SUMMARIZATION OF LARGE SCALE SOCIAL NETWORK ACTIVITY

Yu-Ru Lin Hari Sundaram Aisling Kelliher

Arts Media and Engineering Program, Arizona State University
{yu-ru.lin, hari.sundaram, aisling.kelliher}@asu.edu

ABSTRACT

This paper presents a novel social media summarization framework. Summarizing media created and shared in large scale online social networks unfolds challenging research problems. The networks exhibit heterogeneous social interactions and temporal dynamics. Our proposed framework relies on the co-presence of multiple important facets: *who* (users), *what* (shared concepts and media), *how* (actions) and *when* (time). First, we impose a syntactic structure of the social activity (relating users, media and concepts via specific actions) in our temporal multi-graph mining algorithm. Second, important activities along each facet are extracted as activity themes over time. Experiments on real-world Flickr datasets demonstrate that our technique capture nontrivial evolution of media use in social networks.

1. INTRODUCTION

We present a method for automatically summarizing social media created from online social networks. The proliferation of Web 2.0 social networking sites such as Flickr, YouTube and Facebook, represents a welcome of community-centric experience – by sharing rich media (text, images, video, etc.), day-to-day activities are scattered and feedbacked by a network of peers. For example, in Flickr, photo-sharing activities lead to awareness among peers, development of common interests, bursty discussions, and so on. Such experience, however, requires significant user efforts of tracking the collective activities in community. This work seeks to accommodate users by introducing an automated social media summarization framework for digesting collective activities.

Related work. The problem of summarizing social media is highly challenging because media-sharing activities are constantly changing mishmash of interrelated users and media objects. Recent analysis of social groups and their temporal dynamics [2,3] has focused on dynamic but homogeneous networks, i.e. the edges represent homogeneous actions (e.g. posting). However in online social networks users can interact with each other though actions with respect to different types of media objects. Although heterogeneous interrelated entities have attracted considerable interest [1,4,5,6], these multi-graph mining algorithms do not consider the temporal evolution of the interrelated entities. To summarize the dynamic and diverse activity context of social media, we propose a unified temporal multi-graph framework which extracts activity themes over time.

We propose JAM (Joint Action Matrix Factorization), a novel method for summarizing media created and shared in online social networks. Our proposed framework relies on the co-presence of multiple important facets: *who*, *what*, *how* and *when*. There are two key contributions in the proposed framework: First, we impose a syntactic structure of the social activity (relating users, media and concepts via specific actions) in a temporal multi-graph mining algorithm. Second, important activities along each facet are extracted as activity themes over time. Experiments on real-world Flickr datasets demonstrate that our technique capture nontrivial evolution of media use in social networks.

The rest of the paper is organized as follows. Section 2 defines the problem. Section 3 presents our proposed method. Section 4 shows the experiment results and section 5 presents the conclusions.

2. PROBLEM DEFINITION

Data model. We construct a data model consisting of the interrelated data entities, including media objects, people, comments and tags. This data model (ref. Figure 1 for a summary of notations) is based on Flickr's social groups and can generalize to other social networks. The data includes different object sets – U (users), P (timestamped media objects such as photos), C (timestamped comments on media) and Q (media description such as tags). Tags are assigned by users and have been commonly used to annotate and retrieve the relevant concepts of a photo. Thus we use tags to represent the concepts of a photo. In the rest of the paper, we use "concept", "tag", "term" interchangeably.

There are heterogeneous relationships among these objects, which can be represented by bi-partite graphs and their corresponding matrices (ref. Figure 1). These matrices are basic relationships in our data model, as other relationships can be derived from a combination of the prime matrices.

We formulate the social media summarization problem in terms of extracting temporally representative social activities. The key idea is to extract activity as a composite of multiple important facets that provide a rich context to understand the social meaning of the media.

An activity a is defined as a co-presence of multiple important facets*: user*, *action* and *term*. An action (identified by *media* object and *time*) indicates how these facets are

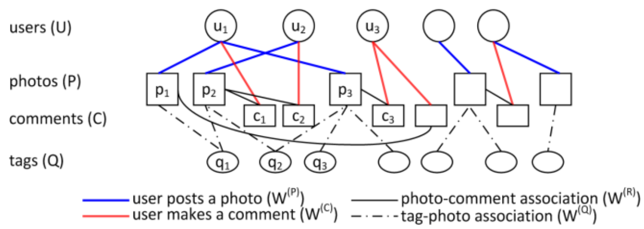


Figure 1: Data model in a photo-sharing space, including four sets of heterogeneous data objects U (users), P (photos), C (comments), Q (tags) and four types of relationships among these objects, which can be represented by matrices: $W^{(P)}$ – user-photo matrix where each entry $W_{ij}^{(P)}$ indicates user u_i posts photo p_j, $W^{(C)}$ – user creating a comment, $W^{(Q)}$ – tag-photo association and $W^{(R)}$ – photo-comment association.

associated, e.g. a user *posts* (action) a photo with respect to some tagged concepts, or *comments* (action) on a photo posted by another user. An *activity theme* A_t for a time t is a set of activities i.e. A_t:=$\{a|t\}$. Finally, a temporal *social media summary* A is a sequence of activity themes, i.e. $A := \{A_1, A_2, \ldots, A_t, \ldots\}$.

Problem. Given interrelated social data entities (U, P, C, Q), extract an activity theme A_t for each time t, to construct the temporal activity-based media summary A.

3. PROPOSED METHOD

In this section we propose a unified matrix factorization framework for extracting activity themes over time.

3.1 Joint-Action matrix factorization

We solve the problem of extracting temporally representative activity themes in a non-negative matrix factorization (NMF) framework. We formalize such theme extraction as a multi-graph clustering problem – each activity theme is a cluster of strongly co-occurring users, actions and terms. The summarization problem requires (1) distinguishing the semantics of different actions, e.g. posting a photo and commenting on a photo have different activity semantics, and (2) differentiating activities across time.

Assume there are K activity themes. We first examine two actions among users and terms. The "post-on" actions comprise user-photo ($\mathbf{W}^{(P)}$) and term-photo ($\mathbf{W}^{(Q)}$) relationships. The "comment-on" actions comprise user-comment ($\mathbf{W}^{(C)}$) and term-comment relationship. The term-comment relationship, denoted by $\mathbf{W}^{(QR)}$, is not directly available from the data model but can be derived in a straightforward manner by combining term-photo and photo-comment relationships, i.e. $\mathbf{W}^{(QR)} = \mathbf{W}^{(Q)}\mathbf{W}^{(R)}$.

Let us begin with the $|Q|\times|P|$ term-photo matrix $\mathbf{W}^{(Q)}$. Similar to a term-document matrix, the K-dimensional latent space can be factorized into a $|Q|\times K$ matrix $\mathbf{Y}$ and a $K\times|P|$ matrix $\mathbf{Z}^{(P)}$, where each column of $\mathbf{Y}$ is the axis of each dimension. Photos can be projected to each dimension by the coefficients in the corresponding rows of $\mathbf{Z}^{(P)}$. We put non-negative constraints on $\mathbf{Y}$ and $\mathbf{Z}^{(P)}$ so that each theme is represented by an additive combination of terms. Using $\mathbf{YZ}^{(P)}$ to approximate the matrix $\mathbf{W}^{(Q)}$, we seek to minimize the following objective function:

$$J(\mathbf{Y},\mathbf{Z}^{(P)}) = D(\mathbf{W}^{(Q)} \,\|\, \mathbf{YZ}^{(P)}) \quad \text{<1>}$$
$$s.t\ \mathbf{Y}\in\Re_+^{|Q|\times K}, \mathbf{Z}^{(P)}\in\Re_+^{K\times|P|}, \sum\nolimits_i \mathbf{Y}_{ij} = 1\ \forall j$$

where $D(\mathbf{A}\|\mathbf{B}) = \sum_{i,j} (\mathbf{A}_{ij} \log \mathbf{A}_{ij}/\mathbf{B}_{ij} - \mathbf{A}_{ij} + \mathbf{B}_{ij})$ is the Kullback-Leibler (KL) divergence between matrices $\mathbf{A}$ and $\mathbf{B}$, which is used to measure how the factorization deviates from the observed data.

The latent space $\mathbf{Y}$ is now solely produced by users' posting actions. We then incorporate different action semantics – users' commenting actions with respect to the terms. Let us use the $|Q|\times|C|$ term-comment matrix $\mathbf{W}^{(QR)}$ to relate terms to the latent space through commenting actions. To combine different action semantics, we use the same latent space $\mathbf{Y}$ to approximate the matrix $\mathbf{W}^{(QR)}$ as follows:

$$J(\mathbf{Y},\mathbf{Z}^{(C)}) = D(\mathbf{W}^{(QR)} \,\|\, \mathbf{YZ}^{(C)}) \quad \text{<2>}$$
$$s.t\ \mathbf{Y}\in\Re_+^{|Q|\times K}, \mathbf{Z}^{(C)}\in\Re_+^{K\times|P|}, \sum\nolimits_i \mathbf{Y}_{ij} = 1\ \forall j$$

Next, we relate users to the latent space. Given the user-photo matrix $\mathbf{W}^{(P)}$ and the user-comment matrix $\mathbf{W}^{(C)}$, if we use the same coefficient matrices $\mathbf{Z}^{(P)}$ and $\mathbf{Z}^{(C)}$ to represent the same set of photos and comments, we can find corresponding K-dimensional latent space which axes are represented by users. Thus, we approximate $\mathbf{W}^{(P)}$ and $\mathbf{W}^{(C)}$ by the following objective functions:

$$J(\mathbf{X},\mathbf{Z}^{(P)}) = D(\mathbf{W}^{(P)} \,\|\, \mathbf{XZ}^{(P)}) \quad \text{<3>}$$
$$s.t\ \mathbf{X}\in\Re_+^{|U|\times K}, \mathbf{Z}^{(P)}\in\Re_+^{K\times|P|}, \sum\nolimits_i \mathbf{X}_{ij} = 1\ \forall j$$

$$J(\mathbf{X},\mathbf{Z}^{(C)}) = D(\mathbf{W}^{(C)} \,\|\, \mathbf{XZ}^{(C)}) \quad \text{<4>}$$
$$s.t\ \mathbf{X}\in\Re_+^{|U|\times K}, \mathbf{Z}^{(C)}\in\Re_+^{K\times|C|}, \sum\nolimits_i \mathbf{X}_{ij} = 1\ \forall j$$

Now we have a combined objective function $J_1 = J(\mathbf{X},\mathbf{Z}^{(P)})+J(\mathbf{X},\mathbf{Z}^{(C)})+J(\mathbf{Y},\mathbf{Z}^{(P)})+J(\mathbf{Y},\mathbf{Z}^{(C)})$. Minimizing this objective function will give two K-dimensional latent spaces that correspond to K activity themes. In the space of $\mathbf{X}$, each dimension axis is represented by a column of $\mathbf{X}$, i.e. $\mathbf{X}_j$, and each entry $\mathbf{X}_{ij}$ indicates the strength of user u_i associating with the j-th activity theme. Similarly in another space $\mathbf{Y}$, each axis is represented by a column $\mathbf{Y}_j$, and $\mathbf{Y}_{ij}$ indicates the strength of term q_j associating with the j-th theme.

To extract themes that have temporal correlation, we introduce time indicator matrices for both actions. First we segment the data duration into T time slots. For $|P|$ posting actions, we construct a $T\times|P|$ matrix $\mathbf{H}^{(P)}$, with each entry $\mathbf{H}_{ti}^{(P)}$=1 indicating that the photo p_i is posted during time t, and 0 otherwise. Similarly for commenting actions, we construct a $T\times|C|$ matrix $\mathbf{H}^{(C)}$. We want the K-dimensional latent spaces to align with these time slots, thus we let K=T. We regularize the objective function by these time indicators:

$$r_1 D(\mathbf{H}^{(P)} \,\|\, \mathbf{Z}^{(P)}) + r_2 D(\mathbf{H}^{(C)} \,\|\, \mathbf{Z}^{(C)}) \quad \text{<5>}$$

where $\gamma_1 \geq 0$ and $\gamma_2 \geq 0$ give the weights for both regularization. Combining with time regularization and different actions, the goal is to minimize the following objective function:

$$
\begin{aligned}
&J(\mathbf{X},\mathbf{Y},\{\mathbf{Z}^{(\kappa)}\}) = \\
&\sum_{\kappa=1}^{M}\left(\begin{aligned} &\alpha_\kappa D(\mathbf{W}^{(u,\kappa)} \,\|\, \mathbf{XZ}^{(\kappa)}) + \beta_\kappa D(\mathbf{W}^{(q,\kappa)} \,\|\, \mathbf{YZ}^{(\kappa)}) + \\ &r_\kappa D(\mathbf{H}^{(\kappa)} \,\|\, \mathbf{Z}^{(\kappa)}) \end{aligned} \right) \qquad \text{<6>} \\
&s.t\ \mathbf{X} \in \Re_+^{|U| \times T}, \mathbf{Y} \in \Re_+^{|Q| \times T}, \mathbf{Z}^{(\kappa)} \in \Re_+^{T \times I_\kappa} \\
&\quad \textstyle\sum_i \mathbf{X}_{ij} = 1\ \forall j, \sum_i \mathbf{Y}_{ij} = 1\ \forall j
\end{aligned}
$$

where $\mathbf{X}$ and $\mathbf{Y}$ represent the latent spaces of users and terms respectively, $\{\mathbf{Z}^{(\kappa)}\}$ is a set of coefficient matrices for each action type κ. $\mathbf{W}^{(u,\kappa)}$ represents the user-action-κ relationships where each entry indicates a user u perform an action of type κ. Similarly $\mathbf{W}^{(q,\kappa)}$ represents the term-action-κ relationships. $\mathbf{H}^{(\kappa)}$ is a time indicator matrix for type κ actions. For example, if κ represents the posting action, the respective matrices are $\mathbf{W}^{(P)}$, $\mathbf{W}^{(Q)}$ and $\mathbf{H}^{(P)}$. α_κ, β_κ, and γ_κ are positive weights for type κ actions. I_κ is the set size (total amount) of type κ actions. The proposed matrix factorization framework is summarized and illustrated as in Figure 2.

3.2 Solution and computational complexity

We provide an iterative algorithm to solve the optimization problem defined by Eq.<6>. We then show that the time complexity of our algorithm is scalable with the total number of data entities in the network.

A solution to Eq.<6> can be found by the update rules:

$$
\begin{aligned}
&\mathbf{X}_{it} \leftarrow \frac{1}{c_X}\left(\sum_\kappa \alpha_\kappa \sum_j \mathbf{W}_{ij}^{(u,\kappa)} \sigma_{ijt}^{(u,\kappa)}\right), \\
&\mathbf{Y}_{kt} \leftarrow \frac{1}{c_Y}\left(\sum_\kappa \beta_\kappa \sum_j \mathbf{W}_{kj}^{(q,\kappa)} \sigma_{kjt}^{(q,\kappa)}\right), \qquad \text{<7>} \\
&\mathbf{Z}_{tj}^{(\kappa)} \leftarrow \frac{1}{c_Z^{(\kappa)}}\left(\alpha_\kappa \sum_i \mathbf{W}_{ij}^{(u,\kappa)} \sigma_{ijt}^{(u,\kappa)} + \beta_\kappa \sum_k \mathbf{W}_{kj}^{(q,\kappa)} \sigma_{kjt}^{(q,\kappa)} + \gamma_\kappa \mathbf{H}_{tj}^{(\kappa)}\right),
\end{aligned}
$$

then normalize such that $\sum_i \mathbf{X}_{ij} = 1\ \forall j, \sum_i \mathbf{Y}_{ij} = 1\ \forall j,$

where:

$$
c_X = \sum_\kappa \alpha_\kappa,\ c_Y = \sum_\kappa \beta_\kappa,\ c_Z^{(\kappa)} = \alpha_\kappa + \beta_\kappa + \gamma_\kappa,
$$

$$
\sigma_{ijt}^{(u,\kappa)} = \frac{\mathbf{X}_{it}\mathbf{Z}_{tj}^{(\kappa)}}{(\mathbf{XZ}^{(\kappa)})_{ij}},\ \sigma_{kjt}^{(q,\kappa)} = \frac{\mathbf{Y}_{kt}\mathbf{Z}_{tj}^{(\kappa)}}{(\mathbf{YZ}^{(\kappa)})_{kj}}.
$$

The above update rules are derived based on the concavity of log function. The proof for the convergence of these rules is skipped due to the space limit.

Scalability. We now investigate the time complexity for each iteration of the updates in Eq.<7>. The most time-consuming part is to compute $\forall\kappa$ $(\mathbf{XZ}^{(\kappa)})_{ij}$ $\forall i,j$, and $(\mathbf{YZ}^{(\kappa)})_{kj}$ $\forall k,j$. However, due to the sparseness of $\mathbf{W}^{(u,\kappa)}$ and $\mathbf{W}^{(q,\kappa)}$, we only need to compute the corresponding $(\mathbf{XZ}^{(\kappa)})_{ij}$, $(\mathbf{YZ}^{(\kappa)})_{kj}$, for each nonzero entry in $\mathbf{W}^{(u,\kappa)}$ and $\mathbf{W}^{(q,\kappa)}$, respectively. Thus the total time complexity is $O(mT)$, where m is the number of non-zero entries in the input matrices, and T is the number of time slots. If we consider the number of time slots T and the degree of nodes (users and terms) in the data is bounded by some constant, the complexity is linear in the total number of entities in the networks.

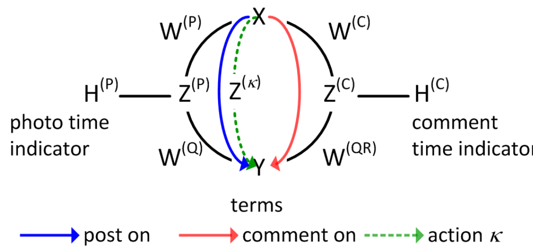


Figure 2: Our proposed joint-action matrix factorization with time regularization. The framework derives two interrelated latent spaces represented by users X and concepts Y through different semantics of actions (post-on and comment-on) represented by matrices $W^{(P)}$, $W^{(Q)}$, $W^{(C)}$ and $W^{(QR)}$, with matrices $H^{(P)}$ and $H^{(C)}$ indicating when the actions occur for enforcing the temporal co-occurrence of actions. The model can be generalized to include other types of actions.

3.3 Theme representation

We begin with constructing the activity theme for each time, and then discuss how to extract theme evolution.

Activity theme. For each activity theme A_t at time t, we want to find a set of users, concept terms, actions and media objects as facets of representative activities for A_t. Because each column of $\mathbf{X}$, i.e. $\mathbf{X}_t$, corresponds to an activity theme A_t, and each entry in the column, $\mathbf{X}_{it}$, indicates how strong the user u_i associates with the theme A_t, we can extract users who have the highest values of $\mathbf{X}_{it}$ as representative users of A_t. Similarly the representative terms are determined based on top K values of $\mathbf{Y}_{it}$. The representative media are determined by $\mathbf{Z}^{(\kappa)}$, e.g. $\mathbf{Z}^{(P)}{}_{tj}$ indicates how strong the photo p_j associates with A_t. Thus we extract top K media based on $\mathbf{Z}^{(\kappa)}$. Finally we extract the actions that exist among the representative users and terms during the time t.

Theme evolution. To determine how a theme evolves into another theme, we compute the similarity between themes by comparing the corresponding axes in both latent spaces, based on a cosine similarity measure. For two activity theme A_i, A_j, the similarity between them is defined as:

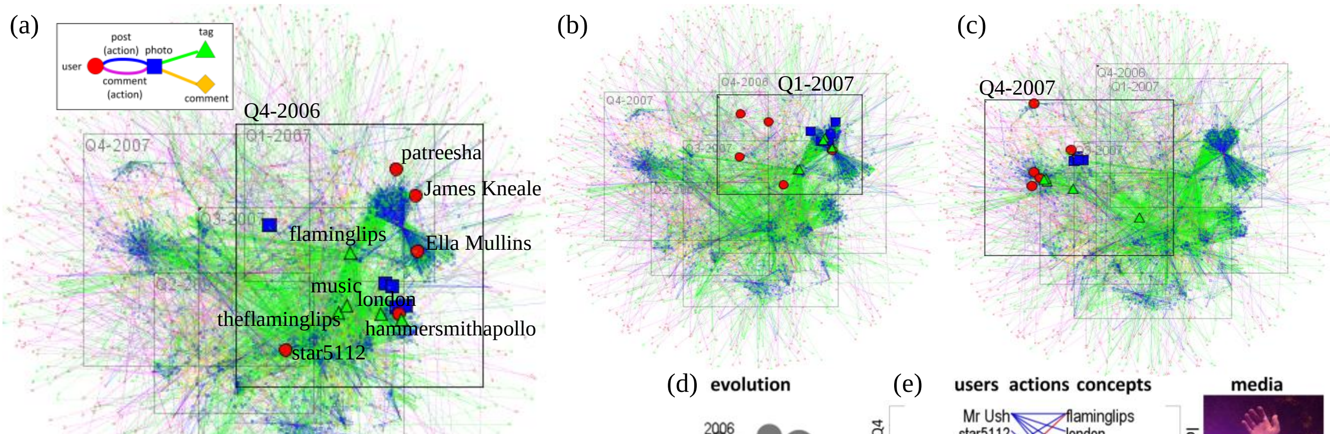
$$
s(A_i, A_j) = (1-\omega)\frac{\mathbf{X}_i \cdot \mathbf{X}_j}{|\mathbf{X}_i\|\mathbf{X}_j|} + \omega\frac{\mathbf{Y}_i \cdot \mathbf{Y}_j}{|\mathbf{Y}_i\|\mathbf{Y}_j|} \qquad \text{<8>}
$$

where $\mathbf{X}_i$ and $\mathbf{Y}_i$ denote the i-th column of $\mathbf{X}$ and $\mathbf{Y}$, respectively; ω is a parameter to determine the weight between user-based and concept-based similarity.

4. EXPERIMENT RESULTS

We discuss the experimental results on real-world data. By using our framework, we have observed interesting stories from automated summarization of Flickr groups. For presentation purpose, we select one group to illustrate that our approach can (1) capture the dynamics of social activities, and (2) give a meaningful social media summary.

The group "The Flaming Lips" is used to demonstrate our technique. This group post and discuss photos about the rock band ("The Flaming Lips"), which represents as an example



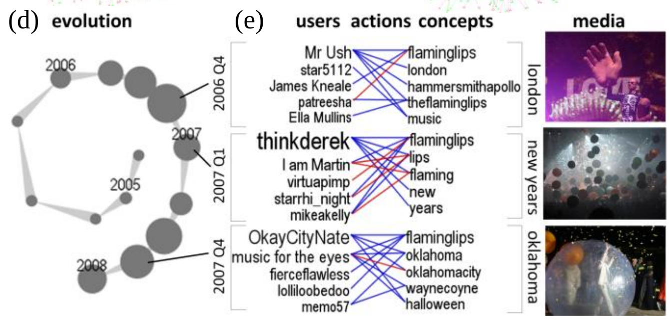


Figure 3: Our method generate social media summary by extracting temporally representative users, actions, concepts and media. Top 5 representative users/concepts/media on aggregated networks for (a) Q4-2006 (b) Q1-2007 and (c) Q4-2004. The semantics of colors and shapes are given in the figure. (d) Activity theme evolution – each bubble represents an activity theme at a time, with size proportional to the intensity of the activities. The distance between bubbles indicates the difference of consecutive themes.
(e) Representative activities extracted from the group "The Flaming Lips" An activity theme is represented by a set of users, concept terms and different actions (blue edges represent "post-on" actions and red edges represent "comment-on" actions).

of the rock fan culture. There are 1007 users involved in this group. The group data contains 2260 photos, with 1286 comments and 1072 unique tags associated with the photos.

Figure 3 (d) and (e) shows the social media summarization results automatically generated by our method. The summary captures important events that interest the group members, including band's tour (London, Q4-2006), concert (New Years Eve, Q1-2007) and honor (Oklahoma City, Q4-2007). This can be verified by cross-checking the band's history[i]. To show that the extracted users, concepts and media are temporally representative, we plot these extracted entities as nodes on the aggregated network (Figure 3 (a-c)). In the network we show the multiple relations (user-photo, user-comment, photo-tag and comment-photo) from Q4-2006 to Q4-2007, where photos posted in the same duration are clustered in corresponding boxes. Over time, the extracted representative entities appear to be a trajectory that spins around the aggregated network.

The evolution plot (Figure 3 (d)) highlights the interesting theme by the bubble size (proportional to the intensity the an activity theme) and the distance between consecutive bubbles (proportional to the similarity of theme concepts). The representative activities (Figure 3 (e)) show the related context (who, what and how). Our method summarizes rich and dynamic social context around media, which cannot be obtained by traditional methods (e.g. tag cloud).

5. CONCLUSION

We propose a method for summarizing and representing social media over time. In our framework, we formulate the summarization problem as extraction of representative activity themes. This summarization framework help identify who (users), what (concepts and media), how (actions) and when (time) to represent the collective activities. We extract activity themes using joint-action matrix factorization with time regularization. Our algorithm is able to differentiate action semantics by the type and the co-occurring time of actions. We show that our method is able to construct representative and meaningful group activity summary.

6. REFERENCES

[1] A. AGARWAL, et al. (2006). *Learning to rank networked entities*, SIGKDD, 14-23, 2006.

[2] R. KUMAR, et al. (2006). *Structure and evolution of online social networks*, SIGKDD, 611-617, 2006.

[3] Y. LIN, et al. (2008). *Facetnet: a framework for analyzing communities and their evolutions in dynamic networks*, WWW, 2008.

[4] X. WANG, et al. (2006). *Latent semantic analysis for multiple-type interrelated data objects*, SIGIR, 236-243, 2006.

[5] W. XI, et al. (2004). *Link fusion: a unified link analysis framework for multi-type interrelated data objects*, WWW, 319-327, 2004.

[6] S. ZHU, et al. (2007). *Combining content and link for classification using matrix factorization*, SIGIR, 487-494, 2007.

[i] http://en.wikipedia.org/wiki/The_Flaming_Lips; http://www.flaminglips.com