TEMPORAL PATTERNS IN SOCIAL MEDIA STREAMS: THEME DISCOVERY AND EVOLUTION USING JOINT ANALYSIS OF CONTENT AND CONTEXT

Yu-Ru Lin   Hari Sundaram   Munmun De Choudhury   Aisling Kelliher

Arts Media Engineering, Arizona State University

Tempe, AZ 85281

Email: {yu-ru.lin, hari.sundaram, munmun.dechoudhury, aisling.kelliher}@asu.edu

ABSTRACT

Online social networking sites such as Flickr, YouTube and Facebook provide a diverse range of functionalities that foster online communities to create and share media content. In particular, Flickr groups are increasingly used to aggregate and share photos about a wide array of topics or themes. Unlike erstwhile photo repositories where images used to be organized with respect to static topics, the photo sharing process as in Flickr often results in miscellaneous time-evolving social and visual patterns. Characterizing such time-evolving patterns can enrich media exploring experience in a social media repository. In this paper, we propose a novel framework that characterizes distinct time-evolving patterns of group photo streams. We use a non-negative joint matrix factorization approach to incorporate image content features and contextual information, including associated tags, photo owners and post times. In our framework, we consider a group as a mixture of themes – each theme exhibits similar patterns of image content and context. The theme extraction is formulated as an optimization problem where the objective is to best explain the observed image content features and associations with tags, users and times. Extensive experiments on a Flickr dataset suggest that our approach is able to extract meaningful evolutionary patterns from group photo streams. We evaluate our method through a tag prediction task. Our prediction results outperform baseline methods, which indicate the utility of our theme based joint analysis.

1 INTRODUCTION

This paper aims at characterizing the time-evolving patterns of visual content and context from community-generated media. The Web 2.0 technology has enabled social networking sites such as Flickr, YoutTube and Facebook to provide a diverse range of functionalities that facilitate social interactions. One such functionality, particularly provided by "groups" on Flickr, allows user generated media to be easily shared among online communities. While on one hand, such sharing enriches end users' experiences through exploring diverse media, the exponential growth of the size of these groups, in terms of both the numbers of photos and users, puts forth several challenges. First, it becomes increasingly difficult for a user to reach to older photos or trace community sharing practices. Second, for a new user, it is difficult to comprehend the changing interests of users in the group or infer the different genres of photos usually posted on the group. Note, the group titles are often not sufficient to understand these changes of patterns over time.

We conjecture that these challenges can be attributed to the time-evolving patterns of the photo content and context in the groups. Hence the primary motivation of this work is to conduct exploratory analysis in group photo streams, through tackling the following questions: Are there typical patterns in the shared photos? Do the patterns grow or diminish over time? How can we extract those patterns, and how can we determine which patterns are followed by particular photos, users, tags, etc. Answers to these questions can be shown to users through an interface as in Figure 1, where the group photos are organized based on patterns over time which we call "themes". The interface allows an effective exploration of group photos with respect to time-dependent themes, and how each theme is contributed by particular users or associated by particular tags.

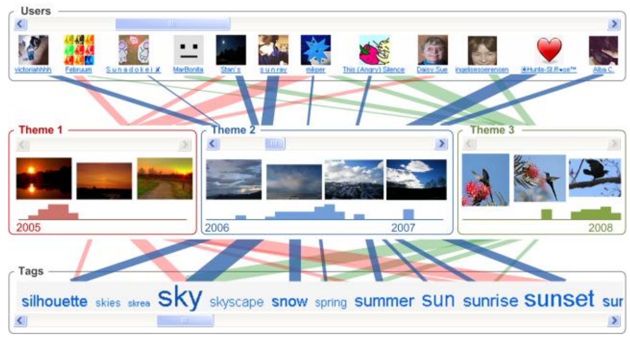

Figure 1: The interface for browsing themes extracted in a Flickr group. A group consists of time-dependent themes having similar patterns of image content and context (users, tags and post times). The thickness of a line between a theme and user or tag indicates how likely a photo in the theme is posted by the user, or is associated with the tag. In each theme, the bars over timeline show how likely a theme photo is posted at the time. The interface allows an effective exploration of typical group photos over time.

Related work. Mining time-evolving patterns of visual content and context in social media streams deals with the interrelatedness of media content features as well as contextual and temporal information associated with the media. To our knowledge, little work has been done to tackle this problem. There has been prior work on mining text corpus by incorporating additional information such as hyperlinks, authors, etc. In text mining, Zhu et al. [12] propose a matrix factorization algorithm combining both the linkage and the document-term matrices to improve the hypertext classification. Mei et al. [8] extract spatiotemporal theme patterns from blogs via a probabilistic approach. On community-generated media analysis, most work has focused on investigating and improving the use of tags, e.g. [1]. In [9], the authors analyze the topical interests of Flickr groups based on the co-occurrences of groups and tags. However, all these studies did not consider the dynamic nature of visual content and context in the media streams generated by community shared practices.

Our approach. In this work, we propose a novel framework that characterizes the time-evolving patterns of group photo streams. We use a non-negative joint matrix factorization approach to incorporate image content features and contextual information, including associated tags, photo owners and post times. In our framework, we consider a group as a mixture of themes – each theme exhibits similar patterns of image content and context. The theme extraction is formulated as an optimization problem where the objective is to best explain the observed image content features and associations with tags, users and times. We provide a scalable iterative algorithm to solve the objectives. Based on the theme extraction results, a group can be represented and explored as shown in Figure 1.

Extensive experiments conducted on a Flickr dataset suggest that our approach is able to extract meaningful evolutionary patterns from group photo streams. In addition to qualitative analysis, we evaluate our method through a tag prediction task – to predict which tags are most likely to be associated on a given image. Our prediction results outperform baseline methods, indicating the utility of our theme extraction approach.

The rest of the paper is organized as follows. Section 2 describes the image content features and context information used in our work. Section 3 presents our theme extraction algorithm. Section 4 discusses the experimental results and section 5 concludes.

2 IMAGE CONTENT AND CONTEXT

This section describes the image content feature extracted from a group photo (section 2.1) and the context information associated with the photo (section 2.2).

2.1 Content Feature Extraction

We use a number of image features that have been found effective in image content analysis. We briefly summarize these features as follows (the details may be referred to in [5,6,7,11]):

1. **Color**: We use two color based features, color histogram and color moments.
2. **Texture**: We use a phase symmetry [11] method for detecting textures from "blobs" in images based on phase information.
3. **Shape**: We use two shape features: radial symmetry [7] and phase congruency [5]. The radial symmetry feature detects points of interest in an image. Phase congruency is an illumination and contrast invariant measure.
4. **SIFT**: We use Scale Invariant Feature Transform (SIFT) [6], a content-based image feature that detects stable keypoint locations in scale space of an image by computing the difference of two nearby scales.

After feature extraction, we construct a D-dimensional feature vector per photo, where D=1064 in this work. Let P be the set of group photos, we obtain a photo-feature matrix $\mathbf{W}^{(F)} \in \Re^{|P| \times D}$, where the i-th row is the feature vector of the i-th photo. We use standard normalization with a logistic function $g(x)=1/(1+\exp(-x))$ to bound the feature values within the range [0,1].

2.2 Image Context

We discuss the context of a shared photo. In social media, a media object such as a photo generally has rich contextual information, e.g. who shares the photo, when the photo is shared, and what additional information is associated with the photo. Three kinds of primary contextual information used in our work are discussed as follows.

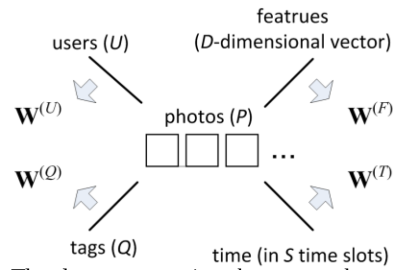

Figure 2: The data representing the group photo stream over time is given as four matrices: photo-feature matrix $\mathbf{W}^{(F)}$, photo-user matrix $\mathbf{W}^{(U)}$, photo-tag matrix $\mathbf{W}^{(Q)}$ and photo-time matrix $\mathbf{W}^{(T)}$.

1. **Users**: The content and concepts of a photo is determined by its owner, and hence the ownership provides the most important contextual information. Let U be the set of users who post the group photos, i.e. photo owners, we construct a photo-user matrix $\mathbf{W}^{(U)} \in \Re^{|P| \times |U|}$, where each entry $\mathbf{W}_{ij}^{(U)}$ =1 if the i-th photo is posted by the j-th user; 0 otherwise.
2. **Tag**: Tags are assigned by users to describe the content of a photo, e.g. "sky", "bird", etc., or provide additional contextual and semantically information, e.g. "summer", "vacation", "nikon", etc. Let Q be the set of tags associated with the group photos, we construct a photo-tag matrix $\mathbf{W}^{(Q)} \in \Re^{|P| \times |Q|}$, where each entry $\mathbf{W}_{ij}^{(Q)}$ =1 if the i-th photo has the j-th tag, and 0 otherwise.
3. **Time**: A photo in Flickr has timestamps indicating when the photo is taken or posted (uploaded). Here we use photo post time since the photo taken time may not be available or accurate for some photos. We segment the timestamps into S time slots and construct a photo-time matrix $\mathbf{W}^{(T)} \in \Re^{|P| \times S}$, where each entry $\mathbf{W}_{ij}^{(T)}$ =1 if the i-th photo is posted during the j-th time slot, and 0 otherwise. We further smooth each row vector of $\mathbf{W}^{(T)}$ using a moving average filter, so that $\mathbf{W}^{(T)}$ contains smooth temporal information.

In our analysis, the data representing the group photo stream over time is given as data matrices described above, including image content features as well as contextual and temporal information of the images. In our notations, a matrix $\mathbf{W}^{(\cdot)}$ is indexed based on its second dimension. Without loss of generality, we normalize $\mathbf{W}^{(\cdot)}$ to ensure $\Sigma_{ij}\mathbf{W}_{ij}^{(\cdot)}$=1. The data representation is illustrated as in Figure 2.

3 THEME EXTRACTION

We now present our method for extracting time-evolving patterns of visual content and context from group photo streams.

We consider a group as a mixture of themes – each theme consists of photos having similar patterns of content and context. In our approach, patterns are deemed to be similar in terms of multiple aspects: visual content, associated tags, photo owners, and post times. We extract themes based on the observations from these aspects – which are represented by the data matrices discussed in the previous session. We shall discuss each of these aspects in the following.

Visual features. In order to extract themes having similar visual patterns, let us assume each theme k has a length D feature vector $\mathbf{z}_k$ where each entry z_{kj} can best represents the significance of the j-th feature of the set of photos in the theme k. Our goal is to determine the coefficient z_{kj} based on how likely a photo i belongs to the theme k. We define p_{ik} to be the probability that a particular photo i belongs to the theme k. The parameters p_{ik} are non-negative numbers satisfying $\Sigma_i \, p_{ik}=1$. Let $\mathbf{Z}^{(F)}=\{z_{ki}\}$ denote a $K\times D$ matrix, $\mathbf{P}=\{p_{ik}\}$ denote a $|P|\times K$ matrix. We use an idea similar to principal component analysis to derive $\mathbf{P}$ and $\mathbf{Z}^{(F)}$ from the data matrix $\mathbf{W}^{(F)}$ as:

$$\mathbf{W}_{ij}^{(F)} \approx \sum\nolimits_k p_{ik} z_{kj} = (\mathbf{P}\mathbf{Z}^{(F)})_{ij} \tag{<1>}$$

This suggests the approximation can be done by minimizing $D(\mathbf{W}^{(F)} \| \mathbf{P}\mathbf{Z}^{(F)})$, given $D(\cdot\|\cdot)$ as a measure of approximation cost between two matrices. In this paper we focus on KL (Kullback-Leibler) divergence between two matrices. The corresponding objective is to minimize:

$$\begin{aligned} &J(\mathbf{P},\mathbf{Z}^{(F)}) = D(\mathbf{W}^{(F)} \,\|\, \mathbf{P}\mathbf{Z}^{(F)}) \\ &s.t.\ \mathbf{P} \in \Re_+^{|P|\times K}, \mathbf{Z}^{(F)} \in \Re_+^{K\times D}, \sum\nolimits_i \mathbf{P}_{ik} = 1 \ \forall k \end{aligned} \tag{<2>}$$

where $D(\mathbf{A}\|\mathbf{B}) = \sum_{ij} (\mathbf{A}_{ij} \log \mathbf{A}_{ij}/\mathbf{B}_{ij} - \mathbf{A}_{ij} + \mathbf{B}_{ij})$ is the KL divergence between matrices $\mathbf{A}$ and $\mathbf{B}$. With the non-negative constraints, this optimization problem is a case of non-negative matrix factorization (NMF) [3].

Users and Tags. Now considering the contextual information of photos, two photos might belong to the same theme not only due to visually similarily, but also because they are posted by the same users, or associated with the same tags. For a set of users U, suppose we have a $K\times|U|$ user coefficient matrix $\mathbf{Z}^{(U)}$, where each entry z_{kj} indicates how likely the j-th user posts a photo that falls in the k-th theme. Similar to the feature matrix factorization, based on the data matrix $\mathbf{W}^{(U)}$ with the (i,j)-entry indicating the photo i posted by the user j, we determine z_{kj} based on how likely the photo i belongs to the theme k. We approximate the data matrix $\mathbf{W}^{(U)}$ by $\mathbf{P}$ and $\mathbf{Z}^{(U)}$ by the following objective:

$$J(\mathbf{P},\mathbf{Z}^{(U)}) = D(\mathbf{W}^{(U)} \,\|\, \mathbf{P}\mathbf{Z}^{(U)}) \tag{<3>}$$

subject to $\mathbf{Z}^{(U)}\in\Re_+^{K\times|U|}$, with other constraints and $D(\cdot\|\cdot)$ defined as in eq. <2>. Similarly, let $\mathbf{Z}^{(Q)}$ be the tag coefficient matrix, where each entry z_{kj} indicates how likely the j-th tag is associated with a photo that belongs to the k-th theme. We approximate the data matrix $\mathbf{W}^{(Q)}$ by $\mathbf{P}$ and $\mathbf{Z}^{(Q)}$ by:

$$J(\mathbf{P},\mathbf{Z}^{(Q)}) = D(\mathbf{W}^{(Q)} \,\|\, \mathbf{P}\mathbf{Z}^{(Q)}) \tag{<4>}$$

s.t. $\mathbf{Z}^{(Q)}\in\Re_+^{K\times|Q|}$, with other constraints defined as in eq. <2>.

Temporal information. To extract themes having temporal trends, we consider two photos to be in the same theme if they are posted during the same time. For S time slots, let $\mathbf{Z}^{(T)}$ be the time coefficient matrix, where each entry z_{kj} indicates how likely a photo posted at time j belongs to the k-th theme. We approximate the data matrix $\mathbf{W}^{(T)}$ by $\mathbf{P}$ and $\mathbf{Z}^{(T)}$ as follows:

$$J(\mathbf{P},\mathbf{Z}^{(T)}) = D(\mathbf{W}^{(T)} \,\|\, \mathbf{P}\mathbf{Z}^{(T)}) \tag{<5>}$$

s.t. $\mathbf{Z}^{(T)}\in\Re_+^{K\times S}$, with other constraints defined as in eq. <2>.

Joint Factorization. Putting together all objective functions with respect to different aspects, we have the following joint objective function:

$$\begin{aligned} &J(\mathbf{P},\{\mathbf{Z}^{(r)}\}) = \sum_{r=\{F,U,Q,T\}} D(\mathbf{W}^{(r)} \,\|\, \mathbf{P}\mathbf{Z}^{(r)}) \\ &s.t.\ \mathbf{P} \in \Re_+^{|P|\times K}, \mathbf{Z}^{(r)} \in \Re_+^{K\times I_r}, \sum\nolimits_i \mathbf{P}_{ik} = 1 \ \forall k \end{aligned} \tag{<6>}$$

where $\{\mathbf{Z}^{(r)}\}$ is a set of coefficient matrices for all the aspects (visual features, users, tags, and times) and I_k denotes the dimensionality of the second dimension of the coefficient matrices. Note, eq. <6> can be easily extended to incorporate additional aspects or to incorporate weights on aspects.

Solution. We provide a solution to the objective defined in eq. <6>. Since eq. <6> is not convex in all variables, it is difficult to guarantee a global minima solution. By employing the concavity of the log function given in the KL-divergence, a local minima solution can be derived as the following updating algorithm:

$$\begin{aligned} &\mathbf{P}_{ik} \leftarrow \sum_r \sum_j \mathbf{W}_{ij}^{(r)} \mu_{ijk}^{(r)}, \quad \mathbf{Z}_{kj}^{(r)} \leftarrow \sum_r \sum_i \mathbf{W}_{ij}^{(r)} \mu_{ijk}^{(r)}, \\ &\text{where } \mu_{ijk}^{(r)} = \frac{\mathbf{P}_{ik}\mathbf{Z}_{kj}^{(r)}}{(\mathbf{P}\mathbf{Z}^{(r)})_{ij}} \end{aligned} \tag{<7>}$$

and then normalize such that $\Sigma_i \, \mathbf{P}_{ik}=1$. This iterative update algorithm extends the algorithm proposed by Lee et al. [3] for solving a single non-negative matrix factorization problem. The proof for the convergence of our algorithm is omitted due to space limit. In eq.<7>, due to the sparseness of the data matrices, the most time-consuming part, computing $(\mathbf{P}\mathbf{Z}^{(r)})_{ij}$ is reduced to computing only the corresponding $(\mathbf{P}\mathbf{Z}^{(r)})_{ij}$ for each non-zero entry. Let n be the number of non-zero entries of all data matrices and let K be the number of themes, the total time complexity per iteration is $O(nK)$. If we consider K is bounded by some constant, the time complexity is linear in $O(n)$. Note that since $\mathbf{W}^{(F)}$ is typically constructed based on fixed length feature vector, it can be considered to have the same degree of sparseness as other data matrices.

4 EXPERIMENTS

This section reports our experimental studies on a collection of Flickr group dataset. We first briefly describe the dataset used in our experiments, and present a qualitative analysis on the group theme extraction. Finally we quantitatively evaluate and discuss the quality of our mining results through a prediction task.

Flickr dataset. We have collected the data using Flickr API. We sample 52 groups based on the group size distribution. We download all photos for each group. In total, there are 50,875 photos, 3,499 unique users, and 54,501 unique tags in this collection. The photo post times range from May 23, 2004 to November 11, 2008, enabling us to analyze long-term temporal patterns in this collection.

Theme discovery and evolution. We discuss the theme extraction results from two groups: (a) “sky's the limit” (2,278 photos), and (b) “Full Frame Sensor group” (3,961 photos). The themes are extracted based on our joint analysis. As shown in Figure 3, there are five themes (A1 – A5) in group (a) and six themes (B1 – B6) in group (b). The theme numbers are automatically determined based on a soft modularity measure defined in [4]. For each theme, we show the top three photos based on how likely a photo i belongs to the theme k, i.e. p_{ik} in $\mathbf{P}$. The most likely users who post the theme photos and most likely tags associated with theme photos can be determined based on the coefficient matrices $\mathbf{X}^{(U)}$ and $\mathbf{X}^{(Q)}$ respectively, which we omit to show due to space limit. Such information can be effectively presented to the end user through an interactive interface as shown in Figure 1. The theme evolution, i.e. how likely a theme appears at a particular time, is determined based on the coefficient matrix $\mathbf{X}^{(T)}$. We plot the theme evolution for group (a) and (b) in Figure 3(c) and (d). As can be seen, some themes only appear at certain time period and then diminish, e.g. A2, B2, etc., while some themes re-appear after first showing up, e.g. A3, A5, etc. We have observed that these themes emerge

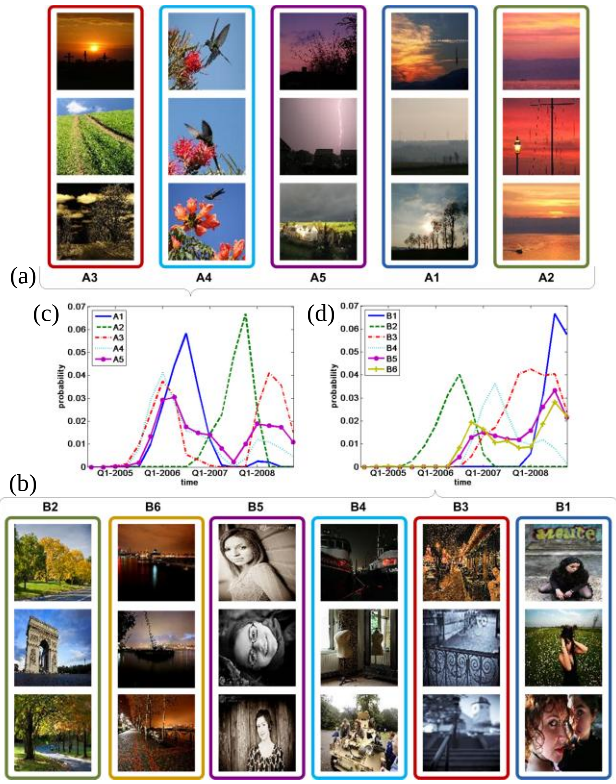

Figure 3: Theme discovery for (a) group "sky's the limit" and (b) "Full Frame Sensor group". Their theme evolutions are shown in plot (c) and (d) respectively. We use our joint analysis to extract five themes in group (a) and six themes in group (b). The results interestingly show that group patterns emerge due to dedicated users, tag co-occurrences, as well as similar visual content.

due to dedicated users (e.g. the "bird" images in A4 are taken by the same user), tag co-occurrences (e.g. "sunset" in A2, "water" in B6, etc.), as well as similar visual content (e.g. A2, A4, A5, B2, B5, B6, etc.). These observations suggest that our analysis captures the dynamics of group patterns and gives meaningful summary of group photo streams.

Prediction task. We design a prediction task to examine the quality of our mining results. For each group, we randomly choose 90% photos to train the theme model, and use the rest 10% as testing set. The task is to predict tags associated with the testing images. Our prediction utilizes the coefficient matrices obtained from the training stage, with an estimation of p_{ik} for the test images using a folding-in technique [10]. We compare our prediction results with three baseline methods: (a) feature-based prediction – predicting tags from photos having most similar visual features; (b) tag-based – predicting tags based on the tag frequency; (c) feature/tag (denoted by "F/T") – predicting tags by only using the feature and tag information in joint factorization. The prediction performance is evaluated based on S@10 (success among the top 10 results) and NDCG [2].

The prediction performance averaged over all groups is shown in Table 1. The result shows that our method significantly outperforms all baselines by 45% – 3.9X on an average. Note, prediction based on visual features alone performs the worst. We observe that our joint analysis that incorporates contextual information (tags, users, post times) with visual features, finds the highest quality tags for the testing images – suggesting the effectiveness of our mining approach.

	Features	Tags	F/T	Joint
NDCG	0.088±0.057	0.299±0.120	0.297±0.122	**0.433±0.105**
S@10	0.378±0.260	0.728±0.193	0.733±0.190	**0.796±0.122**

Table 1: The average tag prediction performance evaluated by NDCG and S@10 metrics.

5 CONCLUSION

In this paper, we propose a framework to characterize the time-evolving patterns of group photo streams. We consider a group as a mixture of themes and use a non-negative joint matrix factorization approach to incorporate image content features and contextual information, including associated tags, photo owners and post times. Extensive experiments on a Flickr dataset show that (a) our analysis is able to capture the dynamics of group patterns, and give meaningful summary of group photo streams; (b) compared with baseline methods, our joint analysis performs the highest quality tag prediction. These results indicate the utility of our theme based joint analysis

As part of our future work, we plan to (a) extract higher level concepts from the theme mining results, and (b) investigate the hierarchical structures of themes. We are also extending the current matrix factorization framework to tensor based analysis.

6 REFERENCES

[1] N. Garg and I. Weber (2008). *Personalized, interactive tag recommendation for flickr*, RecSys, 67-74, 2008.

[2] K. Järvelin and J. Kekäläinen (2000). *IR evaluation methods for retrieving highly relevant documents*, SIGIR, 41-48, 2000.

[3] D. Lee and H. Seung (2001). *Algorithms for non-negative matrix factorization*, NIPS, 556–562, 2001.

[4] Y. Lin, Y. Chi, S. Zhu, H. Sundaram and B. Tseng (2008). *Facetnet: a framework for analyzing communities and their evolutions in dynamic networks*, WWW, 2008.

[5] Z. Liu and R. Laganiere (2007). *Phase congruence measurement for image similarity assessment.* Pattern Recogn. Lett. **28**(1): 166-172.

[6] D. G. Lowe (2004). *Distinctive Image Features from Scale-Invariant Keypoints.* Int. J. Comput. Vision **60**(2): 91-110.

[7] G. Loy and A. Zelinsky (2003). *Fast Radial Symmetry for Detecting Points of Interest.* IEEE Trans. Pattern Anal. Mach. Intell. **25**(8): 959-973.

[8] Q. Mei, C. Liu, H. Su and C. Zhai (2006). *A probabilistic approach to spatiotemporal theme pattern mining on weblogs*, WWW, 533-542, 2006.

[9] R. Negoescu and D. Gatica-Perez (2008). *Analyzing Flickr groups*, CIVR, 417-426, 2008.

[10] A. Schein, A. Popescul, L. Ungar and D. Pennock (2002). *Methods and metrics for cold-start recommendations*, SIGIR, 253-260,

[11] Z. Xiao, Z. Hou, C. Miao and J. Wang (2005). *Using phase information for symmetry detection.* Pattern Recogn. Lett. **26**(13): 1985-1994.

[12] S. Zhu, K. Yu, Y. Chi and Y. Gong (2007). *Combining content and link for classification using matrix factorization*, SIGIR, 487-494, 2007.