
	2009	2010	2011	2012	2013
1	raw art	thierry ehrmann	borderline biennale	jose maria moreno garcia	iphoneography
2	zionist	abode of chaos	randomok	jmmg	southeast asia
3	modern sculpture	portland state university	salamander spirit	www.flickriver. com/josemaria morenogarcia	square format
4	2009	90095	emergence	picas- aweb.google. com/josemaria morenogarcia	asia
5	maison d'artiste	post- apocalyptique	2011	2012	instagram app
6	catalano	hi-res	l'esprit de la salamandre	instagram app	malaysia
7	tulkarem	rodents and rabbits	retro	madridejos. fotos.es	atlantic avenue
8	alchimie	2010	square	www.vimeo.com /madridejos	southeast
9	fatah	raw art	zoom lens	bi	visit
10	hijra	demeure du chaos	vintage	free	house



(a) instagram app

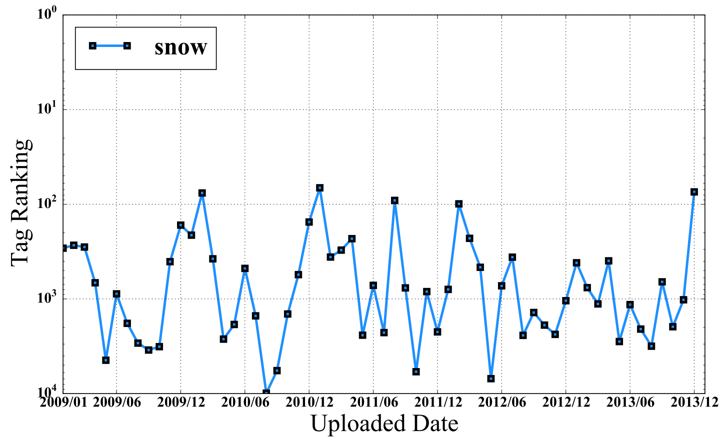


(b) square format

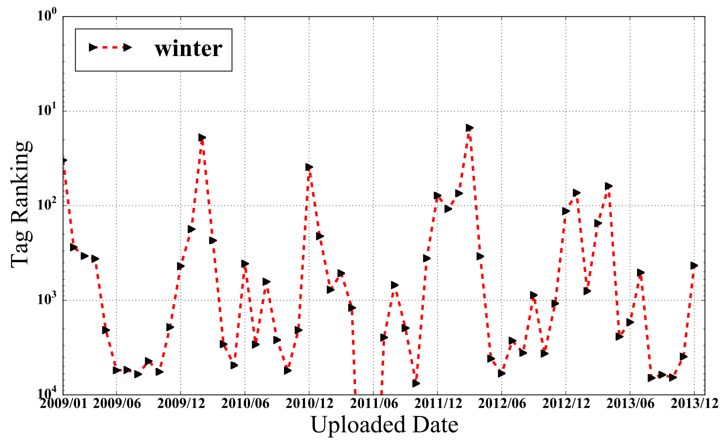


(c) high resolution

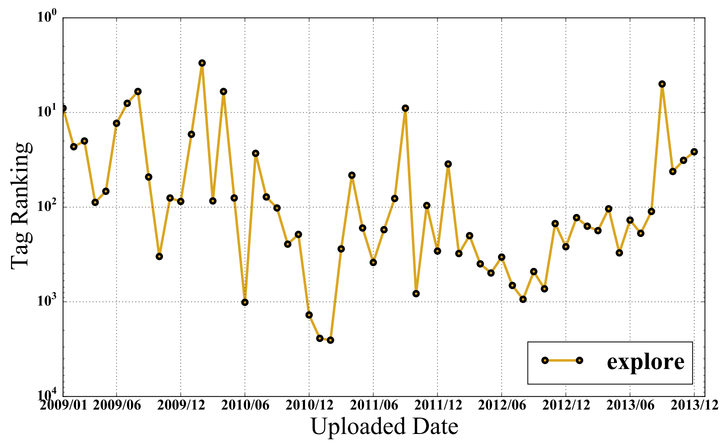
FIGURE 5.3 – The ranking change of typical influential tags: part I.



(d) snow



(e) winter



(f) explore

FIGURE 5.4 – The ranking change of typical influential tags: part II.

The top 10 influential tags by DF-W are listed in Table 5.6. Since Instagram do not provide the number of views, so we use the number of comments as measure of popularity. We can find that the social services differ each other. In Vine, tags related to interesting videos (*wshh*, *remake*, *KingBach*) are popular, while in Instagram tags of high quality photos (*vscocam*, *instagood*) and communities (*iphonesia*, *photooftheday*) are more influential. Most niconico users are a group of people called Otaku, so tags corresponding to games, music, or animation (ゲーム, 音楽, アニメ) have quite high influence on popularity. In Photohito, which is a photo sharing site in Japan, tags associated with nature (風景・自然, 生き物, 花) attract more attention. In contrast, photography technology related tags are more influential. By analyzing the top-ranked tags of various social services, we find that users of different interests and focus are gathered on different services, and influential tags vary greatly from service to service.

Table 5.5 – Datasets of different sharing services.

	Vine	Instagram	niconico	Photohito	Flickr
Data Type	Videos	Photos	Videos	Photos	Photos & Videos
Number of Unique Tags	75,747	271,491	5,328,340	257,783	7,871,020
Total Content with tags	93,629	1,073,349	8,305,696	1,969,189	68,152,028

Table 5.6 – Top 10 most influential tags in different services by DF-W (to the number of comments).

	Vine	Instagram	niconico	Photohito	Flickr
1	wshh	vscocam	もっと評価されるべき	風景・自然	instagram app
2	remake	instagood	ゲーム	生き物	square
3	KingBach	iphonesia	東方	canon	nikon
4	cute	photooftheday	音楽	カワセミ	explore
5	onedirection	hongkong	文字を読む動画	モノクロ	canon
6	voiceover	paris	謎の感動	北海道	archer10
7	worldstar	aditzt	アニメ	女性	hdr
8	Ranked	iphoneography	腹筋崩壊	花	my_gear_and _me_premium
9	videoshop	Instagram	VOCALOID 伝説入り	富士山	my_gear_ and_me
10	EXOplanet-inBKK	nyc	Fullver. へのリンク	映り込み	nature

Chapter 6

Tag Based Popularity Prediction

The dramatically expand of online content shared on social media has intensified the competition of users' attention, resulting that only a small amount of social content become popular while the vast majority is bound to a very limited attention. Predicting the social popularity is of great importance in many areas such as network dimensioning, online advertising, or content recommendation. Besides, it fosters the possibility of appropriate decision making to modify an item and the manner of its publication. As mentioned in Chapter 1, tagging is good predictors for future social popularity and computationally effient. Since we have extracted tags that have a great influence on popularity, in this Chapter we apply them to prediction task. First we describe the framework of popularity prediction using only tags in details, and then compare prediction results using various tag ranking algorithms. In addition, to make further improvement on accuracy, attempts are made in prediction experiments.

6.1. Framework of Popularity Prediction System

Social media allows users to use open-ended words that match their language, preference and interests for tagging. The tags provided by users differ greatly from person to person, which leads to an uncontrolled vocabulary. Apart from that, the quality of tags varies widely, from tags that well describe a web content, to those that are incomplete and ambiguous. As Figure 4.3 (b) in Chapter 4 shows, tags in the long tail are usually too specific or result from spelling mistakes. The fact of social tagging severely limits the application of tags and it is unwise to take all the tags into consideration. Thus in the prediction system, we first do a tag ranking to select tags that greatly influence social popularity, and then a Support Vector Machine (SVM) is applied for learning and predicting.

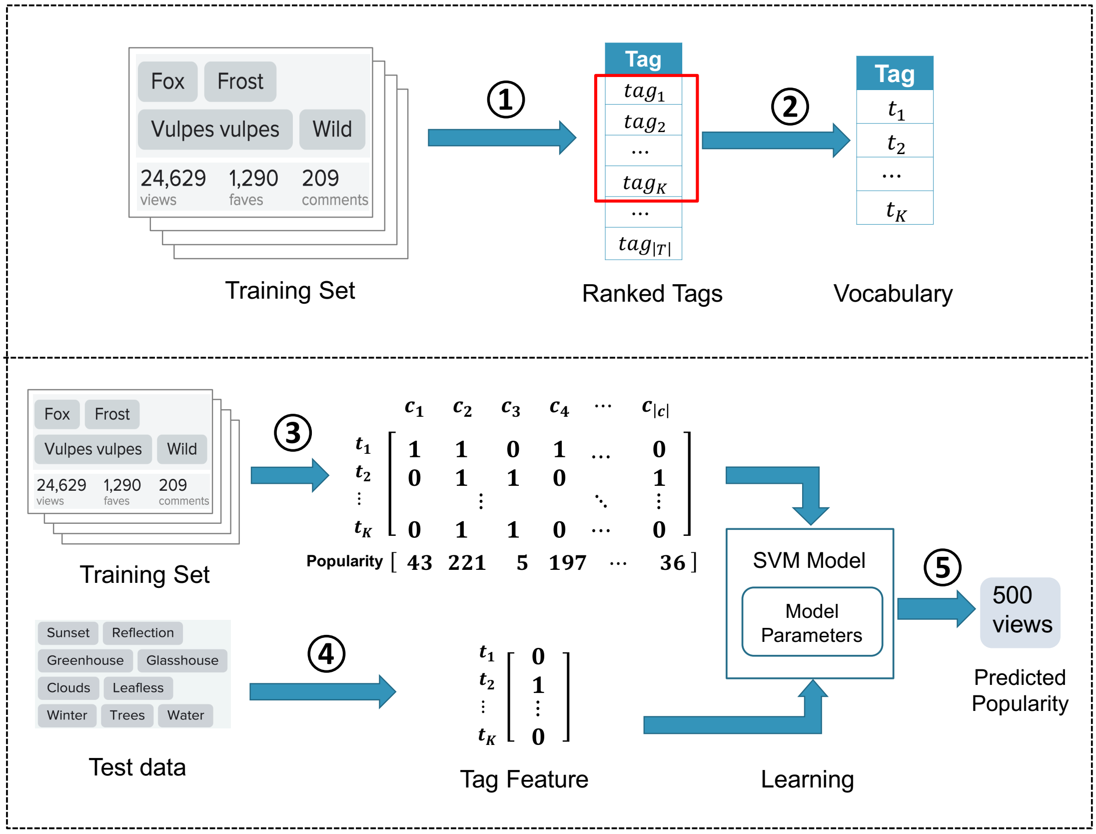


FIGURE 6.1 – The framework of the popularity prediction system.

The main framework of our tag based popularity prediction system is illustrated in Figure 6.4. The details of each step is as follows.

1. Collect all the unique tags, $T = \{t_1, t_2, \ldots, t_{|T|}\}$ in the training set, and perform tag ranking to extract tags with high influence on social popularity.

2. A vocabulary, $T^{'} = \{t_1^{'}, t_2^{'}, \ldots, t_K^{'}\}$ is constructed by the top-ranked K tags.

3. For each content c_i in the training set, a vector of K dimension is obtained as a tag feature. Thus a $\mid C \mid \times K$ matrix is obtained and each entry refers to count of the corresponding tag. Users do not usually use a tag twice to annotate the same content, so here 1 indicates the content is annotated with the corresponding tag, otherwise 0 indicates that the tag is not used. Then an SVM is applied to obtain a trained model.

4. Build a feature vector for test data in the same way.

5. Apply the test feature vector to the trained model and predict the popularity for the test data .

6.2. Comparison of Various Tag Rankings

To evaluate the tag based popularity prediction system, four different tag ranking algorithms are applied. Three rank correlation coefficients, Spearman's ρ, Pearson's γ and Kendall's τ of predicted popularity $P = \{p_1, p_2, \ldots, p_{|C|}\}$ and real-world popularity $Q = \{q_1, q_2, \ldots, q_{|C|}\}$ are used as measure of accuracy. The three rank correlation coefficients are defined as follows:

$$\text{Spearman's } \rho = 1 - \frac{6 \sum d_i^2}{\mid C \mid (\mid C \mid^2 - 1)} \tag{6.1}$$

where d_i is the difference in paired ranks.

$$\text{Pearson's } \gamma = \frac{\sum (p_i - \overline{p})(q_i - \overline{q})}{\sqrt{\sum (p_i - \overline{p})^2} \sqrt{\sum (q_i - \overline{q})^2}} \tag{6.2}$$

$$\text{Kendall's } \tau = \frac{(\text{number of concordant pairs} - \text{number of discordant pairs})}{\mid C \mid (\mid C \mid - 1)^2} \tag{6.3}$$

The experiment results of applying ranking of frequency, DF-W, FolkRank, FPRank are shown in Figure 6.2. The dimensions of the feature vectors, i.e. the size of vocabulary K, is changed from 100 to $35,000$. We can see that the DF-W method achieve highest accuracy in predicting social popularity. FPRank performs a little better than frequency, while FolkRank performs worst among all the ranking methods. We can also observe that the prediction accuracy of $K = 10000$ is almost the same, even a litter better than that of $K = 35000$, which explains the use of tags in long tail will decrease the accuracy of popularity prediction.

6.3. Prediction in Spatial and Temporal Dimension

Furthermore, to improve the accuracy, we try to perform popularity prediction in both spatial and temporal dimension. Here we rank all the tags by DF-W since it works better than other ranking methods. We use the same dataset described in Chapter 5. Figure 6.3 (a) shows the accuracy of social popularity prediction in different areas, using top 1000 influential tags, measured by Spearman's rank correlation. We can observe that the prediction accuracy of divided areas datasets are generally higher than that of

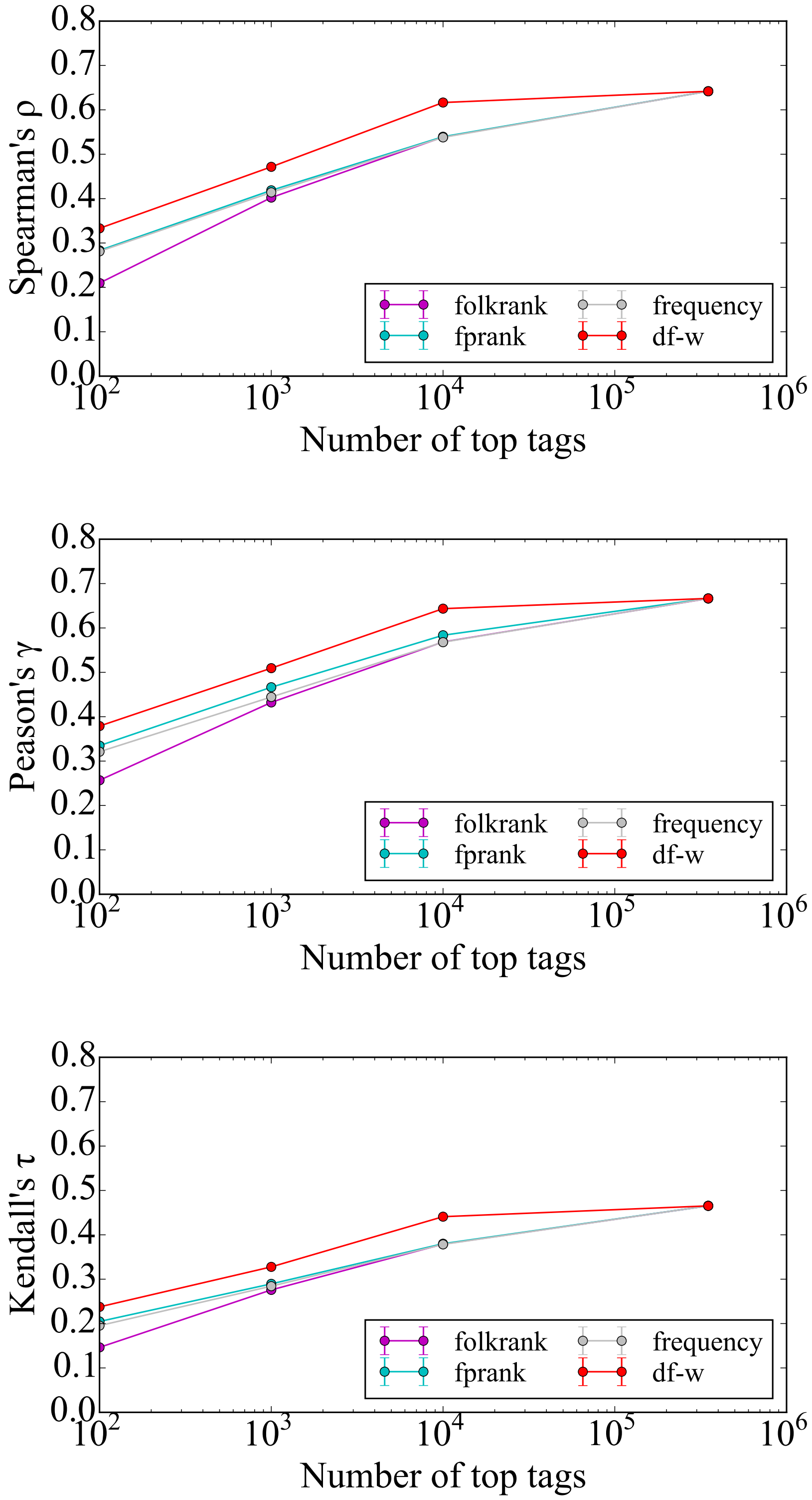


FIGURE 6.2 – Accuracy of popularity prediction using different ranking method with changing size of tag vocabulary.

global dataset. It shows that it is more effective to predict social popularity according to different areas.

Similarly, Figure 6.3 (b) shows the accuracy of social popularity prediction in the temporal dimension, using top 1000 influential tags, measured by the Spearman's rank correlation. Although the accuracy changed greatly over time, we can still observe that in most cases, the prediction accuracy of each month is better than that of the whole period from 2009 to 2013. Thus we can conclude that it will increase the prediction accuracy to analyze the popularity tendency in different time periods.

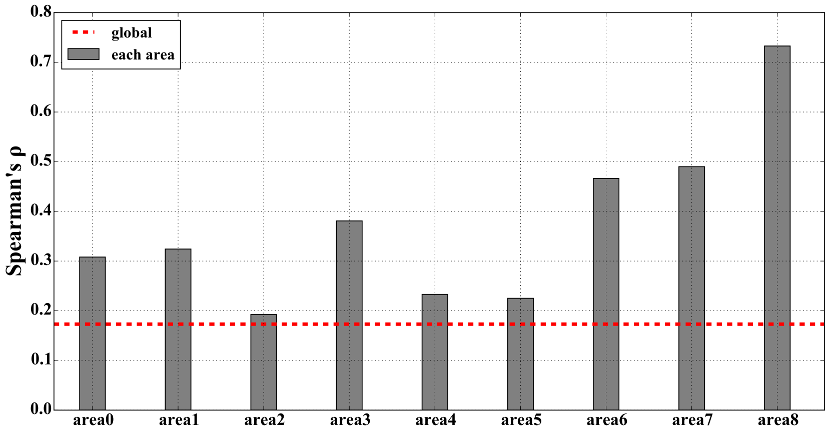
6.4. Multimodal Learning for Popularity Prediction

In the previous sections, we focus on popularity prediction using only the text tags and it achieves good results. However, different modalities typically carry different kinds of information. In Flickr, without considering the image visual content, different images with the same set of tags will be predicted to gain the same prediction result, and what is worse, users may overuse popular tags which are not consistent with the uploaded images just for attracting attention. In real cases, large percentage of uploaded images with few tag annotations present, and our previous method will not have good performance.

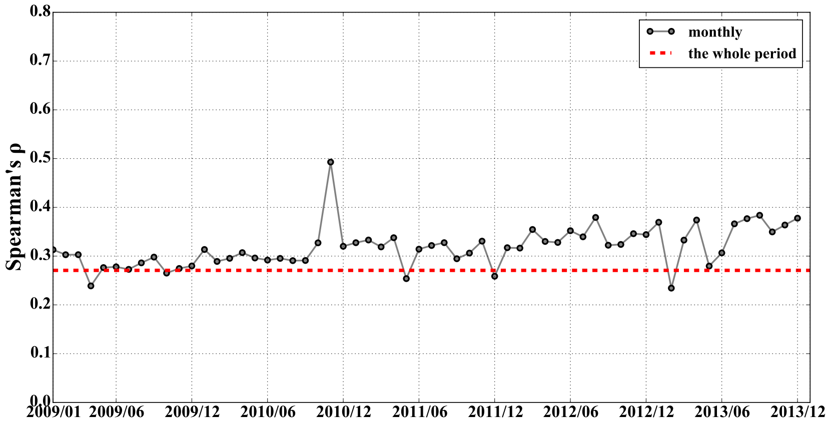
To improve the prediction accuracy, in this section we try to use both tag feature and visual feature in image popularity prediction task. In related work, Khosla et al. [13] and Pedro et al. [15] combine various visual contents and social cues and utilize an SVM to make predictions. In a multimodal settings, each input modality has a different kind of representation and correlational structure. For tag features, sparse word count vectors are used, while for image visual features, real-valued dense vectors are used. Simple concatenation of these features may be inappropriate and cannot get good prediction results. Thus, in this section, we try to use multiple kernel learning framework in our popularity prediction task and compare the experiment results with other unimodal and multimodal learning results.

We perform our experiments on the YFCC100M dataset mentioned in Chapter 4 as well. We choose $10,000$ images from the whole dataset, and tag feature was represented using a vocabulary of the $2,000$ most popular tags, which obtained by the DF-W algorithm.

Deep learning algorithms such as convolutional neural networks (CNNs) have recently become popular as methods for learning image representations. And deep feature shows



(a) in spatial dimension



(b) in temporal dimension

Figure 6.3 – Accuracy of popularity prediction using the top 1000 top ranking tags by applying the DF-W method.

Table 6.1 – Image popularity prediction using visual features extracted by different Caffe pre-trained model.

Caffe Pre-trained Model	LIBLINEAR	LIBSVM
CaffeNet[49]	0.258	0.254
Finetune_Flickr_Style[49]	0.236	0.238
AlexNet[54]	0.251	0.257
RCNN_ILSVRC13[55]	0.254	0.253
VGG_CNN_S[56]	0.270	0.260
VGG_ILSVRC_19[57]	0.262	0.264

good performance in popularity prediction in [13]. In this paper, the Caffe [49] deep learning framework is used to extract visual features for each image, resulting in a $4,096d$ representation of the 7^{th} rectified fully connected layer. In the Caffe Model Zoo[1], it provides many pre-trained deep learning models, so we choose several representative models for visual feature extraction.

In order to predict image popularity, we compare several prediction models: a L2 regularized L2 loss SVR implemented from LibLinear package [50], a L2 regularized L2 loss SVR with RBF kernel implemented from LibSVM package[51], and the Multiple Kernel Learning (MKL) model [52] implemented in Shogun package[53]. And we also compare the SVR models over different combinations of features: visual feature only, tag feature only, simple concatenation of visual feature and tag feature. Multidimensional features were scaled to in the $[-1, 1]$ range, and the parameter of C and γ were fine-tuned in the experiments. 5-fold cross validation is used in our experiments, and the correlation between predicted popularity and the ground truth scores is computed using Spearman's rank correlation.

In Table 6.1, we use visual features extracted by different pre-trained model and compare the prediction results. As reported in the table, different extracted visual features do not show much differences in popularity prediction. It is worthwhile to note the fact that all these deep learning models are trained for visual recognition not for popularity prediction, which may explain why the prediction results seems not so good. Among all the listing model, the VGG_CNN_S model performs a little better than others, so we use it in our following experiments.

[1]https://github.com/BVLC/caffe/wiki/Model-Zoo

TABLE 6.2 – Image popularity prediction using different models and different features.

Prediction Model	Spearman's ρ
Visual-LibLinear	0.260
Visual-LibSVM	0.263
Tag-LibLinear	0.527
Tag-LibSVM	0.619
Concatenated-LibLinear	0.411
Concatenated-LibSVM	0.488
MKL	0.463

Then we evaluate all the unimodal and multimodal prediction models, as shown in Table 6.2. As we can see, tag feature perform significantly better than visual feature, with results more than twice as large as those of visual features. The multimodal models works better than visual feature, but to our surprise, worse than tag feature. And the MKL model perform better than the Concatenated-LibLinear model and a little worse than the Concatenated-LibSVM model. Among all the unimodal and multimodal, the Tag-LibSVM model has the best rank correlation, which means that tag features outperforms all the other unimodal and multimodal learning models.

6.5. A Demo of Popularity Prediction System

We develop a social popularity prediction system, aiming at helping users in the tagging process. Figure 6.4 shows a demo of our system. In the system, we first gather enough text tags to build our model and weight vectors correlated to views, comments and favorites are calculated. Users can search images from Flickr by the indicating the photo ID, along with some metadata, such as tags, the number of views, comments and favorites. The predicted social popularity of each tag will be showed in the web page. Users can modify tags based on the predicted social popularity, to gain more attention.

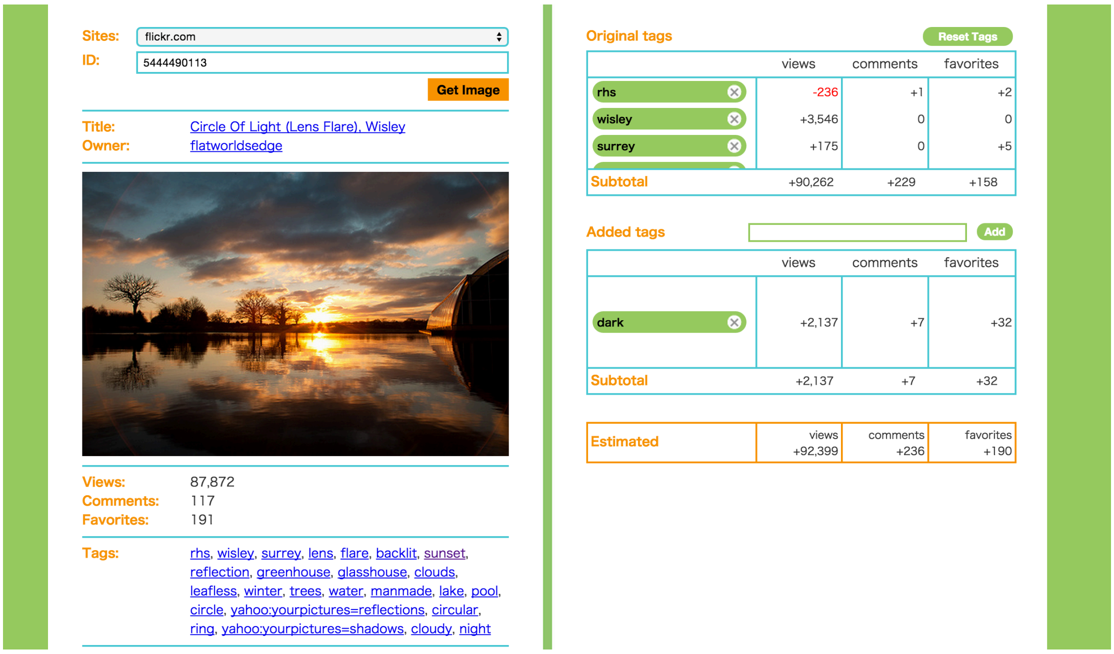


FIGURE 6.4 – A demo of the social popularity prediction system.

Chapter 7

Tag Recommendations for Popularity Boosting

Researches have shown that well tagged images are more likely to become popular, because it makes them more accessible to other users by adding as many tags as possible. However, it is not an easy task for users to annotate their content with tags that are capable of attracting popularity. Thus in this chapter, we propose two recommendation approaches based on DF-W and FPRank, which recommend tags that have high influence over popularity. We then evaluate our methods and several existing tag recommendation strategies on a dataset of Flickr and make a comparison in both efficiency of popularity boosting and tag quality.

7.1. Tag Recommendation

We first describe the procedure of tag recommendation. The experiments were also performed on the YFCC100M dataset. We select the top 0.06 % images that have more than 20 tags and over 5000 views of the whole YFCC100m dataset for training and randomly select images annotated with over 20 tags but different number of views for testing. Table 7.1 gives an overview on the training and testing sets. We can observe that the training set contains 250K unique tags and the images in the training set are really popular in Flickr, which can be good resources for recommendation. The average number of views of test set A is markably large than set B, though the number of unique tags and average number of tags of two test tests do not differ much. Then we perform tag recommendations on these datasets.

We will evaluate a tag recommender based on the proposed tag ranking methods, DF-W and FPRank, and two the well known recommendation methods, Collaborative

TABLE 7.1 – Overview of training and test sets.

Data Set	Number of Unique Tags	Average Number of Views	Average Number of Tags	Total Number of Images
Training	254,734	13,139.5	37.1	60,000
Test A	15,983	15,046.7	37.9	1,000
Test B	17,387	221.4	30.2	1,000

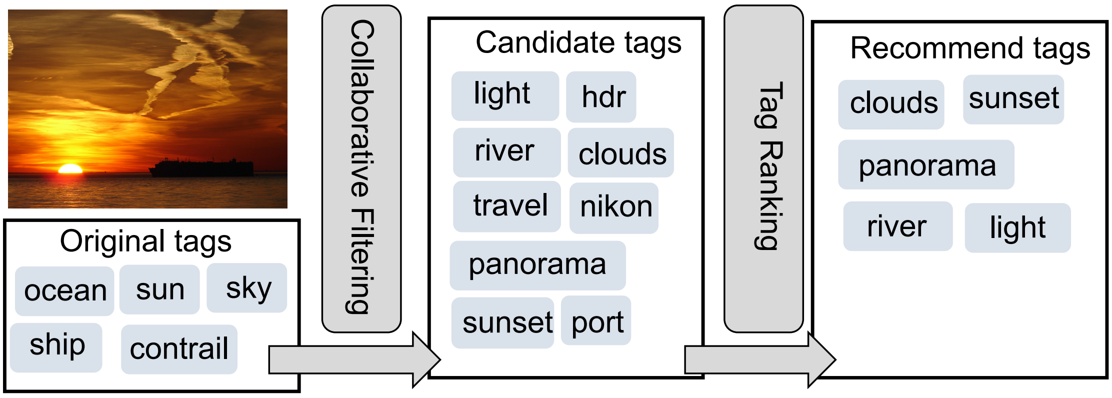


FIGURE 7.1 – Flow of Collaborative Filtering

Filtering (CF) [38, 39] and Tagcoor [17] , as mentioned in Chapter 2. We briefly describe the four methods here.

The idea of Collaborative Filtering for tag recommendation is to suggest new tags based on annotations of similar images, as shown in Figure 7.1. Given an image with user-defined tags, a feature vector is represented by the bag of tags, which is defined as follows:

$$FV_i^{tag} = \{f_{i1}, f_{i2}, \cdots, f_{ij}, \cdots, f_{iT}\}, \tag{7.1}$$

where FV_i^{tag} is the feature vector for the ith image tag sets and f_{ij} represents whether ith image has the jth tag. These vector can be viewed as an approximate representation of the image content and user preference in the corresponding domain. An ordered list of candidate tags is derived based on similarity between tag sets, and the similarity is measured by computing the cosine of the angle formed by the two feature vectors. Then the candidate tags are ranked and the top n tags in the ranking list will be recommended to users. For basic CF method, we rank the candidate tags by frequency to ultimately produce the ranked list of recommended tags. And the DF-W method can be applied to the ranking phase, which is represented as CF_DF-W method.

Tagcoor makes recommendation based on tag co-occurences, defined as follows:

$$P(t_j \mid t_i) = \frac{\mid t_i \cap t_j \mid}{\mid t_i \mid} \tag{7.2}$$

Tag aggregation and promotion strategies are then used to produce the final list of recommended tags. FPRank can directly used for tag recommendation as described in Chapter 3 so we do not repeat the details here.

Then we recommend 10 tags for each image of the two test sets using the four recommendation methods, CF, Tagcoor, CF_DF-W and FPRank respectively. Table 7.2, 7.3, 7.4, 7.5 and 7.6 give five examples of recommendation results by different algorithms.

In existing researches of recommender system, performance is typically evaluated by analyzing precision and recall rates. Namely, some of the tags in the original content were artificially deleted, and the recommender systems attempted to predict as many deleted tags as possible while suppressing the number of unrelated tags. However in this thesis, we focus more on efficiency of popularity boosting than properly annotations. To evaluate which of the three recommendation results can make images receive the most attention, we consider two approaches: (1) Offline: predict the future popularity scores; (2) Online: upload both images and tags to Flickr to compare the results. In the next two subsections, we will illustrate the experiment design and discuss results.

7.2. Offline Evaluation: Popularity Prediction

The DF-W method has been proved to be efficient in tag based popularity prediction task, so we apply popularity prediction method described in Chapter 6 to see how many number of views a image will gain annotated with different sets of tags, original tags and 10 more tags recommended by different recommendation methods. The predicted results of the two test sets are shown in Table 7.7. The table shows that on test set A, Tagcoor performs the best but we must note that the number of views gained by original tags is so large that recommended tags do not make much difference. whereas on test set B, remarkable increase popularity scores can be observed compare to the number of views obtained by original tags. And among all the recommendation, CF_DF-W works the best, which helps to make the images more than twice than before. We also calculate the Spearman's ρ correlation coefficient of predicted and real popularity of original tags,

Table 7.2 – Example I of recommendation results by different algorithms.

(a) Image

(b) Original tags and Recommendations

Recommendation Method	Tags
Original Tags	gravestones, headstones, leaves, sigma, shadows, leicester-shire, park, pdeee454, saint wistans, sunlight, sunshine, tombstones, trees, wistow, 17-70mm, 450d, canon, church, grass, kilby
Tagcoor	light, nature, sky, leaf, landscape, green, hdr, clouds, blue, sun
CF	nature, beach, water, australia, sky, sea, landscape, flowers, photography, clouds
CF_DF-W	landscape, beach, water, nature, girls, skyline, oregon, clouds, australia, sea
FPRank	hdr, light, nature, sky, landscape, sun, clouds, blue, green, water

TABLE 7.3 – Example II of recommendation results by different algorithms.

(a) Image

(b) Original tags and Recommendations

Recommendation Method	Tags
Original Tags	1 times square, 7th avenue, advertising, architecture, architectural lighting, broadway, one times square, art deco, arts decoratifs, aztec style, brightlights, cc-by-sa, curtain wall, deco, georgerexphotography, paramount building, photographygeorgerex, set back, manhattan, midtown, modernism, neon, new york, new york times tower, ny, grxa23, new york city, new york times building, ziggurat style, theatre district, times building, times square, united states of america, usa, wedding cake style
Tagcoor	nyc, gothamist, landmark, city, nrhp, skyscraper, national register of historic places, photography, sign, new york city landmarks preservation commission
CF	nyc, theatre, theater, gothamist, un marble, facade, empire, united nations, columns
CF_DF-W	gothamist, nyc, theatre, theater, united nations marble, facade, columns, le corbusier, aia150
FPRank	national register of historic places, nrhp, nyclpc, new york city landmarks preservation commission, u.s. national register of historic places, skyscraper, nyc, gothamist, landmark, national historic landmark

TABLE 7.4 – Example III of recommendation results by different algorithms.

(a) Image

(b) Original tags and Recommendations

Recommendation Method	Tags
Original Tags	105mm, 40d, bush, ca, california, canon, costume, devine, dress, face, vine, female, fruit, fruity, grape, green, woman, juice, juicy, lady, paint, plant, zoo, pretty, san diego, sigma, smile, tree
Tagcoor	portrait, girl, nikon, beautiful, model, red, people, photo, food, nature
CF	2012, people, photo, photography, portrait, red, white, fun, girl, digital
CF_DF-W	portrait, girl, people, cosplay, red, 2012, candid, white, photo, photography
FPRank	girl, portrait, beautiful, model, women, blonde, red, sexy, cosplay, milf

TABLE 7.5 – Example IV of recommendation results by different algorithms.

(a) Image

(b) Original tags and Recommendations

Recommendation Method	Tags
Original Tags	d80, europe, explore, explored, favoritesonly, flores, flowers, garden, inexplore, jardin, spain, spring, nikond80, orange, pasotraspaso, photography, photos, jesus solana, naranja, naturaleza, nature, nikon, primavera, prision, prison, rosa, rose
Tagcoor	canon, travel, flower, hdr, landscape españa, macro, blue, light, sky
CF	2009, blue, españa, light, sun, march, reflection, easter, santa, boy
CF_DF-W	2009, blue, light, march, boy, sun, españa, reflection, abigfave, easter
FPRank	hdr, españa, flower, water, abigfave landscape, naturesfinest, soe, blue, sunset

Table 7.6 – Example V of recommendation results by different algorithms.

(a) Image

(b) Original tags and Recommendations

Recommendation Method	Tags
Original Tags	creatures, female, image, images, natural, nature, nevada, phalarope, phalaropus, phalaropus tricolor, photo, photograph, photographs, photos, pic, pics, picture, pictures, shore, tricolor, animal, animals, avian, bird, birds, wading, wild, wildlife, wilson's
Tagcoor	photography, canon, water, flickr, nikon, beach, travel, beautiful, girl, landscape
CF	usa, nikon, united states, photography, zoo, arizona, america, forest, project365, north america
CF_DF-W	nikon, zoo, photography, forest, jpg, project365, arizona, jpeg, north america, usa
FPRank	photography, canon, beautiful, macro, travel, green, blue, animalia, water, insect

Table 7.7 – Predicted popularity of different recommended tag sets.

(a) Test set A

Tag Lists	Average Number of Views	Improvement Ratio
Original tags	1684	–
Original tags + CF	1845	9.56%
Original tags + Tagcoor	1873	11.22%
Original tags + CF_DF-W	2009	19.34%
Original tags + FPRank	1861	10.51%

(b) Test set B

Tag Lists	Average Number of Views	Improvement Ratio
Original tags	161	–
Original tags + CF	277	72.05%
Original tags + Tagcoor	340	111.18%
Original tags + CF_DF-W	375	132.67 %
Original tags + FPRank	360	123.60%

though, it is lower than 0.5. Thus we have to conclude that the prediction results are not so reliable.

7.3. Online Evaluation: Upload to Flickr

Online experiments is one of the most convincing approaches to evaluate efficiency in popularity boosting. We created eight new Flickr accounts and we upload both the images of the two test sets and annotated tags to different accounts. To avoid the situation where the search engine of Flickr returns a series of identical images at the same time during the retrieval process, we only upload images with one of the tag lists, delete all the images after 10 days and then try to upload images with different tag lists to a new account. In this way, the subsequent experiments will not be disturbed by previous results, thus it ensures that the experiments are done independently. We record the number of views of all the images every 12 hours and the whole online experiment lasted from November 2016 to January 2017. The results are shown in Figure 7.2 and Figure 7.3.

As we can see from Figure 7.2 (b) and Figure 7.3 (b), FPRank performed best both under the case of average popularity of each image and average popularity obtained by each tag, which proves that FPRank makes better recommendations with a higher level of influence on popularity boosting over the other three tag recommendation methods. And CF_DF-W also worked well, only a little worse than FPRank.

However, the results of Figure 7.2 (a) and Figure 7.3 (a) are beyond our expectations, where images with tag sets of both original tags and recommended tag by Tagcoor achieve lower popularity than those with only original tags. FPRank performed well at the first few days but gradually the growth rate of average popularity became very slow and finally went below than CF. On the contrary, the average popularity of CF_DF-W was lower than CF at first but finally exceeded CF.

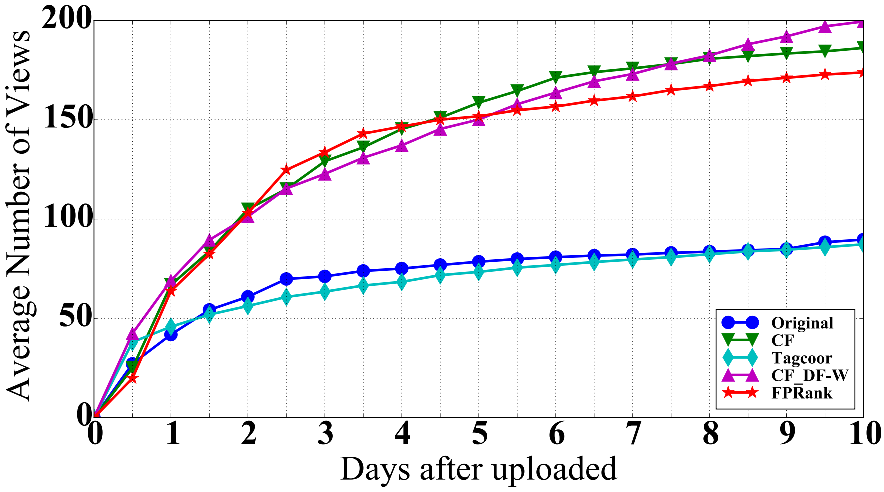
The results on Test set A is so different with those on Test B, since we know that various factors can influence content popularity. Beside content quality and image annotation, user interactions in social media will also cause great impact on popularity, which are really difficult to control. For an example, we find out that one image of an account occasionally favorited by a pro Flickr account, as a result of which this account attracts unexpectedly large number of views.

Another thing we find that affecting the social popularity is the Safe Search option in Flickr[1]. There are three safety levels and they are Safe, Moderate, Restricted. The content of Safe level can be seen by a global public audience, while Moderate and Restricted are only for part of the whole audience. All the new accounts are set to Safe by default. However, we find that there are some photos of people wearing revealing clothes in Test set A, and the safety level of accounts for Test set A were always modified to Moderate by Flickr Administrator at some time during our experiments. This greatly affected the popularity received in these accounts. So the results of accounts for Test set A seems so wired compared to those for Test set B.

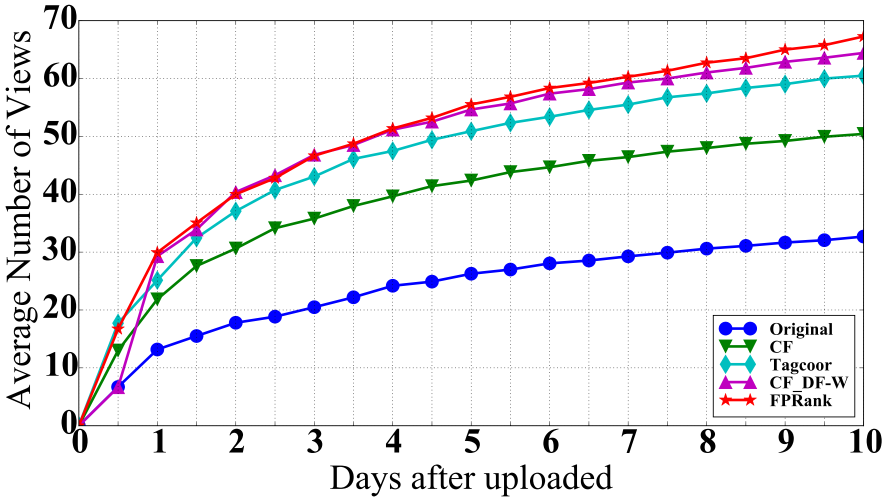
7.4. Tag Quality Assessment

One may doubt that it is possible to assign tags which have a high level of popularity influence but totally unrelated to the image, which will lead to tag spam. In order to eliminate this confusion and prove that our recommendation methods do recommend

[1] https://safety.yahoo.com/SafetyGuides/Flickr/index.htm?.tsrc=lgwn/

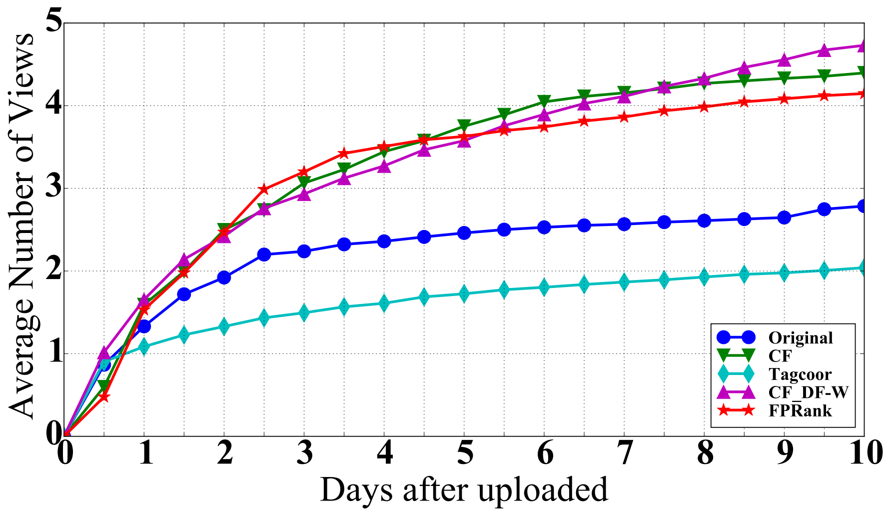


(a) on Test set A

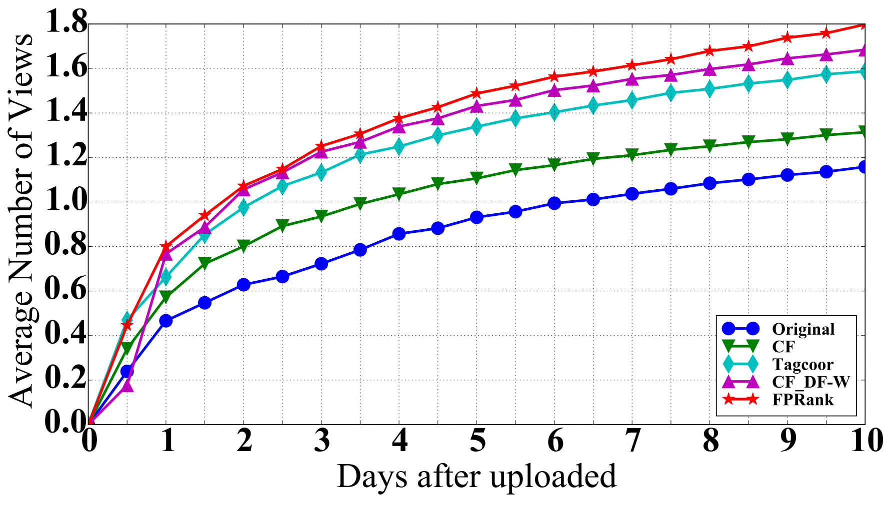


(b) on Test set B

FIGURE 7.2 – Results of online evaluation: average number of views of each image on Test set A and B respectively.



(a) on Test set A



(b) on Test set B

FIGURE 7.3 – Results of online evaluation: average number of views obtained by each tag on Test set A and B respectively.

proper tags based on the content of web objects, we conduct a quality assessment of tags obtained by different recommendation algorithms.

We do our assessment on Amazon's Mechanical Turk [58], which is a famous crowd-sourcing online site enabling individuals and businesses to coordinate the use of human intelligence to perform tasks that computers are currently unable to do. We ask workers to rate a tag on a five-point scale, where 5 stars for good quality and 1 star for bad quality. The content of an image and necessary descriptions are provided for reference, such as title, owner, date taken and etc. Figure 7.4 shows the design of the assessment task. We assign 5 different workers, who has a approval rate of 85 % or higher, to rate each image–tag pair. The whole assessment took about three weeks in January 2017 and more than 1,500 workers did our tasks. We carefully checked all the answers and rejected answers that were uncompleted or of poor quality. A summary of tasks is shown in Table 7.8.

The results of rating scores of tags recommended by different methods are calculated in Table 7.9. And Table 7.10 shows the comparison of every two tag lists. As we can see from the results, original tags are rated the highest, which is without any doubts. The tags recommended by Tagcoor are rated the second highest, and the proposed algorithm rated a little lower than Tagcoor. The results show that FPRank recommend reasonably good tags that can properly describe the images. But the average rating of CF_DF-W is the worst among all the methods, even worse than the basic CF.

Table 7.8 – Summary of tasks done on Amazon Mturk.

(a) on Test Set A

Session	Total Number of Tasks	Number of Workers	Average Work Time (s)
0	1250	270	148.38
1	1250	286	130.73
2	1250	235	109.08
3	1250	121	111.65
4	1250	154	117.05
5	1250	171	116.40
6	1250	106	94.55
7	1250	112	104.50
8	1250	120	92.80
9	1255	117	100.69
10	240	107	152.92
Total	12745	1143	113.34

(a) on Test Set B

Session	Total Number of Tasks	Number of Workers	Average Work Time (s)
0	1350	147	96.39
1	1349	155	100.34
2	1350	125	90.95
3	1350	150	102.53
4	1350	161	104.38
5	1351	117	95.57
6	1350	130	76.00
7	1350	157	80.68
8	1373	195	82.65
9	1030	129	92.80
Total	13203	631	92.20

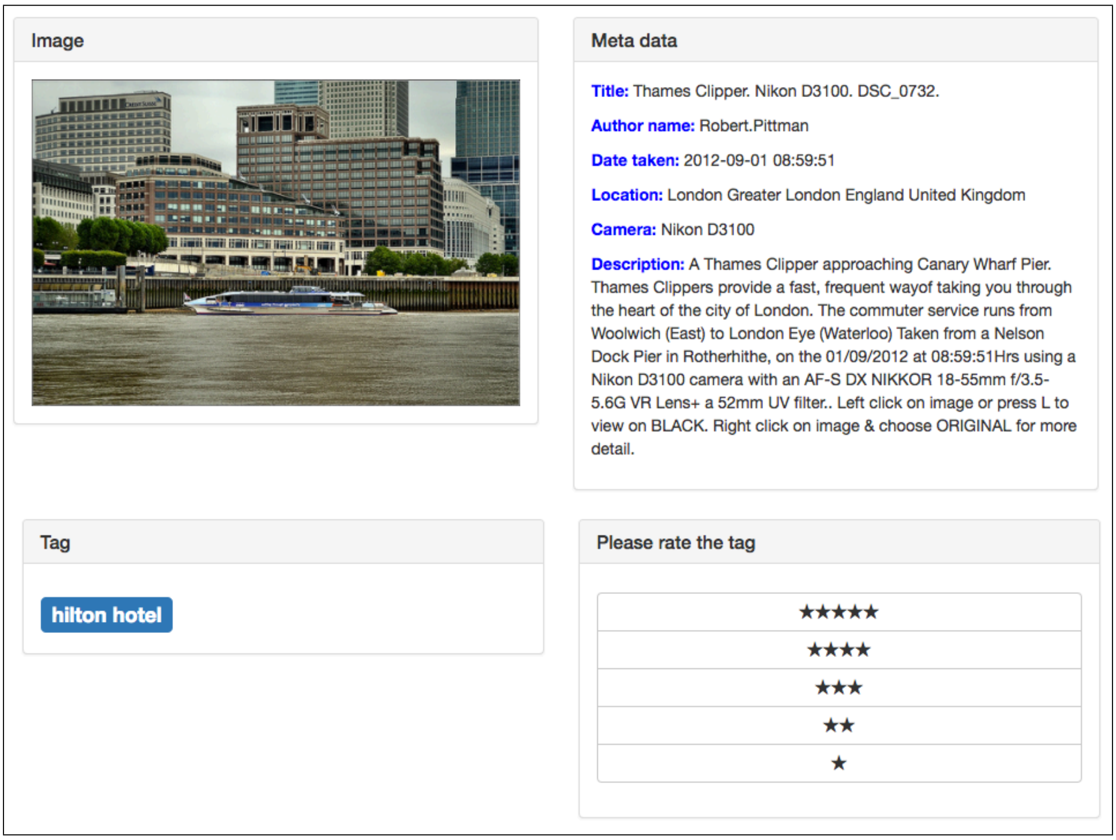


FIGURE 7.4 – Sample design of tag quality assessment task.

TABLE 7.9 – Results of tag quality assessment.

(a) Test set A

Tag Lists	Number of tags	Average Score	Standard Deviation
Original	29,130	3.008	1.303
CF	8,620	2.680	1.310
Tagcoor	8,620	2.887	1.314
CF_DF-W	8,620	2.672	1.310
FPRank	8,620	2.809	1.313

(b) Test Set B

Tag Lists	Number of tags	Average Score	Standard Deviation
Original	26,958	2.886	1.288
CF	9,470	2.695	1.288
Tagcoor	9,470	2.859	1.282
CF_DF-W	9470	2.680	1.289
FPRank	9,470	2.791	1.287

Table 7.10 – Comparison of every two recommendation methods.

(a) Test set A

Method A	Method B	Cnt($\overline{A}_{score} \geq \overline{B}_{score}$)	Cnt($\overline{A}_{score} < \overline{B}_{score}$)
original	cff	552	310
original	cfdfw	564	298
original	tagcoor	482	380
original	fprank	538	324
cff	cfdfw	463	399
cff	tagcoor	336	526
cff	fprank	384	478
cfdfw	tagcoor	333	529
cfdfw	fprank	365	497
tagcoor	fprank	480	382

(b) Test Set B

Method A	Method B	Cnt($\overline{A}_{score} \geq \overline{B}_{score}$)	Cnt($\overline{A}_{score} < \overline{B}_{score}$)
original	cff	571	376
original	cfdfw	559	388
original	tagcoor	497	450
original	fprank	521	426
cff	cfdfw	489	458
cff	tagcoor	384	563
cff	fprank	418	529
cfdfw	tagcoor	400	547
cfdfw	fprank	424	523
tagcoor	fprank	507	440

Chapter 8

Conclusions and Future Work

In this thesis, we presented a study of popularity on image/video sharing sites by analyzing text tags. We propose two tag ranking methods, DF-W and FolkPopularityRank, which can extract tags that highly influence content popularity. We applied the two method to a real-world dataset of Flickr, and compare the top ranked tags by different ranking methods. We then discussed three applications of the DF-W tag ranking method and showed the efficiency of the two proposed algorithms in popularity prediction and tag recommendations.

Firstly, we investigated the spatial and temporal changes of tags that significantly affect popularity, and explored the evolution of community focus from time to time, place to place, which allows us to look into difference of online user behavior and increase the effectiveness of popularity study. Besides, we analyzed top ranking tags over various image/video sharing media and reveal main differences among mainstream media.

Secondly, we presented an approach to predict popularity of web content before uploading them to social services based on our proposed tag ranking methods. We contrasted our methods with simpler tag rankings and experiment results show the efficiency of our methods in popularity prediction. The experiment results showed that DF-W outperformed all the other ranking methods in popularity prediction. To improve the prediction accuracy, we also conducted prediction in spatial and temporal dimension and used a multimodal learning approach for popularity prediction, which try to use both of tag feature and visual feature in social popularity prediction task.

And thirdly, we presented two new tag recommendation strategies based on the two proposed tag ranking methods, which focus more on popularity boosting rather than merely semantics and descriptive annotations. A series of experiments results showed that FPRank performed well in recommending tags that not only properly describe the content and also help to attract more popularity, while a DF-W based method did not work well in tag recommendations.

The spatial and temporal tag analysis is quite interesting and meaning. In the future, we would like to automatically detect important events through tag analysis. The work on multimodal learning for popularity prediction and tag recommendations for popularity boosting are incomplete and more effort is necessary. In the future work of multimodal learning for popularity prediction, we would like to try to use different deep learning models or train our own model, and other visual features, such as color, histogram, GIST feature, SIFT feature. As follow up on tag recommendations for popularity boosting task, we plan to make better experiment design and conduct an online evaluation experiments to see how real popularity differ. Currently, we mainly performed our work on the dataset of Flickr, and future work further includes trying on other social media.

References

[1] A. M. Kaplan and M. Haenlein, "Users of the world, unite! the challenges and opportunities of social media," *Business horizons*, vol. 53, no. 1, pp. 59–68, 2010.

[2] http://www.internetlivestats.com/twitter-statistics.

[3] http://expandedramblings.com/index.php/flickr-stats.

[4] https://www.youtube.com/yt/press/statistics.html.

[5] M. Cha, H. Kwak, P. Rodriguez, Y.-Y. Ahn, and S. Moon, "I tube, you tube, everybody tubes: analyzing the world's largest user generated content video system," in *Proceedings of the 7th ACM SIGCOMM conference on Internet measurement.* ACM, 2007, pp. 1–14.

[6] https://www.instagram.com/press/.

[7] 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.

[8] T. Vander Wal, "Folksonomy," 2007.

[9] T. Vander Wal, "Explaining and showing broad and narrow folksonomies," 2005.

[10] M. Ames and M. Naaman, "Why we tag: motivations for annotation in mobile and online media," in *Proceedings of the SIGCHI conference on Human factors in computing systems*, 2007, pp. 971–980.

[11] K. Lerman and L. Jones, "Social browsing on flickr," *arXiv preprint cs/0612047*, 2006.

[12] A. O. Nwana, S. Avestimehr, and T. Chen, "A latent social approach to youtube popularity prediction," in *IEEE GLOBECOM*, 2013, pp. 3138–3144.

[13] A. Khosla, A. Das Sarma, and R. Hamid, "What makes an image popular?" in *Proceedings of the 23rd international conference on World wide web.* ACM, 2014, pp. 867–876.

[14] R. van Zwol, A. Rae, and L. Garcia Pueyo, "Prediction of favourite photos using social, visual, and textual signals," in *Proceedings of the international conference on Multimedia*, 2010, pp. 1015–1018.

[15] J. San Pedro and S. Siersdorfer, "Ranking and classifying attractiveness of photos in folksonomies," in *Proceedings of the 18th international conference on World wide web*. ACM, 2009, pp. 771–780.

[16] L. Weng, A. Flammini, A. Vespignani, and F. Menczer, "Competition among memes in a world with limited attention," *Scientific reports*, vol. 2, 2012.

[17] B. Sigurbjörnsson and R. Van Zwol, "Flickr tag recommendation based on collective knowledge," in *WWW*, 2008, pp. 327–336.

[18] G. Szabo and B. A. Huberman, "Predicting the popularity of online content," *Communications of the ACM*, vol. 53, no. 8, pp. 80–88, 2010.

[19] F. Figueiredo, F. Benevenuto, and J. M. Almeida, "The tube over time: characterizing popularity growth of youtube videos," in *Proceedings of the fourth ACM international conference on Web search and data mining*. ACM, 2011, pp. 745–754.

[20] E. Ferrara, R. Interdonato, and A. Tagarelli, "Online popularity and topical interests through the lens of instagram," in *Proceedings of the 25th ACM conference on Hypertext and social media*. ACM, 2014, pp. 24–34.

[21] H. Kwak, C. Lee, H. Park, and S. Moon, "What is twitter, a social network or a news media?" in *Proceedings of the 19th international conference on World wide web*. ACM, 2010, pp. 591–600.

[22] T. Zaman, E. B. Fox, E. T. Bradlow *et al.*, "A bayesian approach for predicting the popularity of tweets," *The Annals of Applied Statistics*, vol. 8, no. 3, pp. 1583–1611, 2014.

[23] L. Hong, O. Dan, and B. D. Davison, "Predicting popular messages in twitter," in *Proceedings of the 20th international conference companion on World wide web*. ACM, 2011, pp. 57–58.

[24] A. Tatar, M. D. de Amorim, S. Fdida, and P. Antoniadis, "A survey on predicting the popularity of web content," *Journal of Internet Services and Applications*, vol. 5, no. 1, pp. 1–20, 2014.

[25] G. Chatzopoulou, C. Sheng, and M. Faloutsos, "A first step towards understanding popularity in youtube," in *INFOCOM IEEE Conference on Computer Communications Workshops, 2010*. IEEE, 2010, pp. 1–6.

[26] 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*. IEEE, 2010, pp. 177–184.

[27] J. Berger and K. L. Milkman, "What makes online content viral?" *Journal of marketing research*, vol. 49, no. 2, pp. 192–205, 2012.

[28] S. Stieglitz and L. Dang-Xuan, "Emotions and information diffusion in social media —sentiment of microblogs and sharing behavior," *Journal of Management Information Systems*, vol. 29, no. 4, pp. 217–248, 2013.

[29] F. Figueiredo, J. M. Almeida, F. Benevenuto, and K. P. Gummadi, "Does content determine information popularity in social media?: A case study of youtube videos' content and their popularity," in *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems.* ACM, 2014, pp. 979–982.

[30] J. Ratkiewicz, F. Menczer, S. Fortunato, A. Flammini, and A. Vespignani, "Traffic in social media ii: Modeling bursty popularity," in *Social Computing (SocialCom), 2010 IEEE Second International Conference on.* IEEE, 2010, pp. 393–400.

[31] Y. Borghol, S. Ardon, N. Carlsson, D. Eager, and A. Mahanti, "The untold story of the clones: content-agnostic factors that impact youtube video popularity," in *Proceedings of the 18th ACM SIGKDD international conference on Knowledge discovery and data mining.* ACM, 2012, pp. 1186–1194.

[32] R. Zhou, S. Khemmarat, L. Gao, and H. Wang, "Boosting video popularity through recommendation systems," in *Databases and Social Networks.* ACM, 2011, pp. 13–18.

[33] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *science*, vol. 286, no. 5439, pp. 509–512, 1999.

[34] H. Pinto, J. M. Almeida, and M. A. Gonçalves, "Using early view patterns to predict the popularity of youtube videos," in *Proceedings of the sixth ACM international conference on Web search and data mining.* ACM, 2013, pp. 365–374.

[35] K. Lerman and T. Hogg, "Using a model of social dynamics to predict popularity of news," in *Proceedings of the 19th international conference on World wide web.* ACM, 2010, pp. 621–630.

[36] P. J. McParlane, Y. Moshfeghi, and J. M. Jose, "Nobody comes here anymore, it's too crowded; predicting image popularity on flickr," in *Proceedings of International Conference on Multimedia Retrieval.* ACM, 2014, p. 385.

[37] F. Gelli, T. Uricchio, M. Bertini, A. Del Bimbo, and S.-F. Chang, "Image popularity prediction in social media using sentiment and context features," in *Proceedings of the 23rd Annual ACM Conference on Multimedia Conference.* ACM, 2015, pp. 907–910.

[38] B. Sarwar, G. Karypis, J. Konstan, and J. Riedl, "Item-based collaborative filtering recommendation algorithms," in *WWW*, 2001, pp. 285–295.

[39] D. Jannach, M. Zanker, A. Felfernig, and G. Friedrich, *Recommender systems: an introduction.* Cambridge University Press, 2010.

[40] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme, "Tag recommendations in social bookmarking systems," *Ai Communications*, vol. 21, no. 4, pp. 231–247, 2008.

[41] J. Leskovec, A. Rajaraman, and J. D. Ullman, *Mining of massive datasets.* Cambridge University Press, 2014, ch. Data Mining.

[42] T. Yamasaki, S. Sano, and K. Aizawa, "Social popularity score: Predicting numbers of views, comments, and favorites of social photos using only annotations," in *WISMM*, 2014, pp. 3–8.

[43] L. Page, S. Brin, R. Motwani, and T. Winograd, "The pagerank citation ranking: Bringing order to the web." Stanford InfoLab, Tech. Rep., 1999.

[44] B. Thomee, D. A. Shamma, G. Friedland, B. Elizalde, K. Ni, D. Poland, D. Borth, and L.-J. Li, "The new data and new challenges in multimedia research," *arXiv preprint arXiv:1503.01817*, pp. 3–10, 2015.

[45] https://www.flickr.com/services/api/.

[46] D. Comaniciu and P. Meer, "Mean shift: A robust approach toward feature space analysis," *IEEE Transactions on pattern analysis and machine intelligence*, vol. 24, no. 5, pp. 603–619, 2002.

[47] E. Ferrara, R. Interdonato, and A. Tagarelli, "Online popularity and topical interests through the lens of instagram," *CoRR*, vol. abs/1406.7751, pp. 24–34, 2014.

[48] http://www.nii.ac.jp/dsc/idr/nico/nico.html.

[49] Y. Jia, E. Shelhamer, J. Donahue, S. Karayev, J. Long, R. Girshick, S. Guadarrama, and T. Darrell, "Caffe: Convolutional architecture for fast feature embedding," *arXiv preprint arXiv:1408.5093*, 2014.

[50] R.-E. Fan, K.-W. Chang, C.-J. Hsieh, X.-R. Wang, and C.-J. Lin, "Liblinear: A library for large linear classification," *JMLR*, vol. 9, pp. 1871–1874, 2008.

[51] C.-C. Chang and C.-J. Lin, "Libsvm: A library for support vector machines," *TIST*, vol. 2, no. 3, p. 27, 2011.

[52] M. Gönen and E. Alpaydın, "Multiple kernel learning algorithms," *JMLR*, vol. 12, pp. 2211–2268, 2011.

[53] S. Sonnenburg, G. Rätsch, S. Henschel, C. Widmer, J. Behr, A. Zien, F. d. Bona, A. Binder, C. Gehl, and V. Franc, "The shogun machine learning toolbox," *JMLR*, vol. 11, pp. 1799–1802, 2010.

[54] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," in *Advances in Neural Information Processing Systems 25*, F. Pereira, C. Burges, L. Bottou, and K. Weinberger, Eds., 2012, pp. 1097–1105.

[55] R. Girshick, J. Donahue, T. Darrell, and J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation," in *CVRP*, 2014.

[56] K. Chatfield, K. Simonyan, A. Vedaldi, and A. Zisserman, "Return of the devil in the details: Delving deep into convolutional nets," in *British Machine Vision Conference*, 2014.

[57] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *arXiv preprint arXiv:1409.1556*, 2014.

[58] M. Buhrmester, T. Kwang, and S. D. Gosling, "Amazon's mechanical turk a new source of inexpensive, yet high-quality, data?" *Perspectives on psychological science*, vol. 6, no. 1, pp. 3–5, 2011.

Publications

International Conferences

[1] Toshihiko Yamasaki, Jiani Hu, Kiyoharu Aizawa, Tao Mei. Power of Tags: Predicting Popularity of Social Media in Geo-Spatial and Temporal Contexts. *Pacific Rim Conference on Multimedia (PCM 2015)*, pp.149-158, September 16-18, 2015, Gwangju, Korea (invited).

[2] Jiani Hu, Toshihiko Yamasaki, Kiyoharu Aizawa. Multimodal Learning for Image Popularity Prediction on Social Media. *IEEE International Conference on Consumer Electronics-Taiwan (ICCE-TW 2016).*

Domestic Conferences and Symposia

[3] 胡 佳微, 佐野峻平, 山崎俊彦. 画像共有 SNS における人気へ影響を及ぼすタグの時系列解析. 電子情報通信学会サイバーワールド研究会, CW2014-16, pp. 39-42, Chiba, Chiba, Feb. 25, 2015.

[4] Jiani Hu, Toshihiko Yamasaki, Kiyoharu Aizawa, Tao Mei. Spatial and Temporal Analysis of Influential Tags on Social-Popularity in Image Sharing Services. 電子情報通信学会　マルチメディア・仮想環境基礎研究会, 信学技報, vol. 115, no. 125, MVE2015-20, pp. 73-78, July 2-3, 2015, Tokyo.

[5] Jiani Hu, Toshihiko Yamasaki, Kiyoharu Aizawa, Tao Mei. How Popular Will Your Uploaded Image Be: Social Popularity Estimation in Image Sharing Services Using Text Tags Only. 画像の認識・理解シンポジウム（*MIRU 2015)*, SS2-10, July 26-30, 2015, Osaka.

[6] Jiani Hu, Toshihiko Yamasaki, Kiyoharu Aizawa. Tag Recommendations in Social Media for Popularity Boosting. 画像符号化シンポジウム・映像メディア処理シンポジウム（*PCSJ/IMPS 2016)*, SS2-10, Nov 16-18, 2016, Shuzenji.

[7] Jiani Hu, Toshihiko Yamasaki, Kiyoharu Aizawa. Tag Recommendations in Social Media for Popularity Boosting. 映像情報メディア学会 *ME*研究会, Feb 20-21, 2017, Hokkaido, Submitted.

Technical Tutorials

[8] Jiani Hu, 山崎俊彦: "ゼロから始めるクラウドコンピューティング〜Microsoft Azure 入門編〜" (映像情報メディア学会誌 2015 年度 4 月号) (2015)

[9] Jiani Hu, 山崎俊彦: "ゼロから始めるクラウドコンピューティング〜Microsoft Azure 発展編〜" (映像情報メディア学会誌 2015 年度 5 月号) (2015)

Acknowledgements

I take this opportunity to express my deepest appreciation to my supervisor, Prof. Toshihiko Yamasaki, for the exemplary guidance, monitoring and constant encouragement he has provided throughout my time as his student. I have been extremely lucky to have a supervisor whose expertise and understanding, added considerably to my graduate experience. During the two and half years in the University of Tokyo, he have supported me in developing literacy in science and walked me through all the stages of the completion of this thesis. Without his continued support and excellent suggestions to the undertaking of the research, the completion of this thesis would be impossible.

A special gratitude I give to Prof. Kiyoharu Aizawa, whose contribution in stimulating suggestions and encouragement helped me to coordinate research. Many thanks to the secretary of our lab, Ms. Masaki Matsubayashi, for her thoughtful kindness in both school affairs and daily life. Furthermore I would also like to thank all members of Aizawa-Yamasaki Lab for their personal guidance and the provided valuable information in their respective fields. Special Thanks to the students of the same year for the wonderful days we spent together.

Finally, I would like to thank my beloved family, who always support and understand me all through these days. Last but not the least, I would like to thank my friends for their support in my lift in Tokyo.

Jiani Hu

January 27, 2017