Proceedings of the Seventh International AAAI Conference on Weblogs and Social Media

Leveraging Browsing Patterns for Topic Discovery and Photostream Recommendation

Luca Chiarandini
luca.chiarandini@upf.edu
Yahoo! Research, and
Web Research Group
Universitat Pompeu Fabra
Barcelona, Spain

Przemyslaw A. Grabowicz
pms@ifisc.uib-csic.e
IFISC (CSIC-UIB)
Palma de Mallorca, Spain

Michele Trevisiol
trevi@yahoo-inc.com
Yahoo! Research, and
Web Research Group
Universitat Pompeu Fabra
Barcelona, Spain

Alejandro Jaimes
ajaimes@yahoo-inc.com
Yahoo! Research
Barcelona, Spain

Abstract

In photo-sharing websites and in social networks, photographs are most often browsed as a sequence: users who view a photo are likely to click on those that follow. The sequences of photos (which we call *photostreams*), as opposed to individual images, can therefore be considered to be very important content units in their own right. In spite of their importance, those sequences have received little attention even though they are at the core of how people consume image content. In this paper, we focus on photostreams. First, we perform an analysis of a large dataset of user logs containing several million pageviews, examining navigation patterns between photostreams. Based on observations from the analysis, we build a stream transition graph to analyze common stream topic transitions (e.g., users often view "train" photostreams followed by "firetruck" photostreams). We then implement two stream recommendation algorithms, based on collaborative filtering and on photo tags, and report the results of a user study involving 40 participants. Our analysis yields interesting insights into how people navigate between photostreams, while the results of the user study provide useful feedback for evaluating the performance and characteristics of the recommendation algorithms.

1 Introduction

Social media platforms such as Flickr provide a wide range of functionalities and different ways to share and view content. Given the sequential nature of browsing photographs, it is common for people to share and view images in sequences, whether the photos are arranged in galleries, slideshows, or in groups. In Flickr, in particular, photos uploaded by a user to his account are placed in a "photostream", which in essence is a sequence of photos. Although there are many ways to reach individual photographs, such sequences constitute a fundamental part of the interaction. In the rest of the paper we will refer to such sequences as *photostreams* (or simply streams).

Furthermore, navigation across sequential units of content is present in other fields of social media, e.g. social network, music streaming and microblogging platforms. In popular social networks photos are organized in albums and can be viewed sequentially. Songs in music streaming services can be listened to one after another usually as a part of an album or a playlist. Posts in microblogging platforms are chronologically organized in independent blogs. Therefore, methods developed for photostreams could be adapted to other social media as well.

A key question for social media platforms, then, is how users navigate inside and between various photostreams. In particular, such photostreams may be considered not just as collections of images, but rather as fundamental units of content. On one hand, understanding how users navigate between specific photostreams is crucial in designing interfaces and algorithms that improve user experience, by providing the right content in the right places. On the other hand, analyzing the semantic categories of such streams can also provide important insights on general topics of interest. In addition, investigating how users transition between photostreams allows us to understand how topics may be related.

In this paper, we treat photostreams as content units and analyze a large sample of navigation logs to gain insights into how users navigate between different photostreams. More specifically, we examine user navigation logs containing several millions pageviews in order to create a photostream transition graph to analyze frequent topic transitions (e.g. users often view "train" followed by "firetruck" photostreams). We implement two photostream recommender systems: a collaborative filtering approach based on transitions between photostreams and a content-based method using tag-similarity of photostreams. Finally, we report the results of a user study involving 40 participants to explore the fundamentals for creation of an effective recommender system in a large social media platform.

Our main contributions can be summarized as follows:

- We perform a large-scale analysis of photostream browsing patterns, providing insights into frequent transitions between different topics in image browsing.
- Based on photostreams, we propose collaborative filtering and content-based recommender systems, that are computationally inexpensive due to their aggregate nature.
- By means of a user study, we show that the collaborative filtering method, based on transitions between photostreams, provides more novel content than the tag-based recommender system.

To the best of our our knowledge, this is the first study which analyses photostream browsing as opposed to photo browsing. This is also the first time the problem of photostream recommendation is addressed, in particular by leveraging the navigation patterns of a large number of users.

The rest of this paper is organized as follows. First, we discuss related work in Section 2 and the dataset used in Section 3. Then, in Sections 4 and 5, we present results of the analysis of user browsing patterns and the transitions between photostreams. We follow up with the description of our recommender systems in Section 6. In Section 7 we present results of a pilot user study and of the main user study comparing the recommender systems. Finally, we discuss our findings and conclude the paper.

2 Related Work

Flickr has been used extensively in research, in large part because it provides a public API that has allowed researchers to easily obtain large data-sets. The scope of research on Flickr is rather broad and includes analyses of the Flickr browsing patterns, development of new interfaces, image recommendations and many others. In this section we provide a brief overview of some of the Flickr and non-Flickr works that we consider to be most related.

Several studies investigate image browsing patterns of users. Lerman *et al.* (2006) jointly use information about tags, social network, photo groups and photo views to understand how different people browse photos. Lipczak et al. (2013) performed a similar study in Flickr also considering user behavior. However, they focused their attention on user explicit actions, in particular on favorites. Srikant and Yang (2001) use implicit information extracted from server logs to improve the design of a website. In particular, the authors analyze the server logs in order to suggest modifications to the website link structure, to make content easier to find for the users. Other work extracts user sessions from Flickr server logs to identify navigation patterns (Chiarandini, Trevisiol, and Jaimes 2012), rank photos (Trevisiol et al. 2012) or create new photo browsing interfaces (Chiarandini and Jaimes 2012). None of the above, however, take into account photostreams, nor implement a recommender system.

Various interfaces have been considered for image browsing. Fan et al. (2009) describe *JustClick*, which recommends images via interactive exploratory search. They build a topic network based on Flickr tags, and propose an interactive interface that allows the user to express a query by selecting images. They perform experiments on a big Flickr dataset of 1.5 billion images with 4000 different topics. Xu et al. (2009) present an innovative visual search interface based on topic clustering. Given the query and the results from a search engine, latent topics are detected and clustered and then the clusters are shown in an intuitive layout. Ren and Calic (2009) present an interactive interface for browsing of large-scale image collections. Their system is based on two main parts, an image clustering module and an interface generation component in order to retrieve the images in a more efficiently way. Strong et al. (2010) presented an approach for browsing images based on conceptual and visual similarity, with the main benefit being that the displayed images are grouped together. Zavesky et al. (2008) proposed a new framework called *Visual Island*, a novel organization algorithm for interactively displaying results. The aim is to organize the images in order to improve human comprehensibility and reduce required inspection time. We propose a recommender system that is well-integrated in the standard photo-browsing interface and uses only anonymous browsing and content data.

Multiple studies investigate similarity of photos in photo collections, e.g. (Cao, Luo, and Huang 2008; Yu et al. 2011), where the goal is to organize the images that present similar visual or textual information in groups of the same topic. We use user-generated photostreams instead and split them into batches of photos. In the case of Flickr, Gozali et al. (2012) used a hidden Markov model to split photostreams into groups of similar photos and evaluate different layouts to represent them. In this paper we use a simpler method which is computationally cheap. However, more advanced methods could enhance our recommender systems and can be considered for future work.

A number of reviews summarize the state of the art in the field of recommender systems (Adomavicius and Tuzhilin 2005; Ricci, Rokach, and Shapira 2011; Herlocker et al. 2004). A main classification divides recommender systems into content-based and collaborative filtering (Adomavicius and Tuzhilin 2005; Ricci, Rokach, and Shapira 2011). Both methods have several limitations. Content-based systems tend to recommend items which are too similar and therefore not interesting to the user, namely overspecialized. Collaborative methods require a critical mass of user traces in order to provide meaningful recommendations. This is known as a data sparsity problem and it is particularly challenging in systems with numerous small content pieces like social media platforms, where the number of visits per photo tends to be low. We use both methods to recommend photostreams. Mobasher et al. (Mobasher et al. 2002) proposed a system based on aggregate usage profiles consisting of clustered user transactions. One can draw a parallel to our system: a user browsing a part of a photostream may be seen as a transaction and a cluster of such transactions can be interpreted as a whole photostream. A natural difference is that photostreams are explicit content and structured units. Furthermore Herlocker et al. (2004) list the recommendation of item sequences as one of the possible goals for a recommender system. Nevertheless, this has not attracted much attention of researchers (Cacheda et al. 2011). In this paper we recommend sequences of photos belonging to recommended photostreams in a two-level recommender system.

3 Dataset

For the purpose of this study, we took a sample of the pageviews of more than 10 million anonymous users from 2011. Since Flickr allows users to set specific pages to be private, in our analysis we considered only public pages. All of the data processing was anonymous and performed in aggregate.

3.1 Pageview Filtering and Data Selection

In order to obtain a coherent dataset in terms of both timezone and activity, we focused on users who are located in the United States. We then removed traffic derived from Web crawlers by preserving only the entries corresponding to a well-known browser (e.g., Firefox, Chrome, etc.). In spite of this filtering some users have a very large number of pageviews. The frequency, however, suggests that such server requests could not have been made by humans, but instead were done automatically for malicious crawling. We therefore set a maximum threshold on the total number of pageviews per user. This heuristic filters out around 1% of the total number of users.

3.2 Session Identification and Characteristics

We group pageviews of a user into *sessions*. Sessions are linear chronological sequences of pageviews over a specific period of time. We split the activity of a single user into different sessions when either of the following two conditions holds:

- *Timeout:* the time difference between two consecutive pageviews is longer than 25 minutes.
- *External url:* if a user leaves the Flickr site, and then returns back, the current session ends even if previous visits are within the 25 minute threshold (i.e. we make the assumption that if a user is viewing a page on Flickr and visits another domain, then the session ends).

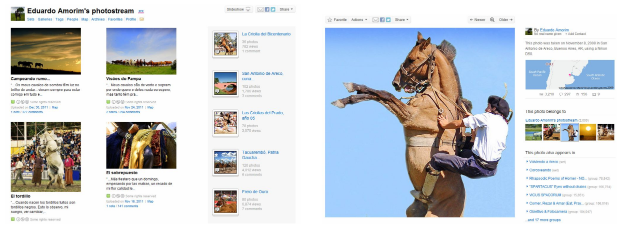
We use the resulting filtered dataset and sessions in the rest of the analysis.

3.3 Tags of Photos

Users of Flickr can create and attach tags to their photos. We gathered tags of all public photos in the dataset from Flickr. We preprocess these tags by discarding the ones that belong to a multi-lingual stop-word dictionary obtaining around 5 million distinct tags.

4 Analysis

In this section we define the main concepts of our study, present statistics on how users browse within sessions and how they transition between photostreams.



(a) grid view (b) photo-focused view

Figure 1: General types of stream views in Flickr.[1]

4.1 Photostream Browsing

Photos in Flickr are organized in photostreams. Each photo in Flickr belongs to a photostream of the owner, but it can belong to other streams of photos as well: groups, sets, galleries, favorites. Apart from favorites, all of these photostreams are either chosen or created by the owner of the photo. Users always view and browse photos in the context of a particular photostream.

There are two main ways of viewing photostreams: *a*) *grid view*, i.e. grid of photos from the stream (see Fig.1a), and *b*) *photo-focused view* i.e. a single zoomed-in photo with a possibility of browsing neighboring photos (see Fig.1b). Although Flickr allows different variations of grid views, they share a common feature, namely that they show several pictures from the browsed stream at a glance. The photo-focused view is the same for all the streams: it shows a large selected photo and, on the right side of it, thumbnails of 4 neighboring photos from the stream are presented, which the user can switch to by clicking on them. This way one can change the focus from the current photo to another one from the currently browsed stream. Below the thumbnails a list of all photostreams that the photo belongs to is shown in the form of hyperlinks, as visible in Fig. 1b.

One can expect that users first enter the grid view of a photostream, and then select one of the photos they like and see it in a photo-focused view. Then, they can continue on browsing other photos from this photostream. The grid view may be used for purposes which seem less involving to the user, e.g. quick browsing many photos from a stream, having an overview of a stream or seeking interesting content. Photo-focused views give the user options of performing many different actions in reference to the photo, e.g. he or she can comment on the photo, favorite it, download it, see it in different sizes or in a light-box setting.

For the purpose of the study, we define a *stream-browsing sequence* as an uninterrupted chronological sequence of pageviews that contains at least one photo-focused view and an indefinite number of grid views of one particular photostream (schematic examples are shown in Fig. 2). Each browsing session can consist of a number of stream-browsing sequences.

The Flickr logs in our dataset contain a total of 264 million pageviews, out of which a considerable part form

[1]Sample Flickr pages from the user http://www.flickr.com/photos/bombeador.

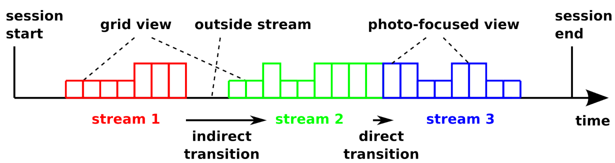


Figure 2: Diagram of possible transitions between streams.

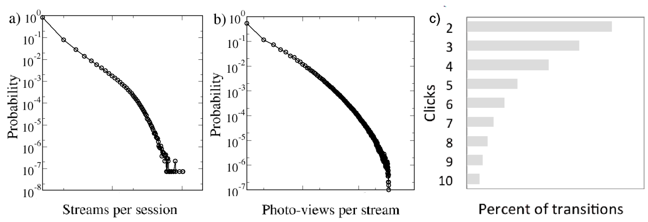


Figure 3: Distributions of number of unique streams per session (a) and a number of photo-focused views per each unique stream in a session (b), in log-log scale. The number of clicks between different streams is shown in (c).

stream-browsing sequences. On average, each sequence consists of 8 pageviews, among which there are photo-focused views and grid views of the photostream. Distributions of both the number of distinct streams viewed per session (Fig.3a), and the number of photo-focused views per stream (Fig.3b), have a heavy-tail showing high variability in user browsing patterns.

4.2 Transitions Between Streams

In the previous subsection we have shown that a large portion of all pageviews corresponds to sequential browsing of photos inside of photostreams. In this context an interesting question to ask is how users switch between streams.

We distinguish two types of transitions (Fig. 2): *a*) *direct transitions*, which happen when the user is in photo-focused view of the stream and chooses one of the listed streams to the right of the photo, as in Fig. 1b, and *b*) *indirect transitions*, in which the user leaves the photo-focused view and enters it again in a different stream after performing a number of clicks (e.g. watching grid views, searching, exploring users' profiles, etc.).

We define a transition from photostream i to j as a sequence of non-photo-focused pageviews from a photo-focused view inside stream i to another photo-focused view inside stream j. This definition implies directionality. One can estimate the number of clicks and actions performed during the transition by counting the number of pageviews between the photo-focused views of the two streams and summing one. Direct transitions only require one click, whereas indirect transitions require more than one action.

In total we have identified 3.6 million transitions between photostreams. Indirect transitions achieved within 2 clicks cover a large portion of all transitions, as shown in Fig.3c. However, even more transitions happen after more than 5 clicks, so many users, before reaching another picture in a photostream pass through many non photo-focused pageviews. Moreover, direct transitions happen much less often than indirect transitions. In the present Flickr interface, photostreams which are reachable from the currently browsed stream with just 1 click are the ones that the displayed photo belongs to. Moreover, in Flickr, only the names of these photostreams are presented, with no thumbnails of pictures shown, which may negatively impact the number of direct transitions between streams.

graph	nodes	degree	strength
full	1530875	4.23	4.68
lcc	972047	5.80	6.46

Table 1: General stats of stream transition graph and its large connected component (lcc).

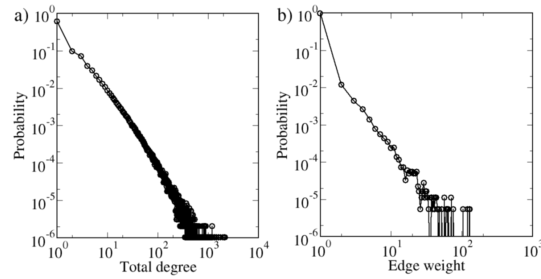


Figure 4: Distributions of degrees (a) and edge weights (b) in the graph of transitions.

4.3 Discussion

Almost half of all pageviews in the dataset form stream-browsing sequences. Users tend to see multiple photos of a photostream either in the photo-focused view or in the grid view before leaving the stream. The vast majority of all transitions between photostreams take place over several clicks. These results suggest that a modified photo-focused interface facilitating direct transitions to other streams could be implemented. Moreover, a system recommending other photostreams within this interface could be an improvement.

5 Graph of Transitions Between Streams

In this section we study the transitions between streams in more detail. Our goal is to show that users tend to browse photos of a given topic and sometimes switch to another topic that is related but not obvious. Such observations play an important role in the design of interfaces and recommender systems.

5.1 General Description

We define the graph of transitions as follows. Each photostream is treated as a node in a network. Edges in the graph represent transitions between photostreams i and j and their weight is equal to the number of such transitions.

The resulting total number of nodes in the transition graph is over 1.5 million, with an average degree of 4.2, as stated in Table 1. The graph is therefore sparse. The average strength of nodes, defined as the sum of the weights of its outgoing and incoming edges is 5.8. The graph is characterized

cluster label	no. of streams	escape ratio	self tag-coh.	global tag-coh.
recent-photogr.	36260	0.03	0.003	0.003
portrait	35689	0.21	0.021	0.008
graffiti	20518	0.06	0.060	0.006
landscape	12073	0.21	0.009	0.005
lego	8015	0.03	0.059	0.006
virtual-reality	6001	0.07	0.030	0.004
public-libraries	5044	0.28	0.006	0.004
bikes	3809	0.14	0.009	0.003
cakes	3748	0.08	0.056	0.007
canon-portrait	3456	0.28	0.021	0.008

Table 2: Statistics of the large clusters of photostreams.

by typical heavy-tailed distributions of degrees (Fig.4a) and weights (Fig.4b). Many nodes of the graph belong to the largest connected component, which covers over 60% of all nodes in the network. Further analysis presented in this section is based on it. The largest connected component has similar characteristics to the whole network, with slightly higher average degree and average strength, as presented in Table 1.

5.2 Clusters of Frequently Co-viewed Streams

In order to investigate if users browse photostreams sharing similar features we first *cluster these streams* using a community detection algorithm. A priori, detected clusters consist of nodes between which connections are dense, therefore they consist of photostreams between which the transitions are frequent. Our goal is to test if clusters of streams share common features.

We used Infomap (Rosvall and Bergstrom 2008; 2011), a state of the art community detection algorithm for weighted and directed networks. This algorithm was found to be one of the best performing methods in a recent review (Fortunato 2010). Infomap detects hierarchical community structure, but for the purpose of this paper we analyze only the highest hierarchical level of communities. The number of clusters found by the algorithm at the top level is over 2000.

In order to illustrate the content of clusters covering a considerable portion of the network, we show properties of some of the largest clusters. A possible way to measure the quality of the detected cluster is by calculating the ratio $\epsilon_i = \frac{l_i^{ext}}{l_i^{ext}+l_i^{int}}$, where l_i^{ext} is the number of edges connecting nodes from the cluster i with external nodes from other clusters, and l_i^{int} is the number of edges connecting internal nodes from the cluster i. In this work we call it *escape ratio*, as in our context it measures likeliness of a user browsing inside of a stream from a particular cluster to escape from this cluster by switching to a stream from another cluster. Generally this ratio should be small for well-defined clusters, however it grows with the number of clusters and their size (Grabowicz et al. 2012). Values of escape ratio for the large clusters found in the largest connected component of the transition graph are shown in Table 2. The results vary on clusters but tend to be very low. Given that the largest cluster accounts for less than 4% of all streams, its escape ratio of 0.03 is much lower than escape ratio of 0.96 that could be expected in a random scenario.

In order to characterize the content of the clusters we aggregate all photo tags which belong to all the streams of each of the clusters. If a photo belongs to several photostreams in one cluster, then we count its tags multiple times. Using this method we create tag clouds for every cluster in the network. We present them in Fig.5, where we plot the 50 most frequent tags for each cluster. The size of each tag from a tag cloud is proportional to the number of its appearances in the cluster. The labels, stated in the figure underneath each of the tag clouds, are chosen manually.

As one can see in the tagclouds, most of them have quite a narrow focus, and only a few have a wide focus: recent-photography, portrait, landscape, public-libraries, canon-portrait. As a side note, the narrow focus of clusters could possibly arise from just a few streams with many tags. To test if this is not the case and to quantify narrowness of cluster topics we use a measure of similarity s_{ij} between streams i and j. We define it as cosine similarity of multidimensional vectors of tag-clouds s_{ij}, where each dimension is a tag and the length is the count in the tag cloud. For every cluster we measure average similarity of its member streams with a) other member streams from this cluster, and b) all streams. We call these averages, correspondingly, *self tag-similarity* and *global tag-similarity*. The former property measures how coherent are streams within a particular cluster, whereas the latter quantifies how coherent these streams are with respect to all streams. As one can see in Table 2 self tag-similarity is several times higher than global tag-similarity for most of the clusters, meaning that indeed streams belonging to the same cluster are similar in content. Moreover, the clusters with narrow focus obtain the best scores as their self tag-similarity is up to 10 times higher than their global tag-similarity. Therefore streams in the clusters tend to be of similar topics.

5.3 Transitions Between Clusters

Since clusters contain streams of similar topic, an interesting question to ask is between which clusters people switch most often. This can be answered by a creating a node in place of every cluster of streams and aggregating edges of all streams belonging to this cluster. In this manner we obtain a directed and weighted network of transitions between clusters from the network of transitions between streams. After the conversion there are self-loops in such a network, which we remove. This network is dominated, however, by the connections between large nodes. To account for the size effect of the nodes and to extract meaningful information about relations between clusters we take the following approach. In the random case, the expected number of connections from node i to node j, having an out-degree k_i^{out} and an in-degree k_j^{in}, is equal to $l_{ij}^{rand} = \frac{k_i^{out}k_j^{in}}{l}$ for a large total number of edges in the network l. If connections between clusters were spread randomly between nodes of known degrees then l_{ij}^{rand} would be expected to be the number of edges between particular nodes. To see which connections between clusters

(a) recent-photography (b) portrait (c) graffiti (d) landscape (e) lego

(f) virtual-reality (g) public-libraries (h) bikes (i) cakes (j) canon-portrait

Figure 5: Tag clouds for the large clusters of photostreams.

are the furthest from random configuration we calculate the ratio $a_{ij} = \frac{l_{ij}}{l_{ij}^{rand}}$ of the actual number of connections l_{ij} and expected l_{ij}^{rand}. We call a_{ij} the abundance ratio. If this ratio is larger than 1 then transitions from stream i to j are overrepresented, while if it is lower than 1 then they are underrepresented. We pick the parts of the network formed by edges with abundance ratio a_{ij} higher than 10 and actual weight l_{ij} also higher than 10. We present most of them in Fig. 6. Here we provide a short description of each of the examples:

(a) Clusters of fans of cars and machinery. From left to right in the figure: the first cluster seems to be on the verge of cars and photography, while the next one is more narrowly about cars, especially classic ones. Users from this cluster tend to switch between both to see photos of trains and railroads, as well as firetrucks.

(b) Event-orientated clusters. From down to up in the figure: photography of live music shows is related to the cluster of journaling, blogging and fisheye photography.

(c) Household-centered clusters. From left to right: clusters of cakes and vintage style, which incorporate elements from previous eras into modern fashion and style, are related to the cluster of sewing and fabrics, to dolls, and then handmade are related. Note that dolls and Disney/Disneyland are also related.

(d) Toys and military. From left to right: photography of lego constructions, mostly of star-wars, is related to army and military photography. This is quite interesting, and shows an interests from toys and plastic soldiers to real ones. The military cluster is related to natural disasters in which often the army and powerful natural forces are involved.

It is also possible to find underrepresented connections between clusters, and it would certainly be interesting to examine more clusters in detail.

5.4 Discussion

On one hand, low escape ratios and high tag-coherence of the clusters of streams show that indeed users browse topically-similar streams. On the other hand, examples of the transitions between the clusters show that the users also switch between streams which are further apart in the topical space, but are still related (e.g. trains and firetrucks, cake and sewing, lego and army). This implies interesting consequences for the design of new interfaces or recommender systems, e.g. the recommended photos should not be topically overspecialized, in order to leave the user possibility of exploration. We investigate it further in the following sections, comparing two recommender systems.

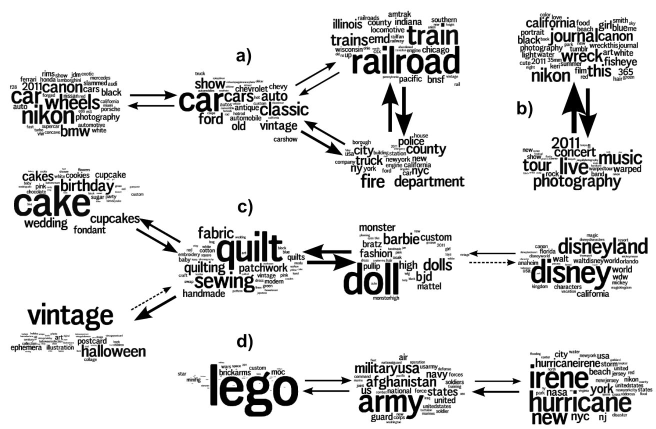


Figure 6: Interesting over-represented transitions between clusters of photostreams. Width of edges corresponds to the abundance ratios.

6 Recommendation of Photostreams

In this section we introduce two recommender systems that suggest photostreams (and the photos belonging to them). The first is based on collaborative-filtering (i.e. the transitions between photostreams from past user browsing sessions) and the second is based on content (i.e. the tags of images). Our aim is to compare them in terms of interestingness, relatedness and novelty of the recommended photos. Although we use well-known recommendation techniques, the novelty of our system lies in the use of photostreams as the main content unit. The systems consist of two levels. First, we recommend photostreams; second, we center on related photos from the photostreams. In this section, we describe the recommender systems in detail, while in the following section we evaluate them with a user study.

6.1 Two-Level Recommender System

This paper proposes two distinct recommender systems based on anonymized traffic data and content data. As such they do not require the user to log in. Each recommender system is built with a top-down approach in mind, meaning that it first analyzes high-level content units (photostreams) and then low-level content units (photos).
It consists of two levels:

1. *Photostream selection*: the system recommends a set of streams to the user based on the streams the user has seen until that moment.
2. *Centering on a photo inside a photostream*: the system chooses which part of the stream (i.e., consecutive photos) will be displayed to the user for each selected stream based on the last seen images.

Due to the fact that both levels function on significantly reduced data, this two-level system is computationally much less demanding than a system working at the level of all photos. For example, the first level effectively reduces the task of giving recommendations among billions of photos to the task of providing recommendations among just millions of streams. Moreover, the two-level design lets us circumvent the problem of data sparsity inherent in highly atomized social media platforms, which commonly store billions of images.

6.2 Photostream Selection

Here, we describe the first level of the two photostream recommendation algorithms. The task at this level is to rank photostreams based on the browsing history of the current user.

Collaborative Filtering Recommender. The first recommender presents to the user those streams which were most often co-viewed in the past sessions with the streams seen by the user in the current session. The algorithm computes the relevance of unseen streams to the current session, in the following way:

1. Given a stream i in the current session and an unseen stream j, we compute c_{ij}, i.e. the number of past sessions in which they appear together.
2. Then, we consider last $N_s = 5$ streams from the current session, and for each unseen stream j we compute its relevance to the current session: $c_j = \sum_{i=1}^{N_s} c_{ij}$.
3. Finally, the recommender selects the streams with the highest value of c_j.

It is difficult to estimate the coverage of this recommendation algorithm, because it is dependent on our limited sample of user traffic data. However, we point out here that over 99% of the photostreams have appeared with at least 3 other streams in the user sessions from our dataset.

Tag-based Recommender. The second recommender is based on similarity photo tags belonging to the streams. The algorithm takes as input the last stream seen by the user, and recommends those that are the most similar.

1. We use a standard information retrieval approach, where streams are documents and tags are words. We use Okapi BM25, as it has been shown to perform better in similar cases (Whissell and Clarke 2011). We create a tag vector for each stream.
2. We compute stream-to-stream similarity for each pair of the streams using cosine similarity.
3. We select recommended streams with the highest similarity.

We take into consideration only streams that contain at least 10 different tags, what gives us 1.8 million distinct streams. Additionally we note, that over 90% of these streams have a non-zero similarity with at least one other stream. This may serve as an estimate of the upper-bound of the coverage of this recommendation algorithm.

6.3 Centering on a Photo Inside a Photostream

Among all photostreams, some of them might contain pictures of various topics. Because only five images are shown to the user (see Fig 2), they play an important role in order to catch one's attention. Therefore, we choose which photos to show to the user for each recommended photostream. First, we split photostream in batches of photos, and then we choose the ones that are the most related to the last photos seen by the user.

The photostreams can contain a large number of photos and cover different themes. It has been observed (Graham et al. 2002) that users tend to load images in batches and that photos of the same batch tend to share similar characteristics (e.g. tags). Following this finding, we first split each photostream in batches based on the photo upload date. To this end, we apply the same method as the one that we used to retrieve sessions from the sequences of pageviews (see Section 3.2). The split points will occur when the time difference between two consecutive uploads is more than 25 minutes. The first pictures of each batch are candidates to be shown to the user.

There are many ways to chose a batch that is the most related to the last seen images. Unfortunately, our browsing data is too sparse for this purpose. We therefore use the tags of the photos to choose the most appropriate batch. We aggregate tags of all photos belonging to each batch. As input the recommender system uses the tags of the last $N_p = 5$ photos seen by the user. Finally, the batch that shares the highest number of tags with the set of last seen photos is displayed to the user.

7 Evaluation

We have introduced two recommender systems: a collaborative filtering method using transitions between streams (CF) and a recommender system based on tags (TB). In this section, we compare the performance of these recommender systems focusing on the experience of the user. We test if

Figure 7: The two user interfaces tested in the pilot user study. (a) Original Flickr. Hyperlinks of the photostreams which the current photo belongs to are listed (red box), thumbnails are displayed only for the current photostream (green box). (b) Additional rows of thumbnails. Three rows of thumbnails from other photostreams are shown (red box), centered on the current photo (green rectangle).[2]

the recommended photos are related and interesting to users, and we compare the levels of novelty and serendipity of the provided recommendations. First, we present results of a pilot user study. Second, we present the results of the comparison of the two recommender systems in the main user study.

7.1 Pilot User Study

To recommend photostreams we slightly modify the original Flickr photo-focused interface, shown in Fig. 7a, to place more emphasis on recommended streams. To this end, we show photo thumbnails from three additional streams, as shown in Fig. 7b. Also, to gain space for the additional rows of thumbnails we hide the map of the place where the photo was taken, visible in Fig.7a.

The interface was tested in a pilot user study. Each user was asked to perform two photo-browsing sessions with each of the two user interfaces described above. The sequence of presentation of the interfaces was randomized for every user. Each photo-browsing session lasted 4 minutes. Users were introduced to each interface at the beginning of each session via written instructions on the screen. After completion of both sessions, the user was asked which of the two sessions they liked most. The user study is implemented as a Google Chrome extension which manipulates the way Flickr pages are displayed and automatically manages all the steps of the study.

In total, we had 33 participants: mostly male (78%) between 26 and 40 years old (84%). Around half of them declared that they used Flickr "a few times a year" (52%), and only 25% use it "a few times a month". The great majority of the users preferred additional rows of thumbnails over the original Flickr interface (79% against 9%, 12% no opinion).

[2]Sample Flickr pages from the user http://www.flickr.com/photos/bombeador.

7.2 Comparative User Study

We conducted the user study to test the following hypotheses:

H1) A recommender system based on transitions among streams could propose related and interesting streams to the user.

H2) Collaborative filtering allows the user to explore more novel content than tag-based recommendation.

In order to evaluate the recommendation algorithms, we integrated them in the interface presented in the previous subsection. Each user was asked to go through two photo-browsing sessions, using each of the two recommendation algorithms in random order. Each of the photo-browsing sessions lasted 6 minutes and started with a grid of 100 images randomly selected among the top-1000 photostreams with the highest number of suggestions in the both recommender systems. Users were able to go back to the grid during the experiment. No task was given to the users, apart from a suggestion to browse the photos freely. Each session began with written instructions on the screen describing the task and ended with an evaluation form. Questions were in the form of a statement, and the subjects were able to express their agreement on a 5-point Likert scale (from "strongly disagree" to "strongly agree"). First, we asked users how *related* and *interesting* the recommended photos were. Relatedness expresses how similar the suggested photos are to the displayed one. Interestingness is related to user curiosity and interests. Recommended items are interesting when they catch ones attention. Second, we asked users about *novelty* and *serendipity*. Novelty is the capability of the recommender system to suggest unfamiliar and non-obvious items (Baeza-Yates and Ribeiro-Neto 1999). Serendipity is a related concept. A serendipitous recommendation algorithm proposes items that are novel but also surprisingly interesting. After completion of both sessions, the participant were asked for a final direct comparison.

Results. In total 40 subjects participated in the study, mostly male (66%) between 26 and 40 years old (89%). Around three quarters of them declared that they use Flickr "never" or "a few times a year" (73%), and only 20% use it "a few times a month".

The random null hypothesis has been rejected by χ^2 test ($p < 5 \cdot 10^{-4}$) for the results of each of the questions. For each comparative result we applied the Sharpiro-Wilk normality test. Since the normality null hypothesis was rejected for each distribution, we provide the p-value of the Wilcoxon signed-rank test.

The majority of users agreed (answers: "strongly agree" or "agree") that the recommender systems suggested related pictures (61% for CF, 75% for TB). For both recommender systems, the suggested images were found to be interesting (75% for CF, 69% for TB). Moreover, users considered the collaborative filtering recommender to suggest more novel content (51% for CF, 29% for TB, $p < 0.04$). On average, the collaborative filtering recommender was more likely to provide serendipitous encounters (55% for CF, 38% for TB). However, the two algorithms do not show a large statistically significant difference ($p < 0.11$). In the comments many people reported that they found interesting photos or photographers they liked but did not know. Finally, 44% of the participants preferred CF over TB, 39% preferred TB over CF and 17% did not express any opinion.

Additionally, we analyze logs of the user study and report on them briefly. On average, during the study users of the collaborative filtering system transitioned between photostreams 8.9 times, while those of the tag-based system transitioned 13.3 times. The average number of distinct photostreams seen per session is 11.2 for CF and 11.9 for TB.

Discussion. Based on the results of the user study, we conclude that both recommender systems provided related and interesting suggestions of photostreams and photos, which gives evidence in support of hypothesis H1. Moreover, the collaborative filtering recommender provided more novel content, and to a lesser extent also more serendipitous content, which confirms hypothesis H2. However, this did not result in a significant user preference to either of the recommender systems.

From the log analysis we can see that, due to the fact that the tag-based recommendations are more related and less novel, users are more willing to browse the photos by switching between streams, instead of just browsing consecutive photos of the same stream. However, users have seen on average the same number of distinct photostreams in the two sessions, meaning that, in the case of the tag-based recommender system, users encounter streams that they have already seen more often.

8 Conclusions and Future Work

In this paper, we worked with photostreams as content units for analyzing user browsing behavior in Flickr. In particular, we presented the results of an analysis of a large sample of Flickr navigation logs to gain insights into how users navigate between photostreams. To analyze frequent stream topic transitions, we created a stream transition graph from over 100 million pageviews. We found interesting aggregate patterns in how users navigate between streams and showed that users tend to browse related streams.

Furthermore, we used these findings to design two photostream recommender systems, one based on collaborative filtering (using transitions between photostreams) and one based on content (using photo tags). Each of the algorithms is a part of a two-level photo recommender system, that first recommends photostreams, and then particular photos from the chosen photostreams. The recommender systems are computationally inexpensive.

We compare the two recommender systems through a user study involving 40 participants. The majority of users found the recommended photos to be interesting and related. Moreover, the survey's results confirmed that the collaborative filtering method based on transitions between streams provides more novel recommendations than the tag-based method. In summary, the user studies were useful in gaining insights on the functionality that can be provided. Feedback was mostly positive making the approach very promising.

Future work includes making the recommendation more sophisticated by leveraging the two algorithms jointly. Moreover, the algorithms could be greatly enhanced by taking into account more content features (e.g. EXIF data, favorites), using more advanced methods (e.g. other stream-splitting techniques) or by enabling personalization. In the comments to the experiment, users expressed their interest in reasons behind the recommendations they were seeing. Knowing which of the user's contacts contributed to a personalized recommendation could be interesting and engaging from the user's perspective, and an interesting direction for future studies.

9 Acknowledgments

This research is partially supported by EU's FP7/2007-2013 under the ARCOMEM and Social Sensor projects, by the CDTI (Spain) under the CENIT program, project CEN-20101037 "Social Media", by MICINN (Spain) through Grant TIN2009-14560-C03-01, and by MINECO (Spain) through project MODASS (FIS2011-247852). P.A.G. acknowledges support from the JAE Predoc program of CSIC (Spain).

References

Adomavicius, G., and Tuzhilin, a. 2005. Toward the next generation of recommender systems: a survey of the state-of-the-art and possible extensions. *IEEE Transactions on Knowledge and Data Engineering* 17(6):734–749.

Baeza-Yates, R. A., and Ribeiro-Neto, B. 1999. *Modern Information Retrieval*. Boston, MA, USA: Addison-Wesley Longman Publishing Co., Inc.

Cacheda, F.; Carneiro, V.; Fernández, D.; and Formoso, V. 2011. Comparison of collaborative filtering algorithms. *ACM Transactions on the Web* 5(1):1–33.

Cao, L.; Luo, J.; and Huang, T. S. 2008. Annotating photo collections by label propagation according to multiple similarity cues. In *Proceedings of the 16th ACM international conference on Multimedia*, MM '08, 121–130. New York, NY, USA: ACM.

Chiarandini, L., and Jaimes, A. 2012. Browsing-based content discovery. In *Proceedings of the Designing Interactive Systems Conference*. ACM.

Chiarandini, L.; Trevisiol, M.; and Jaimes, A. 2012. Discovering social photo navigation patterns. In *Multimedia and Expo (ICME), 2012 IEEE International Conference on*, 31–36. IEEE.

Fan, J.; Keim, D.; Gao, Y.; Luo, H.; and Li, Z. 2009. Justclick: Personalized image recommendation via exploratory search from large-scale flickr images. *Circuits and Systems for Video Technology* 19(2):273–288.

Fortunato, S. 2010. Community detection in graphs. *Physics Reports* 486(3–5):75 – 174.

Gozali, J.; Kan, M.; and Sundaram, H. 2012. Hidden markov model for event photo stream segmentation. In *International Conference on Multimedia and Expo Workshops*, 25–30. IEEE.

Grabowicz, P. A.; Ramasco, J. J.; Moro, E.; Pujol, J. M.; and Eguiluz, V. M. 2012. Social features of online networks: The strength of intermediary ties in online social media. *PLoS ONE* 7(1):e29358.

Graham, A.; Garcia-Molina, H.; Paepcke, A.; and Winograd, T. 2002. Time as essence for photo browsing through personal digital libraries. In *Proceedings of the 2nd ACM/IEEE-CS joint conference on Digital libraries*, 326–335.

Herlocker, J. L.; Konstan, J. a.; Terveen, L. G.; and Riedl, J. T. 2004. Evaluating collaborative filtering recommender systems. *ACM Transactions on Information Systems* 22(1):5–53.

Lerman, K., and Jones, L. 2006. Social browsing on flickr. *CoRR* abs/cs/0612047.

Lipczak, M.; Trevisiol, M.; and Jaimes, A. 2013. Analyzing favorite behavior in flickr. In *Advances in Multimedia Modeling*. Springer. 535–545.

Mobasher, B.; Dai, H.; Luo, T.; and Nakagawa, M. 2002. Discovery and Evaluation of Aggregate Usage Profiles for Web Personalization. *Data Mining and Knowledge Discovery* 6(1):61–82.

Ren, K., and Calic, J. 2009. Freeeye: interactive intuitive interface for large-scale image browsing. In *Proceedings of the 17th ACM international conference on Multimedia*, 757–760. ACM.

Ricci, F.; Rokach, L.; and Shapira, B. 2011. Introduction to recommender systems handbook. *Recommender Systems Handbook* 1–35.

Rosvall, M., and Bergstrom, C. T. 2008. Maps of random walks on complex networks reveal community structure. *Proceedings of the National Academy of Sciences* 105(4):1118–1123.

Rosvall, M., and Bergstrom, C. T. 2011. Multilevel compression of random walks on networks reveals hierarchical organization in large integrated systems. *PLoS ONE* 6(4):e18209.

Srikant, R., and Yang, Y. 2001. Mining web logs to improve website organization. In *Proceedings of the 10th international conference on World Wide Web*, 430–437. ACM.

Strong, G.; Hoque, E.; Gong, M.; and Hoeber, O. 2010. Organizing and browsing image search results based on conceptual and visual similarities. *Advances in Visual Computing* 481–490.

Trevisiol, M.; Chiarandini, L.; Aiello, L. M.; and Jaimes, A. 2012. Image ranking based on user browsing behavior. In *Proceedings of the 31st annual international ACM SIGIR conference on Research and development in information retrieval*, 445–454.

Whissell, J. S., and Clarke, C. L. a. 2011. Improving document clustering using Okapi BM25 feature weighting. *Information Retrieval* 14(5):466–487.

Xu, S.; Jin, T.; and Lau, F. 2009. A new visual search interface for web browsing. In *Proceedings of the second ACM international conference on web search and data mining*, 152–161. ACM.

Yu, J.; Jin, X.; Han, J.; and Luo, J. 2011. Collection-based sparse label propagation and its application on social group suggestion from photos. *ACM Transactions on Intelligent Systems and Technology* 2(2):1–21.

Zavesky, E.; Chang, S.-F.; and Yang, C.-C. 2008. Visual islands: intuitive browsing of visual search results. In *Proceedings of the 2008 international conference on Content-based image and video retrieval*, CIVR '08, 617–626. New York, NY, USA: ACM.