Interfaces For Networked Media Exploration And Collaborative Annotation

Preetha Appan Bageshree Shevade Hari Sundaram David Birchfield
Arts Media and Engineering Program
Arizona State University
Tempe, AZ 85251
email: {preetha.appan, bageshree.shevade, hari.sundaram, david.birchfield}@asu.edu

ABSTRACT

In this paper, we present our efforts towards creating interfaces for networked media exploration and collaborative annotation. The problem is important since online social networks are emerging as conduits for exchange of everyday experiences. These networks do not currently provide media-rich communication environments. Our approach has two parts – collaborative annotation, and a media exploration framework. The collaborative annotation takes place through a web based interface, and provides to each user personalized recommendations, based on media features, and by using a common sense inference toolkit. We develop three media exploration interfaces that allow for two-way interaction amongst the participants – (a) spatio-temporal evolution, (b) event cones and (c) viewpoint centric interaction. We also analyze the user activity to determine important people and events, for each user. We also develop subtle visual interface cues for activity feedback. Preliminary user studies indicate that the system performs well and is well liked by the users.

Categories and Subject descriptors

H.5.1 [**Multimedia Information Systems**]: *Artificial, augmented, and virtual realities,* H.5.4 [**Hypertext/Hypermedia**]: *Architectures, Navigation,* H.5.2 [**User Interfaces**]: *Theory and methods, User-centered design*

General Terms

Algorithms, Design, Human Factors

Keywords

Networked media, communication, collaborative annotation, media exploration, personalized media interaction

1. INTRODUCTION

In this paper we develop a system (ref. Figure 1) that allows for a network of users to explore the activities of a group of friends. The system incorporates a collaborative annotation and a media rich exploration framework. This is an emerging problem in several contexts: (a) sharing of media is important in online social-networks such as Friendster [2], where a set of friends share a set of multimedia experiences. (b) With the ready availability of digital still and video cameras, it has become easy to archive events in our daily lives. The popularity of social networks such as Friendster, as well the formation of moblogs suggests that users are willing to share personal media experiences online.

IUI'05, January 9-12, 2005, San Diego, California, USA.

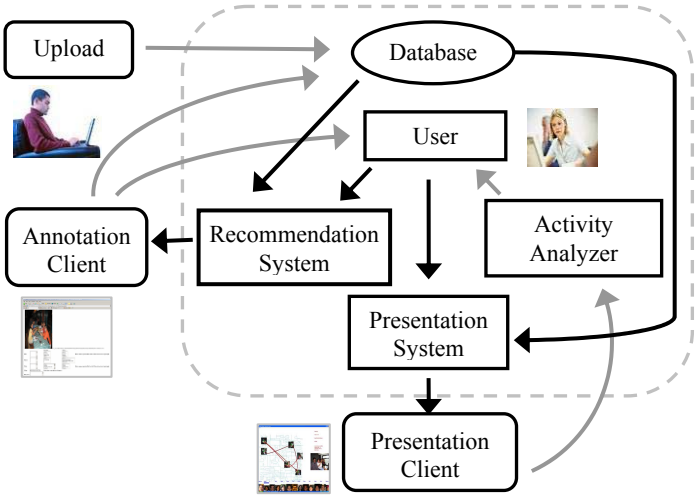

Figure 1: Overall system diagram. Users upload media to database and annotate using the recommendation system that incorporates user-context. The activity analyzer updates the user-context which in turn affects the presentation system. The dark arrows indicate data flow from server to client and the grey arrows, vice versa.

There has been prior work in creating collaborative annotation systems [7,9]. In [7], the authors explore a collaborative annotation system for mobile devices. There they used appearance based recommendations to suggest annotations to mobile users. In [9], the authors describe a collaborative annotation procedure for scientific visualization tasks, that can done remotely. A key innovation in our approach is to augment the feature based recommendation systems with a common sense toolkit [12], thus making the recommendations more useful.

There has been prior work in exploring image collections. Prior work in [10,11], provides a summary view of media through thumbnail display of images. Temporal arrangement of media can be explicitly queried for, and similarly other types of Boolean queries can be used to search for media effectively. However, these visualizations provide no information on spatio-temporal or semantic relationships amongst media. Displaying images as clusters around locations in a map as done in [13], gives an idea of how the media are distributed across space. In this work temporal relationships between images are lost. We now present our approach to both the collaborative annotation and networked exploration.

In our approach, our collaborative recommendation system consists of the following components: (a) media and its features,

(b) user / group context, (c) common sense based recommendations. The user annotates the images using a web-based interface. As the user begins to annotate images, the system provides personalized recommendations using a combination of low-level features and a common sense toolkit [12]. After the user has finished annotating an image, the system creates positive example image sets (or clusters) for the associated annotation words within each field (*who*, *when*, *where, what*). The clusters are based on annotation words/concepts entered by the users and not on automatic grouping of low-level features. These clusters will help the annotation process improve *for all users of the network.*

Our goal is to develop an interactive exploratory framework that enables easy and intuitive exploration of shared media, for a group of users in a social network, through *implicit* rather than *explicit* querying. Our interaction framework has two key components – (a) a visualization subsystem (b) an interaction analysis module. The user interacts with the shared media set, through three novel visualization schemes. Her interactions get captured by the interaction analysis module, which then sends information back to the presentation system. Then, the presentation system provides visual cues to the user that indicates how she interacts with other members. Preliminary user studies indicate that the annotation / exploration system is well liked by members of the network.

The rest of this paper is as follows. In the next section, we present a brief overview of our collaborative recommendation system. In Section 3, we shall describe the features extracted, and how we use ConceptNet. In Section 4, we describe our recommendation system algorithm in detail. In Section 5, we describe the web-based interface for collaborative annotation and show how collaborative annotation is related to networked exploration. Section 6, discusses our networked exploration framework visualizations. Section 7 discusses details on activity analysis and visual feedback. Finally, we present our experiments in Section 8 and conclusions in Section 9.

2. COLLABORATIVE ANNOTATION

The motivation behind collaborative annotation is to share the process of annotating shared personal media within a social network. This is intuitive as members of a small social network share the same media and participate in the same events. The shared personal media in our framework consists largely of images and audio associated with everyday events. This media needs to be annotated before being used in the interactive exploration framework.

Our system requires minimal *authoring* — the users provide meta-data for the images in the form of *who*, *where*, *when* and *what* fields. The system takes advantage of the fact that other users in the network have entered annotation for their media for shared events.

2.1 Context

User context models are crucial to collaborative annotation as they help to give personalized recommendations to each user. The dictionary definition of context is given as: *the interrelated conditions in which something exists or occurs.* These conditions could be the physical location, time, user's activity and past actions, environment etc [14].

In our system, the user's context model comprises (a) the initial user profile entered by the user which includes demographic information like age, background, hobbies/interests etc. (b) statistical information like number of images contributed in the shared social network and (c) usage statistics which includes the words she has used for annotation and their frequency. Maintaining a frequency count of annotations is intuitive as the media that is shared by the members of the social network consists of everyday events (people, places, what) that recur.

We also model the group context using (a) the images uploaded by all the members of the group and (b) the annotation words used by all the members of the group. The group context is therefore, the union of the user context of all the members of the group. Our system uses the group context to provide group recommendations using low-level features.

3. MEDIA ANALYSIS

In this section, we discuss the low-level features used and the use of semantics through ConceptNet.

3.1 Low-level features

In this section, we give a brief overview of the low-level features as well as feature-based image distance used in our recommendation system.

In this work, our feature vector comprises of color, texture and edge histograms. The color histogram comprises of 166 bins in the HSV space. The edge histogram consists of 71 bins and the texture histogram consists of 3 bins. We then concatenate these three histograms with an equal weight to get the final composite histogram. The low-level Euclidean feature distance d between two images i and j is now given as follows:

$$d(i,j) = \sqrt{\sum_{b=1}^{N} (h_i^b - h_j^b)^2} \qquad <1>$$

where N is the total number of bins in the final feature vector and h_i is the histogram for image i.

3.2 Media Semantics

Semantics are incorporated in our framework through the use of ConceptNet. ConceptNet is a large repository of commonsense concepts and its relations [12]. The repository represents twenty semantic relations between concepts like *effect-of*, *capable-of*, *made-of* etc. In our system, semantics are incorporated through the use of ConceptNet. Since the shared personal media consists of everyday events and activities; we believe that using a commonsense knowledge base such as ConceptNet will enhance the quality of the recommendations.

3.3 Semantic distance

In this section, we determine a procedure to compute semantic distance between any two concepts using ConceptNet. The ConceptNet toolkit allows three basic operations on a concept – (a) finding contextual neighborhoods that determine the context around a concept or around the intersection of several concepts, (b) finding analogous concepts, that returns semantically similar concepts for a source concept and (c) finding paths in the semantic network graph between two concepts.

In our system, we use the ConceptNet toolkit to determine the semantic distance between two given concepts e and f in the following manner:

Context of Concepts: Let us assume that the toolkit returns the contextual neighborhood sets C_e and C_f for the concepts. Then the context-based semantic distance $d_c(e,f)$ is now defined as follows:

$$d_c(e,f) = \frac{|C_e \cap C_f|}{|C_e \cup C_f|}, \qquad <2>$$

where | | is the cardinality operator.

Analogous concepts: Let us assume that the returned analogous concept sets are A_e and A_f . Then the semantic distance $d_a(e,f)$ based on analogous concepts is defined as follows:

$$d_a(e,f) = \frac{|A_e \cap A_f|}{|A_e \cup A_f|}. \qquad <3>$$

Number of paths between two concepts: Given two concepts, the system extracts the total number of paths between them as well as the number of hops in each path. The path-based semantic distance $d_p(e,f)$ is then given as follows:

$$d_p(a,b) = \frac{1}{N}\sum_{i=1}^{N}\frac{1}{h_i} \qquad <4>$$

where N is the total number of paths between concepts *e* and *f* in the semantic network graph of ConceptNet and h_i is the number of hops in path *i*.

The final semantic distance between concepts *e* and *f* is then computed as the weighted sum of the above distances. We have defined equal weight for each of the above distances. The final semantic distance is given as follows:

$$d(e,f) = w_c d_c(e,f) + w_a d_a(e,f) + w_p d_p(e,f). \qquad <5>$$

4. RECOMMENDATION ALGORITHM

In this section, we discuss the algorithm for the recommendation system. The goal is to provide recommendations as the user is trying to annotate images uploaded by her. Let us consider the following scenario. Let us assume that the user has entered an initial user profile and the system has been seeded with a few annotated images. When the user chooses to annotate an image, the system provides recommendations based on low-level features, user-context, group-context and ConceptNet.

As the user annotates images, the system creates positive example image sets for the associated annotations. The system forms clusters for each distinct annotation introduced in the system. These clusters grow in number and size as the users in the network annotate their shared media. Thus with a larger participation by the members of the group, the recommendations get more refined and reflect the group as a whole.

4.1 Algorithm Details

In this section, we present the details of the algorithm for the recommendation system. Figure 2 shows the flow chart of the algorithm. Let us assume that the user wishes to annotate an image *a* with the *who*, *where*, *when*, and *what* fields. Let us also assume that the database contains N clusters for annotations within each field.

4.1.1 Frequency based Personal Recommendation

When the user chooses image *a* for annotating, the system provides two kinds of recommendation list for each of the above mentioned fields – (a) personal recommendation list and (b) group recommendation list. The personal list is obtained from the frequency count of the annotation words used by the user. As the user annotates images, the system maintains a frequency count within each field for each annotation word used by the user for

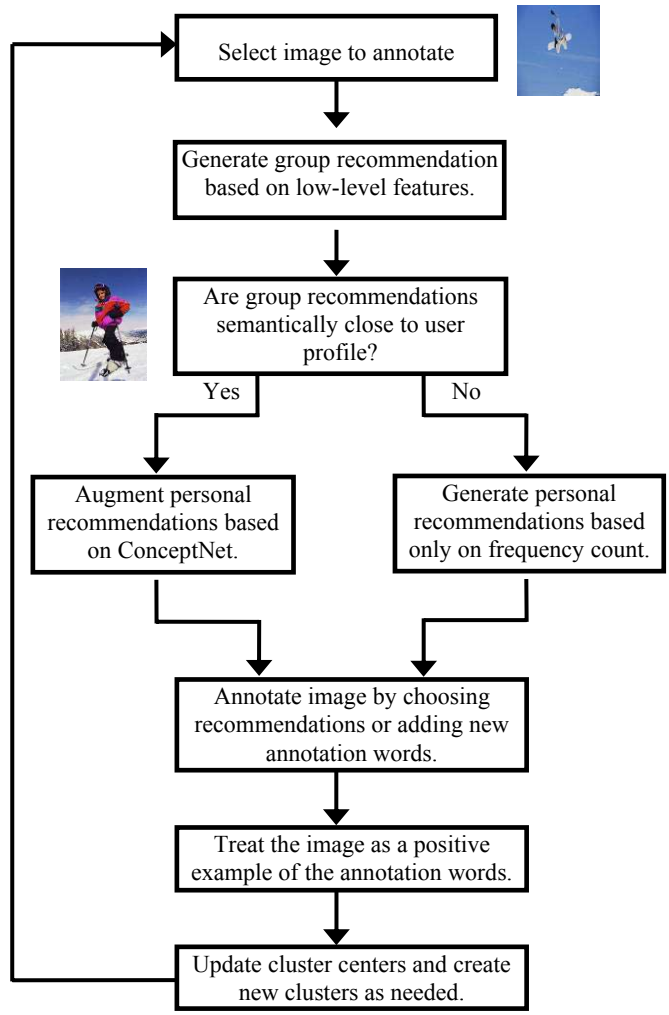

Figure 2: Flow Chart for Recommendation System Algorithm

annotating her images. The system then picks the three most frequently used words within each field to generate the personal list for each field.

4.1.2 Group Recommendation and Concept Filtering

The group recommendation for each field is obtained by computing the low-level feature distance between uploaded image

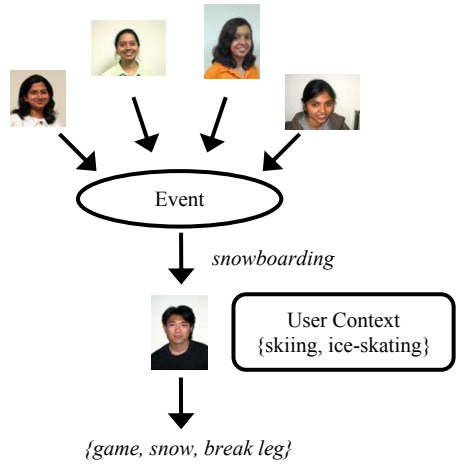

Figure 3: Filtering of group recommendation by user-context using ConceptNet.

and the cluster centers. The system then presents the top three closest cluster center words as recommendations in the group list.

The system also filters this group recommendation list by the user profile of the currently logged in user to get additional personal recommendations for the *what* field of the image. This is done by computing the semantic distance given by equation <5> between every concept in the user's profile and the concepts returned in the group recommendation list. When the semantic distance between the user profile concept and the group recommendation concept is less than one, the system uses the ConceptNet toolkit to

Table 1: Example recommendations obtained by filtering group recommendations with user-profile using ConceptNet. The arrow indicates the sequence of actions taken for the annotations to be personalized.

Photo	Group	User Profile	ConceptNet
	dinner	film	movie, people, chair
	cake	music	at party, activity, group, fun
	golf	tennis	game, ball, sports, activity, fun

get a list of concepts which are in the context of user profile concept biased by group recommendation concept. For e.g. Suppose the user profile concept contained the word *"skiing"* and the group recommendation contained the word *"snowboarding"*. We now use ConceptNet to get a list of concepts in the context of *"skiing"* which are biased by the context of the concept *"snowboarding"*. An example of such a list would contain concepts like *"game"*, *"snow"*, *"break leg"* etc. The system then picks the top five concepts in the list and adds it to the personal recommendation list for the *what* field. This is done for every concept in the user profile that is semantically close to the concept in the group recommendation list. This is shown in Figure 3. Table 1 shows the recommendations that are obtained by filtering group context using user-context.

4.1.3 Updating the System

When the user has annotated image *a* with the recommendations provided or by entering her own annotations, the system treats image *a* as a positive example of all the annotations associated with it. The system thus creates semantic clusters corresponding to all annotations that exist in the system. If the user has introduced a new annotation into the system, then system creates a new cluster for the annotation with only image *a* as the positive example. The system then computes the semantic distance between the newly introduced annotations and all the concepts in all the user profiles present in the system. This is done as a background process to ensure real time interactivity and to make the system scalable. The user can now continue annotating other images and increase the meta-data in the system.

5. THE ANNOTATION INTERFACE

In this section, we describe the web interface used for annotating images. Let us assume that the user has logged into the system and has created a session. The user now uploads some images to the shared media repository. When the user uploads the images, the system scales these images to a fixed size and computes the color, texture and edge histograms on these scaled images.

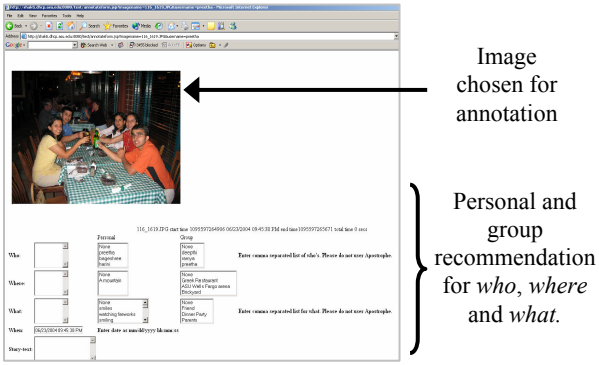

Figure 5: Web interface that provides users with personal and group recommendations that facilitates authoring of media.

When the user has uploaded the images, she is required to group them into *events* [3]. In our framework, events have the following properties associated with them: name, location, time, media elements (set of images, sounds, text), as well as participants. Once the user has uploaded the images, the system presents the user with all the images that she has uploaded in the system so far but has not grouped into events. The user can then group images and creates a new event or she can add images to an already existing event. After the user has created events, the user can now choose to annotate images. Then, the system presents the user with all the images that have been grouped by the user into events but have not been annotated. When she chooses an image for annotation, the system provides her with recommendation for the *who*, *where*, and *what* fields of the image. This is shown in Figure 5. The recommendations are provided based on the algorithm described in the previous section.

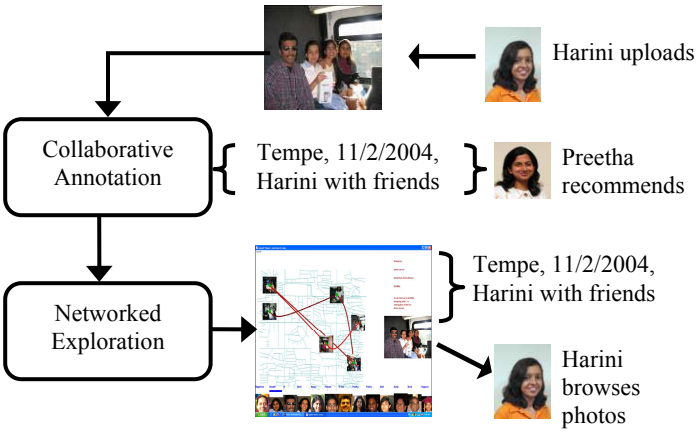

Figure 4: Media flow through Collaborative Annotation and Networked Exploration framework. Collaborative Annotation system helps users annotate media which can then be explored with Networked Exploration framework.

Our web interface also allows the user to add *viewpoints* to events. Viewpoint is defined as a personal narrative about an

event [3]. When the user chooses to add viewpoint, she is asked to select the event to which she wants to add viewpoints. When the event is selected, the system presents all the images which are grouped in that event but have been uploaded and annotated by other users of the network. The user can now choose an image and add her own personal narrative to that image. When the user annotates images, the system updates the cluster centers as described in the previous section. These updates get reflected in the recommendations in the next session. Our web interface also allows the user to browse photographs via the visualization tool. Figure 4 shows an example of how the collaborative annotation framework affects a specific user's event exploration.

6. NETWORKED MEDIA EXPLORATION

In this section, we describe the interactive visualization framework that we have developed, to allow users to explore events and activities of friends in their social network.

6.1 Design Goals

It is a challenging task to build a visualization framework that presents media to users in novel ways, as well as enables them to understand the relationships between people and events. We begin by enumerating our design goals, for this framework:

1. The framework should be interactive, and preferably web-based. This would require new visualizations and interaction mechanisms that differ from current query based photo browsing systems.
2. In addition to browsing, the framework should also allow users to perform certain tasks through implicit querying. Three such tasks are listed below. In the following, we shall use imaginary users – Alice and Bob):
   - *Event conditioning*: What has Alice done since a particular event? (e.g. party at Bob's house)
   - *Event support*: What was the sequence of events that led Alice to meet Bob at particular location? (e.g. Alice meets Bob at the mall).
   - *Interesting events*: What are the set of events of interest to Alice? Note that interest is user context dependent.
3. The authoring environment should place minimal burden on the user. This issue was addressed by our annotation framework.
4. The system should be able to provide feedback to users about their social network, by analyzing user interaction.
5. The communication between users should be two – way, i.e. users should be able to provide feedback to events authored by other users through the system.

We have dealt with design goals one and two in prior work [3]. We are addressing the other three goals in this paper.

6.2 Media Exploration framework

We now present our media exploration framework. While we had presented a preliminary interface in prior work [3], it now significantly augmented by real-time user activity analysis, a new novel visual feedback cues.

6.2.1 Spatio-temporal Evolution

This visualization addresses the task of *event conditioning*, as well the *interesting events* task described in the design goals section. Prior work in [11] show collections of digital media organized either by space or time. Our visualization introduces the idea of evolution of media through both space and time. We use map data available in XML format [1], on which events unfold over time and space.

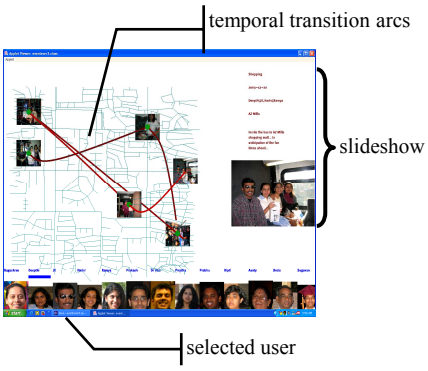

Figure 6: Spatio-Temporal Evolution. For every user, the system shows the events that the user participates in over space and time.

The interaction begins by the user selecting a friend. The system responds by showing the spatio-temporal slideshow of events, over time and by location (e.g. GPS).The temporal transitions between locations is indicated using arcs, and saturation of the arc color is used to indicate time. This is illustrated in Figure 6.

This visualization additionally addresses the problem of visualizing *interesting events*. It is reasonable to assume that users are interested in knowing who in their group of friends went to the same places as themselves. Hence, when the user moves her mouse over a specific event location, then all the friends of the selected user who *participated in the same event* are highlighted in green. The other friends of the selected user, who participated in *other events at the same location,* are highlighted in purple. Users can then click on highlighted members, to see a spatio-temporal slideshow, starting at that event of interest.

6.2.2 Event Cone

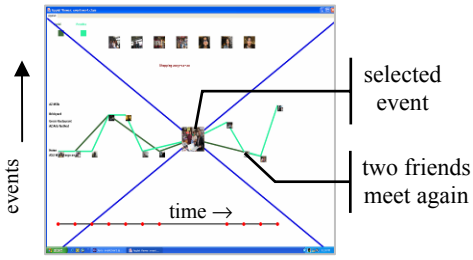

Figure 7: This is a snapshot of all the events associated with the selected users. The events are ordered along the vertical axis and the horizontal axis shows time.

The event cone visualization addresses the task of *event support*. This visualization attempts to provide a summarized snapshot of the event relations over time and space. We call such a summary – *event cone*, as seen in Figure 7. Note that, while the spatio-temporal evolution does enable users to understand what happens to a person over time, it doesn't allow a proper understanding of the event relationships.

When the user selects her event of interest in the spatio-temporal browser, an event cone of that event is shown, which shows the

timeline of events of all the participants of that event, prior to and after that event. Each participant has a distinct path and is individually colored. When the paths intersect, they imply that the friends meet. Users can continue interacting with this visualization by choosing other events from the currently displayed event cone, which will redraw the event cone with the chosen event as the center.

6.2.3 Viewpoint Based Evolution

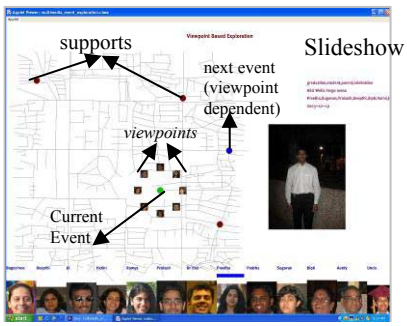

Figure 8: The user can dynamically change the viewpoint, thus changing the slideshow associated with the viewpoint, as well as the future event.

We present our visualization scheme that allows for viewpoint (i.e. personalized narratives) based exploration of events. The idea that users desire *agency* [5] – users proactively interact with the system to control the sequence of events they see , is incorporated in this visualization. Initially, circles representing events associated with the selected user appear on locations in the map.

As seen in Figure 8, the viewpoints associated with an event, are displayed around it. Also the set of preceding events for the current event and one succeeding event are shown as red and blue circles respectively. The set of preceding events for each event are calculated as the union of all events that precede each of the viewpoints of the event. The idea behind this is to provide *context* for the user to understand how the event happened, through each viewpoint's perspective. Now users can select view points around the current event, to see a slideshow of media associated with the event (and the selected viewpoint). Also the succeeding event for each event, changes according to the chosen viewpoint. Thus, the user can continue to explore any event/viewpoint of their interest. Thus this is a very dynamic environment that allows users to explore events/viewpoints in a non linear manner.

7. ANALYSIS AND FEEDBACK

In [8] , the authors discuss why it is vital for the design of user interfaces for online communities to provide the means to communicate social cues and information. We attempt to provide such cues to users about their social network as they interact with our system.

7.1 Activity

Our system allows users to upload media about their events and experiences, and share them with friends in their social network, through the visualization interface. Hence the term "activity" has two different meanings associated with it – (a) User activity by *participation* and (b) User activity by *interaction.* The former refers to actual social activity i.e. users participating in events and sharing them through our system with other friends. This information is limited to the extent that the user uploads to the system. The latter refers to the way users interact with the system, i.e. how they actually view the media in the system, who they pay attention to, etc. This can be captured by keeping detailed logs of user interaction in our system. We would like to use available information about user activity (i.e. by both *participation* and *interaction*) to provide visual feedback to users through the system. Two questions that we attempt to answer for each user through activity analysis are: (a) *Who are the people (within my network), that I interact the most with?* and (b) *Which are the most important events to me?* In the following sub-section, we describe the capture and computation of user activity.

7.2 Activity Capture

The three visualization schemes described in the previous section complement each other. The functionality of each visualization will determine the type of user activity data we collect from them. We shall describe the information collected from each in detail in the following sections. Note that the activity results will be in general different for each member of the network.

7.2.1 Spatio-Temporal Evolution

Since this visualization provides users choice to watch the activities of the social network members they are interested in, we conjecture that the *time spent by the user on each member* is a good indicator of the user interest in that member. Since this time is biased by the number of photographs uploaded by that member, we use a normalized time score ($\bar{t}_{ST}$). See <6>. To calculate the score of the user with respect to each member of the social network, we define an exponential growth function on the normalized time spent on each member, by the user. The intuition behind using a function that varies exponentially with time is that concepts added to memory also grow at an exponential rate. The equation used is given below:

$$\bar{t}_{ST}(U,P_i) = \frac{n(P_i)}{N(P_i)} \bullet t_s(U,P_i), \qquad \text{<6>}$$
$$W_{ST}(U,P_i) = 1 - e^{-\alpha \bar{t}_{ST}},$$

where W_{ST} (U, P_i) is the weight assigned to participant P_i with respect to the currently logged user U in the spatio temporal visualization. $n(P_i)$ is the number of photographs of participant P_i's spatio –temporal slideshow that the user saw, $N(P_i)$ is the total number of photographs for P_i, and t_s is the total time spent by the user on P_i .

7.2.2 Event-Cones

This visualization allows users to see a snapshot of how an event occurred. Thus, in this visualization users focus on the event itself, rather than its participants. Consequently, we conjecture that the time spent by users viewing each event is a good indicator of that event's importance.

We proceed as follows. We record the time spent on the event by the user, and apply the exponential growth equation to get the importance score of each event for the given user. Since every event in the event cone is represented by a single photograph, we do not need to normalize the time score here.

$$W_{EC}(U,E_i) = 1 - e^{-\alpha te_i} \qquad \text{<7>}$$

Where $W_{EC}(U, E_i)$ is the score of the event E_i for the given user U , in the event cone visualization , te_i is the time spent by the given user on event E_i in the event cone visualization.

7.2.3 Viewpoint Centric Visualization

While browsing in this visualization, users can either follow all events of a particular user, or choose a completely non-linear path, by following different viewpoints of arbitrary events. Hence, user activity in this visualization can give us information about both *event importance* and *importance of the members to the user*.

We measure *event importance* as a fraction given by the ratio of number of viewpoints of that event that have been seen, to the total number of viewpoints for that event. Thus, events, for which all viewpoints have been viewed by the user, would have the maximum weight.

$$W_{VP}(U,E_i) = \frac{NV_{seen}(E_i)}{NV_{total}(E_i)} \qquad <8>$$

Where $W_{VP}(U,E_i)$ is the weight of the event E_i for user U in the viewpoint centric visualization, $NV_{seen}(E_i)$ is the number of viewpoints of that event that have been seen , and $NV_{total}(E_i)$ is the total number of viewpoints of that event.

For *member importance*, we record the aggregate of time spent by the user in browsing that participant's viewpoint and normalize it. We also use a similar exponential function as in <6> to arrive at the weight of the friend with respect to the given user, described as follows.

$$\overline{t_{VP}}(U,P_i) = \frac{n(P_i)}{N(P_i)} t_s(U,P_i), \qquad <9>$$
$$W_{VP}(U,P_i) = 1 - e^{-\alpha \overline{t_{VP}}},$$

Where $W_{VP}(U, P_i)$ is the weight assigned to participant P_i with respect to the currently logged user U in the viewpoint centric visualization. $n(P_i)$ and $N(P_i)$ are the total number of photographs of P_i's viewpoint that the user saw, and total number of photographs for P_i's viewpoint respectively, and t_s is the total time spent by the user on P_i .

7.3 Aggregate Activity Scores

We describe how we arrive at aggregate event importance scores and user importance scores, from the data collected on user activity. Since we calculate both event scores and people scores for a given user from different visualizations, we need a way to compare the scores across visualizations. Hence we perform another normalization step, where we obtain the mean and standard deviation of each score per visualization, across sessions. Let us assume that we are calculating the normalized score for the viewpoint centric visualization. We use the raw scores to obtain a re-weighted score, as follows:

$$E_n(i) = \frac{E(i) - \mu_e}{3\sigma_e}, \qquad <10>$$
$$P_n(m) = \frac{P(m) - \mu_p}{3\sigma_p},$$

Where E_N corresponds to the score of the event i, μ_e is the mean of all event scores for that visualization, and σ_e is the standard deviation of the event scores distribution over all sessions. The corresponding normalized people score P_N is also calculated in the same way as given above. Doing this step ensures that the significance of scores obtained across visualizations is comparable. We repeat this calculation for the other two visualizations as well.

To calculate event importance scores for each user, we use data about user activity collected from two visualizations – the event cone and the viewpoint centric visualization. The average of the normalized event importance from these two visualizations ($E_{Ne}(i)$, $E_{Nv}(i)$), gives the final event importance. Here $E_{Ne}(i)$ and $E_{Nv}(i)$ corresponds to event scores from event cone and viewpoint visualizations respectively, computed using <10>.

To calculate importance scores per user, with respect to all the other users in the network, we use activity information collected from spatio-temporal evolution and viewpoint centric visualizations. Similar to event importance, participant importance is also calculated as average of the normalized scores obtained from both the above mentioned visualizations ($P_{Ns}(m)$ *and* $P_{Nv}(m)$).

7.4 Visual cues

We describe in this section, how we use importance scores obtained above, to provide visual feedback to each user, through the system. The purpose of this is two fold – (a) providing feedback to the user about her interaction, gives her interesting insights and (b) the system can learn from user behavior and adapt visualization and interaction to suit user preferences and resource constraints.

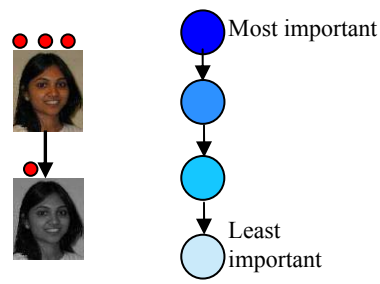

Figure 9: The above figure shows visual cues that indicate important people and important events. Color indicates importance in people, and saturation indicates importance in events. Red dots above photographs also indicate participant importance.

Our goal is to provide subtle visual feedback to users interacting with the system, rather than explicit display of the information calculated. We are motivated by ethnographic studies on Friendster [4]. We use color rather than text, as the primary form of providing feedback, guided by general principles of graphic design given in [15].

We use varying levels of color to indicate importance levels amongst friends in the network, for each user. Photographs of important friends in the network appear the brightest, and friends with whom the user's activity index is low, appear grayed out. The contrast perceived when the network members photographs vary from color to grey scale, communicates their importance or rank very well. We additionally use red dots above member photographs to indicate the top five important people to the user. See Figure 9. To indicate important events to users as they are browsing in the system, we use color saturation. The saturation of the color of the event being rendered increases (as the events become less important. Thus the most important events (according to that user's interaction history), would appear darkest. (See Figure 9).

8. EXPERIMENTS

We had conducted some preliminary experiments in [3] to only evaluate our visualizations. Users were asked to evaluate each of the three visualization schemes – i.e. viewpoint based evolution, spatio-temporal evolution and event cones. We found our visualization schemes were well liked.

In order to evaluate the current system, we asked a group of nine members, seven students and two faculty members to participate in the evaluation. Participants could use the web interface to upload their everyday personal media, or send photos to the system through email from their cell phone. The participants interacted with the annotation and presentation systems for a period of one week and recorded their observations of participation in the social network. We opted for a observer-participant user-study analysis as opposed to a pilot user study, to understand the *practice* of networked media use.

Most users found the idea of collaborative annotation to be interesting. Some of the positive comments are as follows: "The recommendations make things easier for annotation and are often appropriate ", "recommendations of names, using frequency worked well , as images related to a single event are uploaded one after another at one go." Comments about the visualization interface – "the event cone visualization is awesome – really ideal for seeing whose paths crossed when and where" , " viewpoint centric visualization is a good idea, as many times we see a photographs of friends and would like to add your viewpoint about them" .

Some of the negative feedback were –"the slideshow in the spatio – temporal evolution was a little too fast ", "the event cone becomes cluttered when there are more photographs", "we should not be forced to create new events, for events that occur periodically, that mean the same thing but at different time spans – e.g. driving to work, taking dance lessons." We found these observations very helpful and encouraging – we also plan on conducting a long-term user study to gain further insight.

9. CONCLUSIONS AND FUTURE WORK

We now present our conclusions. In this paper, we described our system to enable users in a social network communicate their everyday experiences. Our framework provides (a) an easy to use web interface to upload media, b) a novel recommendation system to help users author the uploaded media, (c) presentation and interactions schemes to share and visualize the media and (d) visual cues to provide feedback to users about their interaction. Our collaborative annotation system uses media features, user-context, group-context and a common sense knowledge base (ConceptNet), to provide recommendations that enable users to author shared media. Our presentation schemes enable users in the social network to visualize and interact with the shared media. The activity analysis subsystem provides visual cues that reflect users' interaction with each other. Our initial study indicates that our interfaces were well liked. We plan to address issues such as allowing users to structure events and create semantic relationships between them. We also plan to conduct extensive user studies to evaluate the system further as well as add more visualizations and use Support Vector Machines [6] to improve feature based group recommendations.

10. REFERENCES

[1] *TIGER maps* http://imagemaps.mle.ie/.

[2] *Friendster* http://www.friendster.com.

[3] P. APPAN and H. SUNDARAM (2004). *Networked event exploration and interaction summarization*, Proc. ACM Multimedia 2004, New York, New York, also AME TR-2004-10, Oct. 2004.

[4] D. BOYD (2004). *Friendster and Publicly Articulated Social Networks*, Conference on Human Factors and Computing Systems, Vienna,

[5] K. M. BROOKS (1997). *Do story agents use rocking chairs? The theory and implementation of one model for computational narrative*, Proceedings of the fourth ACM international conference on Multimedia, 317 - 328,

[6] N. CRISTIANINI and J. SHAWE-TAYLOR (2000). An introduction to support vector machines : and other kernel-based learning methods. Cambridge, U.K. ; New York, Cambridge University Press.

[7] M. DAVIS, N. GOOD and R. SARVAS (2004). *From Context to Content: Leveraging Context for Mobile Media Metadata*, MobiSys 2004 Workshop on Context Awareness, Boston, MA, 2004.

[8] J. S. DONATH (1997). *Inhabiting the virtual city:The design of social environments for electronic communities*, MIT, PhD Thesis.

[9] S. E. ELLIS and D. P. GROTH (2004). *A collaborative annotation system for data visualization*, Proceedings of the working conference on Advanced visual interfaces, ACM Press, 411--414,

[10] J. GEMMELL, G. BELL, R. LUEDER, et al. (2002). *MyLifeBits: fulfilling the Memex vision*, Proceedings of the 10th ACM international conference on Multimedia, Juan Les-Pins, France, pp. 235-238, Dec. 2002.

[11] H. KANG and B. SHNEIDERMAN (2000). *Visualization Methods for Personal Photo Collections : Browsing and Searching in the PhotoFinder*, (ICME2000), New York:, IEEE, 1539-1542,

[12] H. LIU and P. SINGH (2004). *ConceptNet: a practical commonsense reasoning toolkit.* BT Technology Journal **22**(4): pp. 211-226.

[13] C. SHEN, B. MOGHADDAM, N. LESH, et al. (2001). *Personal Digital Historian: User Interface Design*, In Proceedings of Extended Abstracts of Human Factors in Computing Systems (CHI 2001), ACM Press,

[14] H. SRIDHARAN, H. SUNDARAM and T. RIKAKIS (2003). *Context, memory and Hyper-mediation in Experiential Systems*, 1st ACM Workshop on Experiential Telepresence, in conjunction with ACM Multimedia 2003,, Berkeley CA, AME-TR-2003-02, Nov. 2003.

[15] A. WHITE (2002). The elements of graphic design : space, unity, page architecture, and type. New York, Allworth Press.