# Geotag Propagation with User Trust Modeling

Ivan Ivanov, Peter Vajda, Jong-Seok Lee, Pavel Korshunov, Touradj Ebrahimi

## 1 Introduction

Social networks and photo sharing websites have become increasingly popular in recent years. Their services typically focus on building online communities of people who interact with each other by sharing their own interests or activities and exploring shared content of others. Such social networks have become a popular way to disseminate different type of information, such as photo, video, text, and audio. For example, a user uploads wedding album to let other people from the online community comment or rate the photos. This sharing trend has resulted in a continuously growing volume of publicly available photos on such websites like Flickr[1], Picasa[2] or Photobucket[3], as well as social networks like Facebook[4] and Google+[5]. For instance, Photobucket hosts more than 8 billion photos [20], 7 billion photos are hosted on Picasa [20], and 6 billion photos on Flickr [43]. Facebook has more than 250 million photos posted to its network every day [35] and approximately 100

I. Ivanov (✉) · P. Vajda · P. Korshunov · T. Ebrahimi
Multimedia Signal Processing Group (MMSPG), École Polytechnique Fédérale de Lausanne (EPFL)
1015 Lausanne, Switzerland
e-mail: ivan.ivanov@epfl.ch, peter.vajda@epfl.ch, pavel.korshunov@epfl.ch, touradj.ebrahimi@epfl.ch

J.-S. Lee
School of Integrated Technology, Yonsei University
Incheon 406-840, South Korea
e-mail: jong-seok.lee@yonsei.ac.kr

[1] http://www.flickr.com
[2] http://picasa.google.com
[3] http://www.photobucket.com
[4] http://www.facebook.com
[5] http://plus.google.com

billion photos stored on its servers [4], while 3.4 billion photos have been uploaded to Google+ [4] in the first 100 days of it being open to the public. This large volume of multimedia content poses significant challenges for efficient search, retrieval, and processing of the shared content.

Tagging is one of the popular methods to categorize large volume of photos. It is a process by which users assign short textual annotations to photos (in the form of keywords) to describe them and to provide additional information for search engines, online photo albums, and for people browsing the photo collections. Tags, when combined with search technologies, are essential in resolving user queries targeting shared photos. The success of social networks such as Flickr, Google+, and Facebook proves that users are willing to provide tags through manual annotations. Different users annotating the same photo can enrich the information about that photo. However, tagging a lot of photos by hand is a time-consuming task. Users typically tag a small number of the shared photos only, leaving most of the other photos with incomplete metadata. This lack of metadata decreases the precision of search, because photos without proper annotations are typically much harder to retrieve than correctly annotated photos. Therefore, to help people organize and browse large collections of personal photos in an effective way, it is important to develop robust and efficient algorithms for automatic tagging or tag propagation.

Another important challenge in tagging is to identify most appropriate tags for given content, and at the same time, to eliminate noisy or spam tags. The shared photos are sometimes assigned with inappropriate tags for several reasons. First of all, users are human beings and make mistakes. It is also possible that misleading tags can be assigned for advertisement purposes, self-promotion, or to increase the rank of a particular tag in the search engines. Consequently, free-form keywords (tags) assigned to photos carry a significant risk that wrong or irrelevant tags eventually prevent users from the intended benefits of annotated photos. Finally, wrong machine tags, such as longitude and latitude, can be automatically assigned to images captured with cameras equipped with GPS devices due to bad or noisy communication channels with GPS satellites or wireless access points. Kennedy *et al.* [18] analyzed the Flickr website and revealed that the tags provided by users are often imprecise and only around 50% of tags are truly related to an image. Beside the tag-photo association, spam objects can take other forms, i.e. possibly manifesting as a spam photo or a spam user (spammer). Therefore, for the practical tag propagation system, it is important to consider user trust information derived from users' tagging behavior.

Trust provides a natural security policy stipulating that users or photos with low trust values should be investigated or eliminated. Trust can predict the future behavior of users in order to avoid undesirable influences of untrustworthy users. Trust-based schemes can be used to motivate users to positively contribute to social networks and/or penalize adversaries that are trying to disrupt the network. Therefore, the distribution of the trust values associated with either users or photos in a social network can represent the health of the network and used in a spam-free tag propagation algorithm.

Table 1: Summary of representative recent techniques that combine geographical context and visual content for automatic geotagging of images.

| Reference | Descriptor | Method | Application |
|---|---|---|---|
| Hays and Efros [11] | visual features | the probability distribution for the location of an unknown image is found on the globe using a purely data-driven scene matching | non-landmark (scene) location recognition |
| Kennedy and Naaman [19] | visual and textual features | for a given location diverse and representative images are generated based on geotagged community images | visual summary of landmarks |
| Zheng *et al.* [48] | visual features & GPS coordinates | travel blogs and geotagged images are analyzed and a list of tourist landmarks is established based on the information from nearest neighbors | landmark recognition |
| Quack *et al.* [36] | visual and textual features & GPS coordinates | objects and events are retrieved from a large-scale collection of geotagged images using pair-wise similarity | event/scene understanding |

In this chapter, we focus on trust aspects and trust models used in social networks and applicability of these models in automatic tag propagation systems. In Sections 2, we first discuss geotagging and how it is used in various social networks and media retrieval systems. In Section 3, we introduce several techniques used for combatting noise and spam through trust modeling in social tagging systems. In Section 4, we present detailed overview of several trust modeling approaches, which are specific to geotagging systems. And, in Section 5, we demonstrate the advantages of using trust modeling on an example of automatic geotag propagation system in travel related photos. We conclude the chapter with Section 6.

## 2 Geotagging in social networks and sharing websites

The proposed system is related to different research fields including visual analysis, geographic information systems, social networking and tagging. Therefore, the goal of

In this section, we review the most relevant work in the fields of joint analysis of visual content and geographical context, and human tagging.

In the last several years, an important trend in multimedia understanding is modeling and extracting value from geographical context, such as GPS coordinates, and visual content, such as a description of a photo. Different research problems and significant approaches in this field are summarized by Luo *et al.* [27]. In this section, we focus on some of the representative image retrieval approaches that rely on a variety of image or landmark descriptors combined with geographic information. These approaches are summarized in Table 1.

A pioneering paper in this area by Hays and Efros [11] proposed an algorithm called *IM2GPS* to estimate the locations of a single image using a purely data-driven scene matching approach. Given a test image, the algorithm finds the visual nearest neighbors in the database and estimates a geolocation of the image from the GPS coordinates of the tagged nearest neighbors. The estimated image location is represented as a probability distribution over the Earth's surface. However, the IM2GPS approach showed low recognition accuracy due to low-level features. While IM2GPS uses a set of more than 6 million training images, its general applicability is inconclusive because the performance was verified only on 237 hand-selected test images.

Kennedy and Naaman [19] presented a method to search representative landmark images from a large collection of geotagged images. This method uses tags and the geographical location representing a landmark. The visual features (global color and texture features, and Scale Invariant Feature Transform (SIFT)) are analyzed to cluster landmark images into visually similar groups. The method has been proven to be effective for extraction of the representative image sets for a given landmark. But since it cannot be applied to untagged images, its applicability is limited.

The recent work of Zheng *et al*. [48] automatically finds frequently photographed landmarks from a large collection of geotagged photos. The authors perform clustering on GPS coordinates and visual texture features from the image pool and extract landmark names as the most frequent tags associated with the particular visual cluster. Additionally, they extract landmark names from the travel guide articles, such as Wikitravel[6], and visually cluster photos gathered by querying Google Images[7]. However, the test set they use is quite limited – 728 images in total for a 124-category problem, or less than 6 test images per landmark.

Another application that combines textual and visual techniques has been proposed by Quack *et al.* [36]. The authors developed a system that crawls photos on the internet and identifies clusters of images referring to a common object (physical items on fixed locations), and events (special social occasions taking place at certain times). The clusters are created based on the pair-wise visual similarities between the images, and the metadata of the clustered photos are used to derive labels for the clusters. Finally, Wikipedia[8] articles are attached to the images and the validity of these associations is checked. Gammeter *et al.* [9] extends this idea towards object-based auto-annotation of holiday photos in a large database that includes landmark buildings, statues, scenes, pieces of art, with help of external resources such as Wikipedia. In both [36] and [9], GPS coordinates are used to pre-cluster objects which may not be always available.

Most of the photo sharing websites (e.g., Flickr, Picasa, Panoramio[9], Zooomr[10]), provide information about where images were taken in form of maps or groups.

[6] `http://www.wikitravel.com`
[7] `http://images.google.com`
[8] `http://www.wikipedia.org`
[9] `http://www.panoramio.com`
[10] `http://www.zooomr.com`

This information is either provided by an external GPS sensor and stored as image metadata (Exchangeable Image File Format (EXIF) [38], International Press Telecommunications Council (IPTC) [14]), or manually annotated via geocoding.

The main disadvantages of the above systems is that they rely on GPS coordinates to derive geographical annotation, which is not available for the majority of web images and photos, since only a few camera models are equipped with GPS devices. Furthermore, a GPS sensor in a camera provides only the location of the photographer instead of that of the captured landmark, which may be up to several kilometers away. Therefore, the GPS coordinates alone may not be enough to distinguish between two landmarks within a city. Describing landmarks through location names rather than GPS coordinates is not only more reliable but also more expressive. A recent study by Hollenstein and Purves [13] indicated that geotagging should follow the way people actually describe locations, i.e. it is more convenient to use: Church of Saint Sava in Belgrade, rather than: latitude 44.798083, longitude 20.46855. Therefore, there is a growing interest in the research community to derive geographic locations of the scenes in photos based on visual and text features.

## 3 Trust modeling in social media

When information is exchanged on the Internet, malicious individuals are everywhere trying to take advantage of the information exchange structure for their own benefit, while bothering and spamming others. Before social tagging became popular, spam content was observed in various domains: first in e-mail (e.g., [37]), and then in web search (e.g., [8]). Peer-to-peer (P2P) networks have been also influenced by malicious peers, and thus various solutions based on trust and reputation have been proposed, which dealt with collecting information on peer behavior, scoring and ranking peers, and responding based on the scores [30]. Nowadays, even blogs are spammed [39]. Ratings in online reputation systems, such as eBay[11], Amazon[12] and Epinions[13], are very similar to tagging systems and they may face the problem of unfair ratings by artificially inflating or deflating reputations [17]. Several filtering techniques for excluding unfair ratings are proposed in the literature (e.g., [42], [47]). Unfortunately, the countermeasures developed for the e-mail and web spam do not directly apply to social networks and photo sharing websites [12].

In order to reduce or eliminate spams in social networks, various anti-spam methods have been proposed in the state-of-the-art research [15]. Heymann *et al.* [12] classified anti-spam strategies into three categories: prevention, detection, and demotion. *Prevention-based approaches* aim at making it difficult for spam content to contribute to social networks by restricting certain access types through interfaces (such as CAPTCHA [1] or reCAPTCHA [2]) or through usage limits (such as

[11] `http://www.ebay.com`
[12] `http://www.amazon.com`
[13] `http://www.epinions.com`

tagging quota, e.g., Flickr introduced a limit of 75 tags per photo [46]). *Detection approaches* identify likely spams either manually or automatically by making use of, for example, machine learning (such as text classification) or statistical analysis (such as link analysis), and then deleting the spam content or visibly marking it as hidden to users. Finally, *demotion-based approaches* reduce the prominence of content likely to be spam. For instance, rank-based methods produce ordering of a network's content, tags or users based on their trust scores. The prevention-based approaches can be considered as a type of precaution to prevent spammers. However, they cannot completely secure a social network. Some studies, e.g., [32], showed that CAPTCHA systems can be defeated by computers with around 90% of accuracy, using, for example, optical character recognition or shape context matching. Even if prevention methods were perfect, there would be still possibility that the social networks get polluted with spam (malicious) or irrelevant tags. Therefore, detection and demotion via trust modeling are required to keep a network free of noise and spam.

In a social network with tagging capability, spam or noise can be injected at three different levels: spam content (in our case photos, but might be any piece of information - videos, textual documents or web pages), spam tag-content association and spammer [28]. Trust modeling can be performed at each level separately (e.g., [28]) or different levels can be considered jointly to produce trust models, for example, to assess a user's reliability, one can consider not only the user profile, but also the content that the user uploaded to a social network (e.g., [23]). Trust modeling approaches can be categorized into two classes according to the target of trust, i.e. content and user trust modeling.

Content trust modeling is to classify content (e.g., web pages, images, videos) as spam or legitimate. In this case, the target of trust is a content, and thus a trust score is given to each content. Approaches for content trust modeling utilize features extracted from content information, users' profiles and/or associated tags to detect specific spam content. Gyongyi *et al.* [10] proposed an algorithm called TrustRank to semi-automatically separate reputable from spam web pages. TrustRank relies on an important empirical observation called approximate isolation of the good set: good pages seldom point to bad ones. It starts from a set of seeds selected as high-qualified, credible and popular web pages in the web graph, and then iteratively propagate trust scores to all nodes in the graph by splitting the trust score of a node among its neighbors according to a weighting scheme. TrustRank effectively removes most of the spam from the top-scored web pages, however it is unable to effectively separate low-scored good sites from bad ones, due to the lack of distinguishing features. In search engines, TrustRank can be used either solely to filter search results, or in combination with PageRank and other metrics to rank content in search results. Wu *et al.* [44] proposed a computer vision-based technique that discriminates spam images from legitimate ones. By assuming that images containing text are likely to be spam (e.g., banners), they identified a number of useful low-level image features detecting embedded text and computer-generated graphics. Then, pattern classification using support vector machines (SVMs) was performed to classify spam and non-spam images. Although they reported a high detection rate

with a low false positive rate, this approach has limitations in that the discriminant capability of the used features may be limited and, moreover, the assumption that images containing text or computer-generated images are likely to be spam may not be true in some cases.

In user trust modeling, trust is given to each user based on the information extracted from a user's account, his/her interaction with other participants within the social network, and/or the relationship between the content and tags that the user contributed to the social network. Given a user trust score, the user might be flagged as a legitimate user or spammer.

Most of user trust modeling techniques use machine learning approaches applied to features specific to considered social network domains. Krause *et al.* [23] employed a machine learning approach to identify spammers in BibSonomy[14]. They investigated features considering information about a user's profile (e.g., number of digits in the username and the email address), location (e.g., number of spam users with the same IP), bookmarking activity (e.g., number of tags per post), and context of tags (e.g., user co-occurrences with spammers related to tags, content and tag-content pairs). By making use of these features and SVM or naive Bayes classifier, they were able to distinguish legitimate users from malicious ones. It was found that the co-occurrence features describing the usage of a similar vocabulary and content usage are the most promising.

Markines *et al.* [28] proposed six different tag-, content- and user-based features for automatic detection of spammers in BibSonomy. First, tag- and content-based features are averaged across each user's posts, then combined with user-based features, and finally fed into a supervised learning algorithm (such as LogitBoost or AdaBoost) to discriminate spammers from legitimate users. It was shown that TagSpam feature (probability that a particular tag is used to spam, aggregated across all tags assigned to a content) is the best predictor of spammers among all other features, because spammers tend to use certain "suspect" tags more than legitimate users. DomFp feature (likelihood that a content is spam based on its structure) also appeared important, but may not be available since it relies on an infrastructure to enable access to the content, and therefore its feasibility depends on the circumstances of a particular social tagging system.

Noll *et al.* [34] introduced the time of tagging as an additional dimension for assessing the trust of a user in Delicious[15]. They proposed a graph-based algorithm, called SPEAR (SPamming-resistant Expertise Analysis and Ranking). It computes the expertise score of a user and the quality score of a content which are dependent on each other. The time of tagging is considered so that the earlier a user tags a content, the more expertise score he/she receives. These two scores are calculated iteratively in a similar way to that of the Hyperlink-Induced Topic Search (HITS) algorithm. It was shown that SPEAR produces better ranking of users than the HITS method. SPEAR was able to demote different types of spammers (flooders, promoters and trojans [34]) and remove them from the top of the ranking.

---

[14] `http://www.bibsonomy.org`

[15] `http://www.delicious.com`

It can be noted that approaches based on user trust modeling are more common than content trust modeling. One reason is that the user-centered model is simpler to describe than content-centered. Also, user trust models can quickly adapt to the constantly evolving and changing environment in social systems due to the type of features used for modeling, and thus be applicable longer than content trust models, without need for creation of new models. On the other hand, user trust modeling has a disadvantage of "broad brush", i.e. it may be excessively strict if a user happens to post one bit of questionable content on otherwise legitimate content. Trustworthiness of a user is often judged based on the content that the user uploaded to a social system, and thus "subjectivity" in discriminating spammers from legitimate users remains an issue for user trust modeling as in content trust modeling.

## 4 Trust modeling in geotagging applications

From the general trust modeling described in the previous section, we now shift the discussion to a more specific problem of geotagging the shared content and efficient propagation of such tags throughout the untagged content. In this section, we present and discuss several techniques for combatting noise and spam through trust modeling in social tagging systems. First, we introduce the model of a social tagging system. Then we present in details the five recent techniques for trust modeling that are suitable for geotagging and can be used in geotag propagation systems.

The model of a social tagging system [29] consists of *users* who interact with the system, *content* (resources or documents) which might be any piece of information (e.g., photos, videos, textual documents, or web pages) and *tags* which are descriptions assigned to the piece of the content by users. The action of associating a tag to a content by a user is usually referred to as *tag assignment* [25]. Depending on the system under consideration, a user can assign one or several tags to each type of content. Following notations are used in formal description of the trust models: $U$ is a set of users $u$, $D$ denotes a set of documents (content) $d$, $T$ is a set of tags $t$, and a set of tag assignments $p$ is denoted as $P \in U \times D \times T$.

The selected trust modeling techniques are summarized in Table 2 summarizes five trust modeling approaches, which we then describe in more details (in the same order as they are presented in the table). These methods are different in the targeted media content, for which the geotagging is intended, the application they are used in, and the required level of participation from the users of the geotagging system.

### *4.1 A coincidence-based model*

Koutrika *et al.* [22] were the first to explicitly discuss methods of tackling spamming activities in social tagging systems. The authors studied the impact of spamming through a framework for modeling social tagging systems and user tagging behav-

Table 2: Summary of five trust modeling techniques used for combatting noise and spam in social tagging systems.

| Reference | Media Content | Method | Dataset |
|---|---|---|---|
| Koutrika *et al.* [22] | bookmarks | a coincidence-based model for query-by-tag search, which estimates the level of agreement among different users in the system for a given tag | Delicious, real & simulated |
| Liu *et al.* [25] | bookmarks | an iterative approach to identify spam content by its information value extracted from the collaborative knowledge | Delicious, real |
| Xu *et al.* [45] | bookmarks | an iterative approach to compute the goodness of each tag with respect to a content and the authority scores of the users | MyWeb 2.0, real |
| Krestel and Chen [24] | bookmarks | a TrustRank-based approach using features which model tag co-occurance, content co-occurance and co-occurance of tag-content | BibSonomy, real |
| Ivanov *et al.* [16] | images | an approach based on the feedback from other users who agree or disagree with a tag associated with an image | Panoramio, real |

ior. They proposed a method for ranking content matching a tag based on taggers' reliability in social bookmarking service Delicious. Their coincidence-based model for query-by-tag search estimates the level of agreement among different users in the system for a given tag. A bookmark is ranked high if it is tagged correctly by many reliable users. A user is more reliable if his/her tags more often coincide with other users' tags.

In more formal way, the following calculations are performed:

$$c(u) = \sum_{d,t:\exists P(u,d,t)} \sum_{u_i \in U: u_i \neq u} |p : \exists P(u_i,d,t)| \tag{1}$$

$$score(d,t) = \frac{\sum_{u:\exists P(u,d,t)} c(u)}{\sum_{u \in U} c(u)} \tag{2}$$

$$trust^{Koutrika}(u) = \sum_{d,t:\exists P(u,d,t)} score(d,t) \tag{3}$$

where $c(u)$, coincidence factor of the user $u$, is the number of other users $u_i$ who assigned the same tag $t$ to the same document $d$ as the user $u$ did. Score of the document $d$ with respect to the tag $t$, denoted as $score(d,t)$, is calculated as a normalized value of $c$ over all users who assigned $t$ to $d$. Finally, a trust value of the user $u$, $trust^{Koutrika}(u)$, is the sum of $score(d,t)$ over all tag assignments by $u$.

Koutrika *et al.* performed a variety of evaluations of their trust model on controlled (simulated) dataset by populating a tagging system with different user tagging behavior models, including a good user, bad user, targeted attack model and several other models. Using controlled data, interesting scenarios that are not covered by real-world data could be explored. It was shown that spam in tag search

results using the coincidence-based model is ranked lower than in results generated by, e.g. a traditional occurance-based model, where content is ranked based on the number of posts that associate the content to the query tag.

### 4.2 A wisdom of crowds model

Liu *et al.* [25] proposed a simple but effective approach for detecting spam content in Delicious, by harvesting the wisdom of crowds. An information value of a bookmark is defined as the average number of times that each tag of the content is assigned by different users. A low information value of a bookmark indicates a divergence from crowds, which can be considered as a spam content. Furthermore, this method was extended to user trust modeling by aggregating the information values for each user.

All measures are defined as follows:

$$it(d,t) = \frac{|u : \exists P(u,d,t)|}{\sum_{t' \in T} |u : \exists P(u,d,t')|} \tag{4}$$

$$ic(u,d) = \frac{\sum_{t:\exists P(u,d,t)} it(d,t)}{|t : \exists P(u,d,t)|} \tag{5}$$

$$I(d) = \frac{|u : \exists P(u,d,.)|}{\sum_{d' \in D} |u : \exists P(u,d',.)|} \tag{6}$$

$$trust^{Liu}(u) = \sum_{d:\exists P(u,d,.)} I(d) \cdot ic(u,d) \tag{7}$$

where $it(d,t)$ represents the tag's $t$ tagging information value with respect to document $d$, and $ic(u,d)$ is the information value of the content (document) $d$ with respect to user $u$. The importance of the document $d$ is defined with $I(d)$. Finally, a trust value of the user $u$, $trust^{Liu}(u)$, is calculated as the weighted average of the information value of the content tagged by user $u$, with the importance of the document as weight.

An interesting point is that, for the time being, Liu *et al.* collected the largest dataset for trust modeling by crawling Delicious [15]. This dataset had around 82 thousand users, 1.1 million tags, 9.3 million bookmarks and 17.4 million tag-bookmark associations.

### 4.3 An "authority" model based on goodness of tags

Xu *et al.* [45] introduced the concept of "authority" in social bookmarking systems, where they measured the goodness of each tag with respect to a content by the sum of the authority scores of the users who have assigned the tag to the content. Authority scores and goodness are iteratively updated by using HITS algorithm, which was initially used to rank web pages based on their linkage on the web [21].

Following measures are defined and iteratively calculated:

$$s_{i+1}(d,t) = \sum_{u:\exists P(u,d,t)} trust_i^{Xu}(u) \tag{8}$$

$$trust_i^{Xu}(u) = \frac{\sum_{d,t:\exists P(u,d,t)} s_i(d,t)}{|t : \exists P(u,.,t)|} \tag{9}$$

where $i \in [1 \dots Q]$, $s_i(d,t)$ is the goodness of each tag $t$ with respect to a content $d$, and $trust_i^{Xu}(u)$ represents a trust value (authority score) of the user $u$. Initial settings in this iterative approach are: $s_0(d,t) = 0, \forall t,d$ and $trust_0(u) = 1, \forall u$. The number of iterations is set to $Q = 100$.

### *4.4 A co-occurance model*

In contrast to the approach of Xu *et al.* [45], Krestel and Chen [24] iteratively updated scores for users only. The authors proposed to use a spam score propagation technique to propagate trust scores through a social graph in BibSonomy, where edges between nodes (in this case, users) indicate the number of common tags supplied by users, common content annotated by users and/or common tag-content pairs used by users. Starting from a manually assessed set of nodes labeled as spammers or legitimate users with the initial spam scores, a TrustRank metric is used to calculate and iteratively update spam scores for all users. TrustRank metric is previously introduced in Section 3.

All measures are calculated as follows:

$$\begin{aligned} W(u_1,u_2) = {} & |t : \exists P(u_1,.,t), P(u_2,.,t)| + |d : \exists P(u_1,d,.), P(u_2,d,.)| + \\ & + |d,t : \exists P(u_1,d,t), P(u_2,d,t)| \end{aligned} \tag{10}$$

$$Tr(u_1,u_2) = \frac{W(u_1,u_2)}{\sum_{v \in U} W(u_1,v)} \tag{11}$$

$$trust_i^{Krestel}(u) = \alpha \cdot \sum_{v \in U} Tr(u,v) \cdot trust_{i-1}^{Krestel}(v) - (1-\alpha) d(u) \tag{12}$$

where $i \in [1 \dots Q]$, $W(u_1,u_2)$ is the weight of the edge between users $u_1$ and $u_2$ in the social graph and $Tr(u_1,u_2)$ is the corresponding transition matrix. A trust value of the user $u$, $trust_i^{Krestel}(u)$, is iteratively calculated. Initial setting in this iterative approach is: $trust_0(u) = d(u), \forall u$, where $d(u)$ represents the trust values of the seed users. The number of iterations is set to $Q = 100$.

The approach of Krestel and Chen is more sophisticated than the approach of Xu *et al.* [45] in that multiple relationships, such as tag co-occurance, content co-occurance and tag-content co-occurance, can be taken into account, rather than considering only the tag-content pairs shared by users.

### *4.5 User reliability based model*

In this section, we describe our own approach for user trust modeling in image tagging, which was proposed in [16]. First, we evaluate the trust or reliability of users by making use of their past behavior in tagging. We want to distinguish between users who provide reliable geotags, and those who do not. After user evaluation and trust model creation, tags will be propagated to other photos in the database only if the user is trusted. Assuming that there are $L$ users who tag $M$ training images, a matrix $R_{i,u}$, $i \in [1,M]$ and $u \in [1,L]$, is defined as:

$$R_{i,u} = \begin{cases} 1, \text{ if user } u \text{ tags image } i \text{ correctly;} \\ 0, \text{ otherwise.} \end{cases} \tag{13}$$

The process of comparing the propagated tags to ground truth tags can be done automatically using tag similarity measures, for example WordNet [5] or Google distance [7] measures. Nevertheless, we considered only manually defined ground truth for our experiments.

A trust value for user $u$, $trust^{Ivanov}(u)$, is computed as the percentage of the correctly tagged images among all images tagged by user $u$:

$$trust^{Ivanov}(u) = \frac{\sum_{i=1}^{M} R_{i,u}}{M}. \tag{14}$$

Only tags from users who are trusted are propagated to other photos in the dataset. In other words, if the user trust value $trust^{Ivanov}(u)$, exceeds a predefined threshold $\hat{T}$, then all his/her tags are propagated. Otherwise, none of his/her tags are propagated.

In this approach, ground truth data are used for the estimation of the user trust value. However, for a practical photo sharing system, such as Panoramio, it is not necessary to collect ground truth data since user feedback can replace them. The main idea is that users evaluate tagged images by assigning a true or a false flag to the tag associated with an image. If the user assigns a false flag, then he/she needs to suggest a correct tag for the image. The more misplacements a user has, the more untrusted he/she is. By applying this method, spammers and unreliable users can be efficiently detected and eliminated. Therefore, the user trust value is calculated as the ratio between the number of true flags and all associated flags over all images tagged by that user. The number of misplacements in Panoramio is analogous to the number of wrongly tagged images in our approach.

In case that a spammer attacks the system, other users can collaboratively eliminate the spammer. First, the spammer wants to make other users untrusted, so he/she assigns many false flags to the tags given by the trusted users and sets new wrong tags to these images. In this way, the spammer becomes trusted. Then, other users correct the tags given by the spammer, so that the spammer becomes untrusted and all of his/her feedbacks in the form of flags are not considered in the whole system. Finally, previously trusted users, who were untrusted due to spammer attack, recover their status. Following this scenario, the user trust value can be constructed by making use of the feedbacks from other users who agree or disagree with the

tagged location. However, due to the lack of a suitable dataset which provides user feedback, the evaluation of the user trust scenario is based on the simulation of the social network environment as described in details in [16].

## 5 An automated geotagging propagation system

Based on the user reliability trust modeling described in Section 4.5, we built the solution for geotag propagation between images. The main innovation of such system is the combination of object duplicate detection and user trust modeling for accurate and reliable geotag propagation. The system architecture has been proposed previously in [16] and is illustrated here in Figure 1. It contains three functional modules, each of which has a specific task: object duplicate detection, tag propagation, and user trust modeling. As the focus of this chapter is on trust modeling, the object duplicate detection [41] and tag propagation [16] modules are only summarized briefly below.

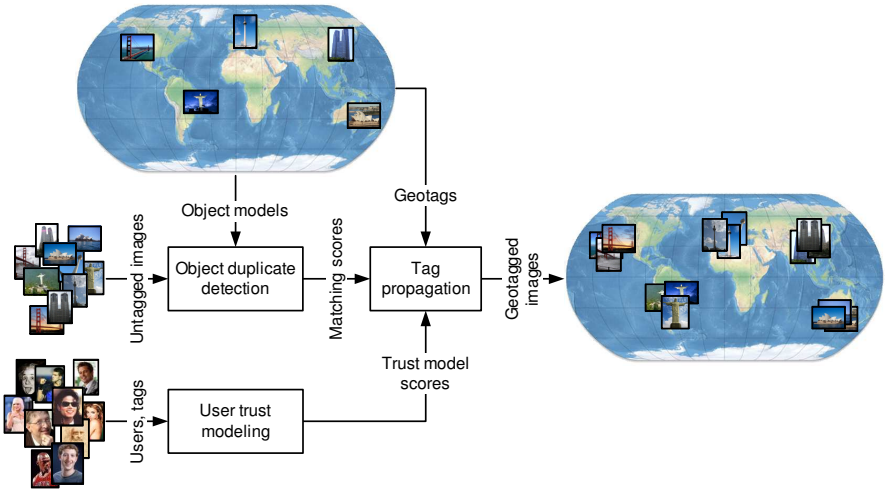


Fig. 1: Overview of the system for geotag propagation in images. The object duplicate detection is trained with a small set of images with associated geotags. The created object (landmark) models are matched against untagged images. The resulting matching scores serve as an input to the tag propagation module, which propagates the corresponding tags to the untagged images. Given a user trust model, only the tags from reliable users are propagated.

The system takes a small set of training images with associated geotags to create the corresponding object (landmark) models. These object models are used to detect objects duplicated in a set of untagged images. As a result, matching scores between

the models and the images are obtained. According to the scores, the tag propagation module makes decisions about which geotags should be propagated to the individual images. Given a user trust model which describes the tagging reliability of each user, only the tags from the users who are trusted are propagated to the photos in the dataset.

### *5.1 Object duplicate detection*

The goal of the object duplicate detection module is to detect the presence of a target object in an image based on an object model created from training images. Duplicate objects may vary from their perspective, have different size, or be modified versions of the original objects after minor manipulations, as long as such manipulations do not change their identity. This is especially true for images related to travel, where tourists tend to take a lot of photos from different distances and viewpoints around a famous landmark. The basic idea of applying object duplicate detection for geotag propagation is that travel images typically depict distinctive landmarks (buildings, mountains, bridges, etc.), which can be considered as object duplicates.

Training is performed as follows: given a set of images, features are extracted and a spatial graph model describing the object, i.e. landmark, is created for each of the landmarks. In our case, one training image per landmark is used to create a graph model. First, regions of interest (ROIs) in an image are extracted using the Hessian affine detector [31] and each of these regions is described using SIFT features [26]. These features are robust to arbitrary changes in viewpoints. Then, hierarchical k-means clustering [33] is applied to the features, to group them based on their similarity. The result of the hierarchical clustering is used for the fast approximation of the nearest neighbor search, to efficiently resolve feature matching in the test phase. Finally, a spatial graph model is constructed to improve the accuracy of the feature matching with a test image. The graph model considers the scale, orientation, position, and neighborhood of features. The nodes of the graph are the features of the training images. The edges of the graph connect features with their spatial nearest neighbors. The attributes of edges are the distance and orientation of the neighbors. These attributes are important for the matching step in the test phase.

To detect the presence of the landmark within a test image, the features are extracted from the image in the same way described above. These features are matched to those in the graph model derived from the training images. Feature matching is performed using a one-to-one nearest neighbor matching, where the hierarchical clustering is used to efficiently resolve the nearest neighbor search. Considering only matched features and their positions, a spatial graph model of the query image is constructed in the same way described in the training phase. Then, graph matching is applied between two graph models to identify the local correspondences between regions in the training and the test image. Finally, for the global object matching and matching score computation, the general Hough transform [3] is applied on the

nodes of the matched graph. The matching scores represent the pair-wise comparison of training and test images.

More details about the proposed object duplicate detection approach are presented in [41, 40].

## *5.2 Tag propagation*

The goal of the tag propagation module is to propagate the geotags from the tagged to the untagged images according to the matching scores, provided by the object duplicate detection module. As a result, labels from the training set are propagated to the same object found in the test set.

The geographical metadata (geotags) embedded in the image file usually consist of location names and/or GPS coordinates, but may also include altitude, viewpoint, etc. Two of the most commonly used metadata formats for image files are EXIF and IPTC. In this paper, we consider the existing IPTC schema and introduce a hierarchical order for a subset of the available geotags, namely: city (name of the city where image was taken) and sublocation (area or name of the landmark), for example, Paris (Eiffel Tower) and Budapest (Parliament).

It was shown in [16] that tag propagation module supports two application scenarios: closed and open set problem. In the *closed set problem*, each test image is assumed to correspond to exactly one of the known (trained) landmarks. Therefore, the image gets assigned to the most probable trained landmark, based on the matching scores provided by the object duplicate detection module, and the corresponding tag is propagated to the test image. However, in the *open set problem* the test picture may correspond to an unknown landmark, and then either one geotag or none will be propagated to the test image.

## *5.3 Experiments and results*

In Ivanov *et al.* [16], we argued that our approach to user trust modeling requires a small number of images to learn models for geotag propagation. We evaluated the approach on a dataset of 1320 images of famous landmarks (such as Bird's Nest Stadium, Sagrada Familia, Reichstag, Golden Gate Bridge and Eiffel Tower) downloaded from Google Images, Flickr and Wikipedia. All landmarks were split into different groups, such as castles, churches, bridges, towers/statues, stadiums and ground structure. More details on the dataset are available in [16].

At first, we evaluated the automatic geotag propagation algorithm without including users and their mistakes in the annotation process. We showed that the object duplicated detection approach performs the best for the landmarks like castles or other buildings which have more salient regions, while landmarks that belong to tower and stadium groups perform worse because these landmarks do not have

enough discriminative features or due to large variety of viewpoints. The accuracy measured as an average recognition rate across all landmarks is 71%. The recognition errors are solely caused by the object duplicate detection.

Then, the users are introduced in the system in order to simulate a real social network and evaluate the algorithm, which combines object duplicated detection with user trust modeling. The methodology used in this experiment is to extract a sub network from a large social network, in a way that every user in this sub network annotates every landmark in the subset of the dataset. In our experiments, each of 47 users is asked to annotate 66 images. Upon this sub network, we build up an automatic propagation system in order to decrease the annotation time and increase the accuracy of the system. In this case, our system relies on user-provided tags, which may sometimes be spam annotations given on purpose or wrong tags given by mistake. The users are evaluated and only tags from users whose trust model exceeds a predefined threshold are propagated to other images of the database.

Fig. 2 shows the accuracy of the system and the percentage of the number of propagated tags versus the threshold set for the user trust modeling. The optimal accuracy using object duplicate detection for geotag propagation is 71%. However, in this scenario the error of the user tagging step leads to a decrease of the performance. This error is caused by wrong tags given by the users. The optimal results can be reached if we set the threshold $\hat{T}$ to a high value, but then the number of propagated tags becomes very low. On the other hand, when the threshold is low, more tags are propagated. These curves could be used to determine an appropriate threshold for the proposed user trust model. The higher the threshold for the user trust model is, the more reliable the geotag propagation system is. At a threshold of 0, the accuracy of the system is equal to that without a user trust model, since all the user tags are propagated. In this case the accuracy of the system is 34%. The figure also shows the average user trust value of 52%, which is the same as the accuracy when the users tag all the images in the dataset (1320 images) and not only 66 images. Therefore, if we consider a large social network system where landmarks and users are selected in a way that each landmark is annotated by each user, our system shows that the best performance is achieved by choosing the most trusted user and propagating his/her annotations through the whole database of images. More precisely, in our dataset, the user annotates $1320/66 = 20$ times less images and the performance of the system (recognition rate) increases from value of 52% to 65%. As a conclusion, by using the proposed model less manual tagging is needed, while the performance of the system increases significantly.

Fig. 3 illustrates the relationship between the accuracy of the tag propagation system and the number of propagated tags by plotting them against each other. The maximum number of propagated tags can be much higher than the number of images, since several tags can be assigned to an image by different users. The black marker indicates the average tagging accuracy of the system without the user trust model and tag propagation presented in this chapter. In this case, if users tag $47 \times 66 = 3102$ photos (47 users in our experiments and each of them tags 66 images), the average accuracy of 52% can be achieved. This is equivalent to what we currently have in Flickr or Panoramio, where users simply tag photos independently

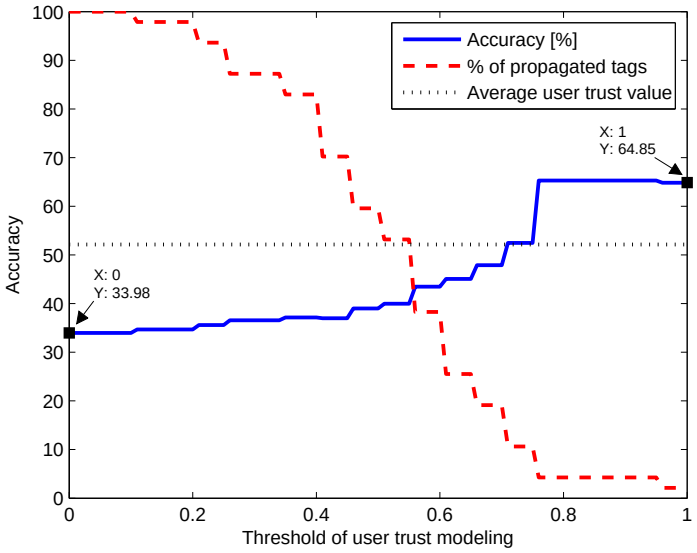


Fig. 2: The recognition rate of the geotag propagation system and the percentage of the propagated tags versus the threshold $\hat{T}$ for the user trust modeling.

and these tags are not being propagated. However, by introducing a user trust model and tag propagation into the system, we can improve the accuracy of the system and propagate more correct tags to untagged images in the dataset. This is depicted with the left part of the blue curve, which is above the dashed line, we can still propagate around 8000 tags from trusted users, while keeping accuracy higher than 52%.

## 6 Conclusion

In this chapter, we have presented different approaches for automatic geotagging and trust modeling in social tagging systems. The problem of having trustworthy geotags of the content is important in social networks, because of their increasing popularity as means of sharing interests and information. Especially photo sharing and tagging is becoming more and more popular. Among other tags, geotags in form of geographical locations provide efficient information for grouping or retrieving images. Since manual annotation of these tags is time consuming, automatic tag propagation based on visual similarity offers a very interestingly good solution.

The particular focus of this chapter is on the system for automatic geotag propagation by associating locations with distinctive landmarks and using object duplicate detection for tag propagation. The adopted graph based approach reliably

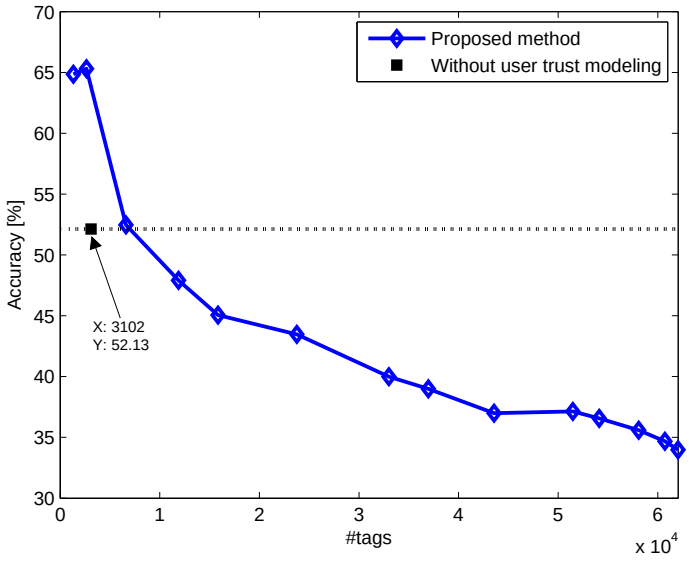


Fig. 3: The recognition rate of the geotag propagation system versus the number of the propagated tags.

establishes the correspondence between a small set of tagged images and a large set of untagged images. Based on these correspondences and a trust value of the model derived for each user, only reliable geotags are propagated, which leads to a decrease of tagging efforts. We have analyzed the influence of wrongly annotated tags, which causes even more wrongly propagated tags in the database. By considering user trust models the accuracy of the system could be considerably improved. In this way, the proposed user trust model can be generalized to photo sharing platforms such as Panoramio or Flickr.

Most of the current techniques for noise and spam reduction focus only on textual tag processing and user profile analysis, while visual features of multimedia content can also provide useful information about the relevance of the content and content-tag relationship. In the future, a promising research direction would be to combine multimedia content analysis with conventional tag processing and user profile analysis.

**Acknowledgements** This work was supported by the Swiss National Foundation for Scientific Research in the framework of NCCR Interactive Multimodal Information Management (IM2), the Swiss National Science Foundation Grant “Multimedia Security” (number 200020-113709), and partially supported by the European Network of Excellence PetaMedia (FP7/2007-2011).

## References


1. von Ahn, L., Blum, M., Hopper, N.J., Langford, J.: CAPTCHA: Using hard AI problems for security. In: Proc. Eurocrypt, pp. 294–311 (2003)
2. von Ahn, L., Maurer, B., Mcmillen, C., Abraham, D., Blum, M.: reCAPTCHA: Human-based character recognition via web security measures. Science **321**(5895), 1465–1468 (2008)
3. Ballard, D.H.: Generalizing the Hough transform to detect arbitrary shapes. Pattern Recogn. **13**(2), 111–122 (1981)
4. Barnett, E.: 3.4 billion photographs on Google+ in 100 days (2011). Available at: `http://www.telegraph.co.uk/technology/google/8838196/3.4-billion-photographs-on-Google-in-100-days.html`
5. Budanitsky, A., Hirst, G.: Evaluating WordNet-based measures of lexical semantic relatedness. Comput. Linguist. **32**(1), 13–47 (2006)
6. Cao, L., Luo, J., Huang, T.S.: Annotating photo collections by label propagation according to multiple similarity cues. In: Proc. ACM MM, pp. 121–130 (2008)
7. Cilibrasi, R.L., Vitanyi, P.M.B.: The Google similarity distance. IEEE Trans. on Knowl. and Data Eng. **19**(3), 370–383 (2007)
8. Fetterly, D., Manasse, M., Najork, M.: Spam, damn spam, and statistics: Using statistical analysis to locate spam web pages. In: Proc. ACM WebDB, pp. 1–6 (2004)
9. Gammeter, S., Bossard, L., Quack, T., van Gool, L.: I know what you did last summer: Object level auto-annotation of holiday snaps. In: Proc. ICCV), pp. 614–621 (2009)
10. Gyongyi, Z., Garcia-Molina, H., Pedersen, J.: Combating web spam with TrustRank. In: Proc. VLDB, pp. 576–587 (2004)
11. Hays, J., Efros, A.A.: IM2GPS: Estimating geographic information from a single image. In: Proc. IEEE CVPR, pp. 1–8 (2008)
12. Heymann, P., Koutrika, G., Garcia-Molina, H.: Fighting spam on social web sites: A survey of approaches and future challenges. IEEE Internet Comput. **11**(6), 36–45 (2007)
13. Hollenstein, L., Purves, R.: Exploring place through user-generated content: using Flickr to describe city cores. Journ. of Spatial Information Science pp. 1–29 (2010)
14. International Press Telecommunications Council: IPTC Photo Metadata Standard, IPTC Core 1.1 and IPTC Extension 1.1. Tech. rep. (2009)
15. Ivanov, I., Vajda, P., Lee, J.S., Ebrahimi, T.: In tags we trust: Trust modeling in social tagging of multimedia content. IEEE Signal Proc. Mag. **29**(2) (2012)
16. Ivanov, I., Vajda, P., Lee, J.S., Goldmann, L., Ebrahimi, T.: Geotag propagation in social networks based on user trust model. MTAP **56**(1), 155–177 (2012)
17. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decision Support Syst. **43**(2), 618–644 (2007)
18. Kennedy, L.S., Chang, S.F., Kozintsev, I.V.: To search or to label?: Predicting the performance of search-based automatic image classifiers. In: Proc. ACM MIR, pp. 249–258 (2006)
19. Kennedy, L.S., Naaman, M.: Generating diverse and representative image search results for landmarks. In: Proc. WWW, pp. 297–306 (2008)
20. Kessler, S.: Mashable Infographics – Facebook Photos by the Numbers (2011). Available at: `http://www.mashable.com/2011/02/14/facebook-photo-infographic`
21. Kleinberg, J.M.: Authoritative sources in a hyperlinked environment. JACM **46**(5), 604–632 (1999)
22. Koutrika, G., Effendi, F.A., Gyöngyi, Z., Heymann, P., Garcia-Molina, H.: Combating spam in tagging systems: An evaluation. ACM TWEB **2**(4), 22:1–22:34 (2008)
23. Krause, B., Schmitz, C., Hotho, A., G., S.: The anti-social tagger: Detecting spam in social bookmarking systems. In: Proc. ACM AIRWeb, pp. 61–68 (2008)
24. Krestel, R., Chen, L.: Using co-occurence of tags and resources to identify spammers. In: Proc. ECML PKDD, pp. 38–46 (2008)
25. Liu, K., Fang, B., Zhang, Y.: Detecting tag spam in social tagging systems with collaborative knowledge. In: Proc. IEEE FSKD, pp. 427–431 (2009)

26. Lowe, D.G.: Distinctive image features from scale-invariant keypoints. International Journal of Computer Vision **60**(2), 91–110 (2004)
27. Luo, J., Joshi, D., Yu, J., Gallagher, A.: Geotagging in multimedia and computer vision - a survey. MTAP **51**(1), 187–211 (2011)
28. Markines, B., Cattuto, C., Menczer, F.: Social spam detection. In: Proc. ACM AIRWeb, pp. 41–48 (2009)
29. Marlow, C., Naaman, M., Boyd, D., Davis, M.: Ht06, tagging paper, taxonomy, flickr, academic article, to read. In: Proc. ACM HT, pp. 31–40 (2006)
30. Marti, S., Garcia-Molina, H.: Taxonomy of trust: Categorizing P2P reputation systems. Comput. Netw. **50**(4), 472–484 (2006)
31. Mikolajczyk, K., Schmid, C.: An affine invariant interest point detector. In: Proc. ECCV, pp. 128–142 (2002)
32. Mori, G., Malik, J.: Recognizing objects in adversarial clutter: Breaking a visual CAPTCHA. In: Proc. IEEE CVPR, pp. I–134 – I–141 (2003)
33. Nister, D., Stewenius, H.: Robust scalable recognition with a vocabulary tree. In: Proc. IEEE CVPR, pp. 2161–2168 (2006)
34. Noll, M.G., Yeung, C.A., Gibbins, N., Meinel, C., Shadbolt, N.: Telling experts from spammers: Expertise ranking in folksonomies. In: Proc. ACM SIGIR, pp. 612–619 (2009)
35. Parr, B.: Mashable Infographics – Facebook by the Numbers (2011). Available at: `http://www.mashable.com/2011/10/21/facebook-infographic`
36. Quack, T., Leibe, B., Van Gool, L.: World-scale mining of objects and events from community photo collections. In: Proc. IEEE CIVR, pp. 47–56 (2008)
37. Sahami, M., Dumais, S., Heckerman, D., Horvitz, E.: A Bayesian approach to filtering junk e-mail. Tech. Rep. WS-98-05 (1998)
38. Technical Standardization Committee on AV & IT Storage Systems and Equipment: Exchangeable image file format for digital still cameras: Exif Version 2.2. Tech. Rep. JEITA CP-3451 (2002)
39. Thomason, A.: Blog spam: A review. In: Proc. CEAS (2007)
40. Vajda, P., Goldmann, L., Ebrahimi, T.: Analysis of the limits of graph-based object duplicate detection. In: Proc. Symposium on Multimedia, pp. 600–605 (2009)
41. Vajda, P., Ivanov, I., Goldmann, L., Lee, J.S., Ebrahimi, T.: Robust duplicate detection of 2D and 3D objects. Int. Journ. of Multimedia Data Engineering and Management **1**(3), 19–40 (2010)
42. Whitby, A., Jøsang, A., Indulska, J.: Filtering out unfair ratings in bayesian reputation systems. In: Proc. IEEE AAMAS, pp. 106–117 (2004)
43. Wikimedia Foundation Inc.: Wikipedia – Flickr (2012). Available at: `http://en.wikipedia.org/wiki/Flickr`
44. Wu, C.T., Cheng, K.T., Zhu, Q., Wu, Y.L.: Using visual features for anti-spam filtering. In: Proc. IEEE ICIP, vol. 3, pp. III – 509–512 (2005)
45. Xu, Z., Fu, Y., Mao, J., Su, D.: Towards the semantic web: Collaborative tag suggestions. In: Proc. ACM WWW (2006)
46. Yahoo! Inc.: Flickr – Tags (2011). Available at: `http://www.flickr.com/help/tags`
47. Yang, Y., Sun, Y.L., Kay, S., Yang, Q.: Defending online reputation systems against collaborative unfair raters through signal modeling and trust. In: Proc. ACM SAC, pp. 1308–1315 (2009)
48. Zheng, Y.T., Zhao, M., Song, Y., Adam, H., Buddemeier, U., Bissacco, A., Brucher, F., Chua, T., Neven, H.: Tour the World: Building a web-scale landmark recognition engine. In: Proc. IEEE CVPR, pp. 1085–1092 (2009)