

# Unified Entity Search in Social Media Community\*

Ting Yao  
City University of Hong Kong  
Kowloon, Hong Kong  
tingyao.ustc@gmail.com

Chong-Wah Ngo  
City University of Hong Kong  
Kowloon, Hong Kong  
cscwngo@cityu.edu.hk

Yuan Liu  
Ricoh Software Research  
Center Co., Ltd.  
Beijing, P. R. China  
yuanliu.ustc@gmail.com

Tao Mei  
Microsoft Research Asia  
Beijing, P. R. China  
tmei@microsoft.com

## ABSTRACT

The search for entities is the most common search behavior on the Web, especially in social media communities where entities (such as images, videos, people, locations, and tags) are highly heterogeneous and correlated. While previous research usually deals with these social media entities separately, we are investigating in this paper a unified, multi-level, and correlative entity graph to represent the unstructured social media data, through which various applications (e.g., friend suggestion, personalized image search, image tagging, etc.) can be realized more effectively in one single framework. We regard the social media objects equally as “entities” and all of these applications as “entity search” problem which searches for entities with different types. We first construct a multi-level graph which organizes the heterogeneous entities into multiple levels, with one type of entities as *vertices* in each level. The *edges* between graphs pairwise connect the entities weighted by *intra-relations* in the same level and *inter-links* across two different levels distilled from the social behaviors (e.g., tagging, commenting, and joining communities). To infer the strength of intra-relations, we propose a *circular propagation* scheme, which reinforces the mutual exchange of information across different entity types in a cyclic manner. Based on the constructed unified graph, we explicitly formulate entity search as a global optimization problem in a unified Bayesian framework, in which various applications are efficiently realized. Empirically, we validate the effectiveness of our unified entity graph for various social media applications on million-scale real-world dataset.

## Categories and Subject Descriptors

H.3.3 [Information Search and Retrieval]: Retrieval models

## General Terms

Algorithms, Performance, Experimentation.

\*This work was performed when Ting Yao was visiting Microsoft Research Asia as a research intern.

## Keywords

Entity search, social media community, friend suggestion, image tagging, personalized image search.

## 1. INTRODUCTION

The Web is organized by entities. Entity search, that is, returning “object” (e.g., people, media objects, locations, products, organizations, etc.) in response to users’ information needs, has received considerable attention recently in various research communities.

In social media communities, such as Flickr, YouTube and Facebook, people are sharing their experiences and interests through substantial amounts of various contents, such as photos, videos, people, tags, geo-locations and documents, which are referred to *entities* in this paper. In addition, these contents are strong correlated by a wide variety of social behaviors, e.g., tagging, commenting, and joining communities. With the large-scale, heterogeneous and correlative sets of data available, there has been increasing expectation for mechanisms that organize and search these entities effectively and efficiently in social networks. To motivate, consider following scenarios, for user Henry who likes taking photographs at the seaside.

**Scenario 1 (friend suggestion):** Henry wants to find friends who have the similar interests to himself. To begin with, he may use the keywords like “seaside” as the query to find some photos in Flickr, and then he needs to sift through the photostream (all photos shared by a user) of photographers one by one and dig out the interesting ones as his friends. This overall process can be unnecessarily time consuming.

**Scenario 2 (image (geo-)tagging):** Henry sees a beautiful picture when he browses a webpage and he wants to know where this picture was taken. He may formulate some descriptive words to search similar images via web image search engines and then read through the surrounding words, but it is hard to find the duplicate images in this way due to the sheer volume of Internet images.

**Scenario 3 (personalized image search):** Henry wants to search photos with sunrise scene. He may use “sunrise” as the query in a web image search engine and many images with sunrise scene are returned. But Henry prefers photos taken by the seaside, to get his expected photos he might have to go through all the images, or formulate the other

keywords for hitting promising search images. This could be a very laborious process.

In these scenarios, like every user in many similar situations, Henry is looking for a particular type of *entities* not pages as “relevant documents” to read. This demand has challenged us to face a dilemma: on one hand, the proliferating data on the Internet leads to surge of research activities in social media search, for instance, personalized image search [12][15][19], tag ranking [10], Yahoo people search [20], friend suggestion [8][14][22], and ImageCLEF (Image Retrieval in CLEF) [5]. All of these social media searches are independent and each of them is designed especially for searching a certain type of entities; on the other hand, in real-world social media communities, these existing entities always influence each other in explicit or implicit way. As in scenarios mentioned above, people always look for entities by using the other types of entities; that is to say, the various entity searches are essentially closely related, and the common goal is to discover the relationships among these entities. Therefore, it is more natural and efficient to develop a unified framework which can accomplish these entity searches in a single way, which directly motivates our work in this paper.

Towards this goal, we propose a multi-level graph model to integrate all the entities together. The relevance of an entity for a given query is estimated by the cost of nearest navigated route. As shown in Fig.1(b), the red lines with arrows present a navigated route from query  $\mathbf{q}$  to entity  $v_k^b$  (will be described in Definition 1) in a three-level graph. The cost comes from two sources, i.e., *intra-relation* between every two same types of entities in each level of graph, and *inter-link* across different types of entities in every two levels of graphs. To leverage the two sources of the route cost, the entity search is formulated from the probabilistic perspective in a Bayesian framework. The intra-relations of entities within the searching pool are modeled as a conditional prior which indicates the ranking score consistency between similar entities; and the inter-links are modeled as a likelihood which reflects the correspondence of the query and entities in searching pool. In the Bayesian framework, the entity search is formulated as maximizing the product of the conditional prior and the likelihood. The foundation of the unified entity search is the proposed *circular propagation* which iteratively refines the intra-relations of entities, through exchanging information across different entity types in a cyclic manner. Instead of exploring similarity metrics used in existing entity search, the procedure encourages interaction among multiple entities to seek for consensus that are useful for entity search. Fig.1 illustrates the unified entity search framework based on the proposed integral multi-level graph.

The main contributions of this work can be summarized as follows:

- We propose to integrate various types of entities in a multi-level graph model, upon which multiple social media applications can be implemented in a single way.
- We propose the circular propagation to exchange information across different entity types in a cyclic manner. In this way, the relationships of various entities are mutually reinforced and refined.
- We explicitly formulate entity search as a global optimization problem in a Bayesian framework, which

well leverages the various relationships in social media community.

The remaining sections are organized as follows. Section 2 describes the related work. Section 3 presents the integral multi-level graph construction including circular propagation and its convergence property, while Section 4 formulates the problem of unified entity search and its solution over the integral graph. Section 5 takes a popular photo-sharing site as an example to detail the entity graph constructed from a social media community. Section 6 provides empirical justifications, followed by the conclusions in Section 7.

## 2. RELATED WORK

We briefly group the related works into two categories: search in social media community and multi-modality fusion. The former draws upon research in searching over extracted entities in social media community, and the later learns the contribution of a modality (entity) in social search applications.

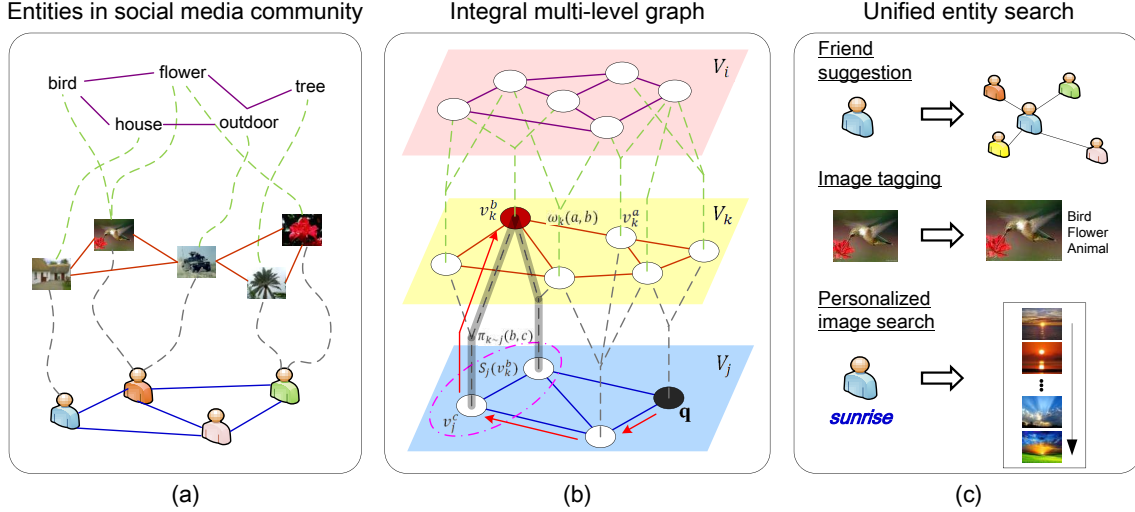
### 2.1 Search in Social Media Community

There exists rich research on search in social media community, such as friend suggestion (user search), image tagging (tag search) and personalized image search (image search).

Finding potential friends with similar interest will improve photo-sharing and browsing experiences in the social community. Roth *et al.* suggest friends based on an implicit social graph, which is formed by user interaction with contacts and groups of contacts [14]. In another work by Li *et al.* [8], user similarity is firstly mined by using location history and then potential friends are recommended. As an online people search service, Yahoo people search [20] is to find friends and family with whom people have lost touch. People can employ it to search for friends using name and location or phone numbers, or search using email addresses.

Attaching textual or semantic linkage to images will significantly facilitate Web image search and organization. Li *et al.* [7] present a real-time automatic image tagging method. The method takes as input a set of labeled images and tries to learn which low level visual features correspond to higher level semantic labels. The mapping can then be applied to annotate the new unlabeled images. Similar in spirit, Chen *et al.* [1] propose a tag recommendation approach that directly predicts the possible tags with models learned from training data. Therefore, the approach can only recommend the tags from a predefined set. Later in [10], Liu *et al.* first estimate relevance scores for the tags to images based on probability density estimation and then a random walk over a tag similarity graph is performed to refine the relevance. Based on the refined relevance, the associated tags can be ranked in a descending order of their relevance. ImageCLEF (Image Retrieval in CLEF) [5] aims to analyze a collection of Flickr photos in terms of their visual and/or textual features in order to detect the presence of one or more concepts. The detected concepts can then be used for the purpose of automatically annotating the images or for retrieving the best matching images to a given concept-oriented query.

Generating the search results according to the modified user search intents is proved to be an effective mean for enhancing Web search experience. Lu *et al.* [12] utilized a co-clustering method to extract latent interest dimensions, and rerank the images by coming latent interest based user



**Figure 1: Unified entity search framework** (For better viewing, please see original color pdf file). (a) Entities in social media community, such as tags, photos and users. (b) Integral multi-level graph with three types of entity sets, i.e.,  $V_i$ ,  $V_j$  and  $V_k$ : solid colorful lines in each level of graphs indicate the *intra-relations* of entities of the same types, while dash gray lines indicate the *inter-links* between entities of the different types; black vertex is the query  $q$  of entity type  $j$  ( $\mu_j = 1, \mu_i = 0 (i \neq j)$ ) and red vertex is the random entity  $v_k^b$  to be ranked in the searching pool  $V_k$ . Shadow lines denote the links between entity  $v_k^b$  and the subset of entities in  $V_j$ , enclosed in the purple ellipse. Red lines with arrows present a navigated route from query  $q$  to entity  $v_k^b$ . (c) Based on the integral multi-level graph, various entity searches can be formulated within a unified framework.

preference and query relevance. Another typical work is performed by Xu *et al.* [19], in which the overall ranking score is not only based on term similarity matching between the query and the documents but also topic similarity matching between the user’s interests and the documents’ topics.

In short, while these approaches focus on the mining of various entities for different social media search applications, the interaction among entities is not exploited. Our work in this paper contributes by studying not only historically exploring interaction (or consensus) among all the entities, but also integrating all the social media search applications in a unified framework.

## 2.2 Multi-modality Fusion

Multi-modality fusion on social media community has been intensively studied in recent years. Sang *et al.* [15] propose a tensor factorization model that integrates users, images and tags together for annotation prediction. With this, personalized search can be easily formulated by user-specific topic modeling. In [22], Yao *et al.* characterize the user relationship by combining the visual and geographic information in a contextual graph modeling. Negoescu *et al.* present to discover hypergroups in Flickr, i.e., communities consisting of groups of Flickr groups, while such hypergroups could serve to enrich Flickr’s traditional group search [13]. Zheng *et al.* provide a tensor decomposition-based group recommendation model to combine semantic tags with social relation and to help people more easily engage in group activities [24]. Lin *et al.* model the community discovery through analysis of time-varying and multi-relational data in rich social media community [9]. These approaches, nevertheless, deal with the multiple modalities independently and the interaction among modalities is not exploited. In another work by Zhuang *et al.* [26], social strength is measured in a two-

stage learning framework. The first stage is to learn the optimal fusion weights of the multiple proximity graphs for a kernel target alignment; with the learnt kernel, the second stage is to estimate the social strength by a kernel-based learning to rank approach. However, it is still implemented in supervised way, which needs expensive human effort.

## 3. INTEGRAL MULTI-LEVEL GRAPH

### 3.1 Notation

To make our formulation more clear, we firstly define a few terms below and some are illustrated in Fig.1(b), an exemplary three-level graph.

**DEFINITION 1:** Represented by multi-level graph model, the social network  $\mathcal{G} = \{G_k : k = 1, 2, \dots, N\}$  consists of  $N$  graphs and each graph is represented by  $G_k = (V_k, E_k)$ , where  $k$  is the entity type id,  $V_k$  and  $E_k$  are sets of vertices and edges. Specifically,  $V_k = \{v_k^1, v_k^2, \dots, v_k^{Z_k}\}$ , where  $Z_k$  is the number of entities in  $V_k$ ; and each edge in  $E_k$  connects two vertices, with weight function  $w_k : E_k \rightarrow \mathbb{R}^+$  mapping edges to nonnegative real-valued weights, here we call it *intra-relation*. For instance,  $w_k(a, b)$  measures the *intra-relation* strength between entities  $v_k^a$  and  $v_k^b$ . The affinity matrix  $\mathbf{W}_k$  is defined by taking  $w_k(a, b)$  as its  $(a, b)$ -element.

**DEFINITION 2:** The link set  $\mathcal{L} = \{L_{k \rightarrow j} : j, k = 1, 2, \dots, N\}$  consists of links mutually connecting entities from two levels of graphs. Of entity set  $V_j$ , the subset  $S_j(v_k^b) = \{v_j^1, v_j^2, \dots, v_j^M\}$  contains  $M$  entities, which are linked to the entity  $v_k^b$  in entity set  $V_k$ , i.e.,  $S_j(v_k^b) \subseteq V_j$ . The links are weighted by link function  $\pi_{k \rightarrow j} : L_{k \rightarrow j} \rightarrow \mathbb{R}^+$  mapping links to nonnegative real-valued weights and we describe it as *inter-link*. For instance,  $\pi_{k \rightarrow j}(b, c)$  measures the

inter-link strength between entity  $v_k^b$  and  $v_j^c$ . For self-link, we set  $\pi_{k \sim k}(\cdot, \cdot) = 1$ .

### 3.2 Circular Propagation

To facilitate interaction among different entity relationships, we construct an integral graph, over which the circular propagation is proposed to mutual reinforce each other. Suppose there are  $N$  entity types, and thus we can have  $N$  views of relationship among these entities by constructing a  $N$ -level of affinity graphs. In each level of graphs, the vertices and edges represent the entities and intra-relations of them ( $w_k$ ), respectively. The entities of different types are mutually linked with inter-link ( $\pi_{k \sim j}$ ). In this way,  $N$  runs of propagations can be conducted separately on each graph to refine the intra-relation of  $N$  entity type. The spirit of circular propagation is to encourage entity interaction in a way that an entity starts propagation by using the affinity graph of another entity, while preserving its original estimated intra-relation measures. By arranging the  $N$  entities linearly, such that the  $k$ -th order entity uses the graph of the  $(k-1)$ -th order entity, this forms a circular ring that each entity, once completing propagation, will propagate the new result to influence its next linked entities.

Let  $w_k(a, b) \equiv w_k^{ab}$ ,  $\pi_{k \sim j}(b, c) \equiv \pi_{k \sim j}^{bc}$ , the circular propagation is formulated as following:

$$\begin{cases} w_1^{ab(t)} = \theta_1 \sum_{ij} \pi_{1 \sim N}^{ai} w_N^{ij(t-1)} \pi_{1 \sim N}^{bj} + (1 - \theta_1) w_1^{ab(0)} \\ w_2^{ab(t)} = \theta_2 \sum_{ij} \pi_{2 \sim 1}^{ai} w_1^{ij(t)} \pi_{2 \sim 1}^{bj} + (1 - \theta_2) w_2^{ab(0)} \\ \dots \\ w_k^{ab(t)} = \theta_k \sum_{ij} \pi_{k \sim k-1}^{ai} w_{k-1}^{ij(t)} \pi_{k \sim k-1}^{bj} + (1 - \theta_k) w_k^{ab(0)} \\ \dots \\ w_{N-1}^{ab(t)} = \theta_{N-1} \sum_{ij} \pi_{N-1 \sim N-2}^{ai} w_{N-2}^{ij(t)} \pi_{N-1 \sim N-2}^{bj} + (1 - \theta_{N-1}) w_{N-1}^{ab(0)} \\ w_N^{ab(t)} = \theta_N \sum_{ij} \pi_{N \sim N-1}^{ai} w_{N-1}^{ij(t)} \pi_{N \sim N-1}^{bj} + (1 - \theta_N) w_N^{ab(0)} \end{cases} \quad (1)$$

where the superscript  $(t)$  denotes the iteration, and the tradeoff parameter  $\theta_k (0 \leq \theta_k \leq 1)$  weights the importance of the propagated and initial estimated intra-relations. Note that the first term in each equation represents information exchange from neighboring entity, while the second term is the initial similarity. In this circular propagation, the intra-relations of  $N$  entities are produced simultaneously. Instead of estimating the similarity of entities with the same type separately, the circular propagation enforces additional constraint that two entities are similar if the entities of different type which are linked to them are connected strongly.

### 3.3 Convergence Property

Eq.(1) can be expressed in a matrix form as following,

$$\begin{cases} \mathbf{W}_1^{(t)} = \theta_1 \mathbf{\Pi}_{1 \sim N} \mathbf{W}_N^{(t-1)} \mathbf{\Pi}_{1 \sim N}^T + (1 - \theta_1) \mathbf{W}_1^{(0)} \\ \mathbf{W}_2^{(t)} = \theta_2 \mathbf{\Pi}_{2 \sim 1} \mathbf{W}_1^{(t)} \mathbf{\Pi}_{2 \sim 1}^T + (1 - \theta_2) \mathbf{W}_2^{(0)} \\ \dots \\ \mathbf{W}_k^{(t)} = \theta_k \mathbf{\Pi}_{k \sim k-1} \mathbf{W}_{k-1}^{(t)} \mathbf{\Pi}_{k \sim k-1}^T + (1 - \theta_k) \mathbf{W}_k^{(0)} \\ \dots \\ \mathbf{W}_{N-1}^{(t)} = \theta_{N-1} \mathbf{\Pi}_{N-1 \sim N-2} \mathbf{W}_{N-2}^{(t)} \mathbf{\Pi}_{N-1 \sim N-2}^T + (1 - \theta_{N-1}) \mathbf{W}_{N-1}^{(0)} \\ \mathbf{W}_N^{(t)} = \theta_N \mathbf{\Pi}_{N \sim N-1} \mathbf{W}_{N-1}^{(t)} \mathbf{\Pi}_{N \sim N-1}^T + (1 - \theta_N) \mathbf{W}_N^{(0)} \end{cases} \quad (2)$$

where  $\mathbf{W}_k = [w_k^{ab}]_{Z_k \times Z_k}$  and  $\mathbf{\Pi}_{k \sim j} = [\pi_{k \sim j}^{bc}]_{Z_k \times Z_j}$  is affinity and link matrix respectively, correspondingly composed

of propagated intra-relations and inter-link. Circular propagation is guaranteed to converge for having the following property:

$$\lim_{t \rightarrow \infty} (\mathbf{W}_k^{(t+1)} - \mathbf{W}_k^{(t)}) = 0 \quad (3)$$

**Proof:**

$$\begin{aligned} & \mathbf{W}_k^{(t+1)} - \mathbf{W}_k^{(t)} \\ &= (\theta_k \mathbf{\Pi}_{k \sim k-1} \mathbf{W}_{k-1}^{(t+1)} \mathbf{\Pi}_{k \sim k-1}^T + (1 - \theta_k) \mathbf{W}_k^{(0)}) \\ & \quad - (\theta_k \mathbf{\Pi}_{k \sim k-1} \mathbf{W}_{k-1}^{(t)} \mathbf{\Pi}_{k \sim k-1}^T + (1 - \theta_k) \mathbf{W}_k^{(0)}) \\ &= \theta_k \mathbf{\Pi}_{k \sim k-1} (\mathbf{W}_{k-1}^{(t+1)} - \mathbf{W}_{k-1}^{(t)}) \mathbf{\Pi}_{k \sim k-1}^T \\ &= \theta_k \mathbf{\Pi}_{k \sim k-1} (\theta_{k-1} \mathbf{\Pi}_{k-1 \sim k-2} \mathbf{W}_{k-2}^{(t+1)} \mathbf{\Pi}_{k-1 \sim k-2}^T) \mathbf{\Pi}_{k \sim k-1}^T \\ & \quad - \theta_k \mathbf{\Pi}_{k \sim k-1} (\theta_{k-1} \mathbf{\Pi}_{k-1 \sim k-2} \mathbf{W}_{k-2}^{(t)} \mathbf{\Pi}_{k-1 \sim k-2}^T) \mathbf{\Pi}_{k \sim k-1}^T \\ &= \theta_k \theta_{k-1} \mathbf{\Pi}_{k \sim k-1} \mathbf{\Pi}_{k-1 \sim k-2} (\mathbf{W}_{k-2}^{(t+1)} - \mathbf{W}_{k-2}^{(t)}) \\ & \quad \mathbf{\Pi}_{k-1 \sim k-2}^T \mathbf{\Pi}_{k \sim k-1}^T \\ &= \dots \\ &= \theta_k \dots \theta_1 \theta_N \dots \theta_{k+1} \mathbf{\Pi}_{k \sim k-1} \dots \mathbf{\Pi}_{1 \sim N} \mathbf{\Pi}_{N \sim N-1} \dots \mathbf{\Pi}_{k+1 \sim k} \\ & (\mathbf{W}_k^{(t)} - \mathbf{W}_k^{(t-1)}) \mathbf{\Pi}_{k+1 \sim k}^T \dots \mathbf{\Pi}_{N \sim N-1}^T \mathbf{\Pi}_{1 \sim N}^T \dots \mathbf{\Pi}_{k \sim k-1}^T \\ &= (\theta_k \dots \theta_1 \theta_N \dots \theta_{k+1} \mathbf{\Pi}_{k \sim k-1} \dots \mathbf{\Pi}_{1 \sim N} \mathbf{\Pi}_{N \sim N-1} \dots \mathbf{\Pi}_{k+1 \sim k})^t \\ & (\mathbf{W}_k^{(1)} - \mathbf{W}_k^{(0)}) (\mathbf{\Pi}_{k+1 \sim k}^T \dots \mathbf{\Pi}_{N \sim N-1}^T \mathbf{\Pi}_{1 \sim N}^T \dots \mathbf{\Pi}_{k \sim k-1}^T)^t \end{aligned} \quad (4)$$

It is easy to see Eq.(3) can be derived when each row of  $\mathbf{\Pi}_{k \sim k-1}$  is normalized to 1, and  $0 \leq \theta_N, \theta_{N-1}, \dots, \theta_1 \leq 1$ .  $\square$

## 4. UNIFIED ENTITY SEARCH

**DEFINITION 3:** The query of an entity search is represented by a vector  $\mathbf{q} = [v_1^0, v_2^0, \dots, v_N^0]^T$ , which corresponds to another vector  $\mathbf{u} = [\mu_1, \mu_2, \dots, \mu_N]^T$ , indicating the importance of different entity types in the query  $\mathbf{q}$ ,  $\sum_j \mu_j = 1$ .

**DEFINITION 4:** The objective of an entity search (e.g., searching entities of type  $k$ ) is to give a ranking score list  $\mathbf{r}_k = [r_k^1, r_k^2, \dots, r_k^{Z_k}]^T$ , which is a vector of the ranking scores, corresponding to the entity searching pool  $V_k$ , for instance,  $r_k^b$  is the ranking score of  $v_k^b$ .

**DEFINITION 5:** An entity search function is defined as follows,

$$\mathbf{r} = f(\mathbf{q}, V_k). \quad (5)$$

Permuting the entities according to this function is called *entity search*. In this paper, we formulate the entity search problem from the probabilistic perspective. Supposing the ranking score list is a random variable, entity search can be regarded as a process to derive the most probable score list given the certain entity query. From the probabilistic perspective, entity search is to derive the optimum  $\mathbf{r}^*$  with the maximum a posterior probability given the query  $\mathbf{q}$  and entity searching pool  $V_k$ ,

$$\mathbf{r}_k^* = \operatorname{argmax}_{\mathbf{r}_k} \Pr(\mathbf{r}_k | \mathbf{q}, V_k). \quad (6)$$

According to Bayesian formula and the independency assumption of the query  $\mathbf{q}$  and entity searching pool  $V_k$ , the posterior is proportional to the product of the *conditional prior probability* and the *likelihood*,

$$\Pr(\mathbf{r}_k | \mathbf{q}, V_k) \propto \Pr(\mathbf{r}_k | V_k) \times \Pr(\mathbf{q} | \mathbf{r}_k), \quad (7)$$

where  $\Pr(\mathbf{r}_k | V_k)$  is the *conditional prior* of the score list given the entity searching pool  $V_k$ .  $\Pr(\mathbf{q} | \mathbf{r}_k)$  is the *likelihood*, which expresses how probable the query  $\mathbf{q}$  is given the optimal ranking score list  $\mathbf{r}_k$ . Replacing the posterior in Eq.(6) with Eq.(7), we formulate the entity search as maximizing the product of a conditional prior and a likelihood,

$$\mathbf{r}_k^* = \operatorname{argmax}_{\mathbf{r}_k} \Pr(\mathbf{r}_k | V_k) \times \Pr(\mathbf{q} | \mathbf{r}_k). \quad (8)$$

For the *conditional prior*  $\Pr(\mathbf{r}_k|V_k)$ , we formulate it with a regularization term,

$$\Pr(\mathbf{r}_k|V_k) = \frac{1}{H_1} \exp(-\text{Reg}(\mathbf{r}_k, V_k)), \quad (9)$$

where  $H_1 = \sum_{\mathbf{r}} \exp(-\text{Reg}(\mathbf{r}_k, V_k))$  is a normalizing constant.  $\text{Reg}(\mathbf{r}_k, V_k)$  is defined as follows following Normalized Laplacian regularizer [25],

$$\text{Reg}(\mathbf{r}_k, V_k) = \sum_a \varphi_a(\mathbf{r}_k, V_k), \quad (10)$$

where  $\varphi_a(\mathbf{r}_k, V_k)$  defined over entity  $v_k^a$  for measuring the visual consistency on its neighbors  $v_k^b$ ,

$$\varphi_a(\mathbf{r}_k, V_k) = \frac{1}{2} \sum_b w_k(a, b) \left( \frac{r_k^a}{\sqrt{d_k^a}} - \frac{r_k^b}{\sqrt{d_k^b}} \right)^2, \quad (11)$$

and  $d_k^a = \sum_b w_k(a, b)$ . Then Eq.(10) can be expressed in a matrix form as following

$$\text{Reg}(\mathbf{r}_k, V_k) = \mathbf{r}_k^T \mathbf{S}_k \mathbf{r}_k, \quad (12)$$

where  $\mathbf{S}_k = \mathbf{D}_k^{-1/2} \mathbf{W}_k \mathbf{D}_k^{-1/2}$  in which  $\mathbf{D}_k$  is a diagonal matrix with its  $(a, a)$ -element ( $d_k^a$ ) equal to the sum of the  $a$ -th row of  $\mathbf{W}_k$ .

For the *likelihood*  $\Pr(\mathbf{q}|\mathbf{r}_k)$ , we formulate it as a *shortest-paths problem* on a weighted graph,

$$\Pr(\mathbf{q}|\mathbf{r}_k) = \frac{1}{H_2} \exp(-\rho' \times IR(\mathbf{q}, \mathbf{r}_k)), \quad (13)$$

where  $IR(\mathbf{q}, \mathbf{r}_k) = \sum_a (r_k^a - \delta(\mathbf{q}, v_k^a))^2$  measures the degree of the disagreement between the ranking scores and the shortest paths in graph.  $H_2$  is the normalizing constant and  $\rho' > 0$  is a scaling parameter.

**DEFINITION 6:** In a shortest-paths problem, for a weighted graph with entities of the same type  $j$ , the weight of path  $l = \langle v_j^m, v_j^{m+1}, \dots, v_j^{m+n} \rangle$  is the sum of the weights of its constituent edges:

$$\eta_j(l) = \sum_{i=1}^n w_j(m+i-1, m+i). \quad (14)$$

The shortest-part weight from  $v_j^a$  to  $v_j^b$  is defined by,

$$\delta'(v_j^a, v_j^b) = \begin{cases} \min\{\eta_j(l) : v_j^a \xrightarrow{l} v_j^b\} & \text{if } \exists v_j^c \rightarrow v_j^b \\ \infty & \text{otherwise} \end{cases}, \quad (15)$$

where  $v_j^a \rightarrow v_j^b$  denotes the path from  $v_j^a$  to  $v_j^b$ . Considering all path weights are nonnegative, we exploit Dijkstra's algorithm to solve the shortest-paths problem in Eq.(13) [17]. When the query type is different from the entities in searching pool, for example, for the path from  $v_j^d$  to  $v_k^b$ , the shortest-part weight is defined by,

$$\delta'(v_j^d, v_k^b) = \min\{\pi_{k \sim j}(b, c) \delta'(v_j^d, v_j^c) : v_j^c \in S_j(v_k^b)\}. \quad (16)$$

For a query  $\mathbf{q}$  with multiple entity types, the shortest part to an entity  $v_k^b$  is derived in linear fusion way,

$$\delta(\mathbf{q}, v_k^b) = \sum_j \mu_j \delta'(v_j^0, v_k^b), \quad (17)$$

then we have  $IR(\mathbf{q}, \mathbf{r}_k) = \sum_a (r_k^a - \delta(\mathbf{q}, v_k^a))^2$ .

Let us further introduce a vector  $\mathbf{\Delta}_k$  with its  $b$ -th element  $\delta(\mathbf{q}, v_k^b)$ . With Eq.(9) and Eq.(11), the entity search formulation in Eq.(8) is equivalent to minimizing the following

energy function,

$$E(\mathbf{r}_k) = \text{Reg}(\mathbf{r}_k, V_k) + \rho \times IR(\mathbf{q}, \mathbf{r}_k), \quad (18)$$

where the two terms on the right-hand side correspond to the conditional prior and the likelihood, respectively. The  $\rho = \frac{H_1}{H_2} \rho'$  is a trade-off parameter to the two terms. Given that the energy function in Eq.(18) is quadratic, we minimize  $E(\mathbf{r}_k)$  solving  $\nabla E(\mathbf{r}_k) = 0$ , and have the closed form,

$$\mathbf{r}_k^* = (\mathbf{I} - \rho \mathbf{S}_k)^{-1} \mathbf{\Delta}_k. \quad (19)$$

## 5. AN EXAMPLE: FLICKR

Without loss of generality, we take the most popular photo sharing site, Flickr, as the social media platform to illustrate our multi-level graph modeling, which mainly has four types of entities, i.e., user ( $V_u$ ), photo ( $V_p$ ), tags ( $V_t$ ) and geo-locations ( $V_g$ ). It is worth noting that although we take Flickr as an example, the proposed multi-level graph model can be applied to any kind of community, such as Facebook and LinkedIn. Due to the missing and noises of geo-locations for most of photos, we perform the circular propagation by using the entity sets  $V_u$ ,  $V_p$  and  $V_t$  as follows, similar to Eq.(2),

$$\begin{cases} \mathbf{W}_t^{(t)} = \theta_t \mathbf{\Pi}_{t \sim u} \mathbf{W}_u^{(t-1)} \mathbf{\Pi}_{t \sim u}^T + (1 - \theta_t) \mathbf{W}_t^{(0)} \\ \mathbf{W}_p^{(t)} = \theta_p \mathbf{\Pi}_{p \sim t} \mathbf{W}_t^{(t)} \mathbf{\Pi}_{p \sim t}^T + (1 - \theta_p) \mathbf{W}_p^{(0)} \\ \mathbf{W}_u^{(t)} = \theta_u \mathbf{\Pi}_{u \sim p} \mathbf{W}_p^{(t)} \mathbf{\Pi}_{u \sim p}^T + (1 - \theta_u) \mathbf{W}_u^{(0)} \end{cases}, \quad (20)$$

where  $\mathbf{W}_t^{(t)}$ ,  $\mathbf{W}_p^{(t)}$  and  $\mathbf{W}_u^{(t)}$  is the affinity matrix of tags, photos and users at the  $t$ -th iteration, respectively. When  $t = 0$ , the affinity matrixes are initial ones before propagation.  $\mathbf{\Pi}_{t \sim u}$ ,  $\mathbf{\Pi}_{p \sim t}$  and  $\mathbf{\Pi}_{u \sim p}$  denote the three link matrixes measuring the link strength of the two levels of graphs, i.e., from tag graph to user graph, from photo graph to tag graph, and from user graph to photo graph.  $\theta_t$ ,  $\theta_p$  and  $\theta_u$  are tradeoff parameters.

### 5.1 Initial Intra-relations

- Initial user intra-relation  $\mathbf{W}_u^{(0)}$

In a social media community, the initial user similarity is mostly determined by people interaction [18], including explicit interaction (one gives comments to the other's photos or one's photos get comments from others) and implicit interaction (two users simultaneously give comments on the same photos). In addition, photos taken in near locations are often shared by the people with similar interests [3], the intra-relation  $w_u^{ab(0)}$  between user  $v_u^a$  and user  $v_u^b$  is defined as:

$$w_u^{ab(0)} = (f_b(a) + f_a(b) + g(a, b)) \times w_g^{ab}, \quad (21)$$

where  $f_a(b)$  denotes the number of the comments given by the user  $v_u^b$  on the photos shared by user  $v_u^a$ , vice versa.  $g(a, b)$  denotes the number of photos on which user  $v_u^a$  and  $v_u^b$  simultaneously comment.  $w_g^{ab}$  is the relation between user  $v_u^a$  and  $v_u^b$  caused by geo-locations, and it is given by

$$w_g^{ab} = \begin{cases} \pi & \text{if } \min_{gd}(v_u^a, v_u^b) < Th \\ 1 & \text{else} \end{cases}, \quad (22)$$

where  $\min_{gd}(v_u^a, v_u^b)$  is the minimal geo-location distance between photos of user  $v_u^a$  and  $v_u^b$ ,  $Th$  is a fixed threshold and  $\pi$  ( $\pi > 1$ ) is a positive number. In this paper,  $Th$  is set to  $20km$  and  $\pi$  is set to 2.

- Initial photo intra-relation  $\mathbf{W}_p^{(0)}$

We follow the approach in [11] and adopt scale-invariant feature transform (SIFT) descriptor with a Difference of Gaussian (DoG) interest point detector for extracting the images’ visual patterns. The interest point is referred to as local salient patch, each associated with a 128-dimensional feature vector. We further use K-means to cluster the similar patches into “visual words,” and use Bag-of-Word (BoW) to represent each image as it has proven to be effective for object and scene retrieval [6][16][21][23]. Let  $\mathbf{f}_a$  and  $\mathbf{f}_b$  denote the BoW feature vector of photo  $v_p^a$  and  $v_p^b$ , respectively, then the initial similarity  $w_p^{ab(0)}$  between the two photos is calculated by

$$w_p^{ab(0)} = \exp \left\{ -\frac{Ed(\mathbf{f}_a, \mathbf{f}_b)}{2\sigma_1^2} \right\}, \quad (23)$$

where  $Ed(\mathbf{f}_a, \mathbf{f}_b)$  denotes the Euclidean distance between  $\mathbf{f}_a$  and  $\mathbf{f}_b$ ,  $\sigma_1$  is the positive radius parameter, and it is estimated by the median value of all the Euclidean distances.

- Initial tag intra-relation  $\mathbf{W}_t^{(0)}$

Tags supplied by users describe the content of photos while providing additional contextual modalities about the photos. As a preprocessing step, the standard stemming and stop word removal are firstly applied. The  $M$  tags which have highest frequency are selected for computing their similarity.

Analogous to the principle of Google similarity distance [2], we define the tag distance  $w_t^{ab(0)}$  between tag  $v_t^a$  and  $v_t^b$ , as follows:

$$Td(v_t^a, v_t^b) = \frac{\max(\log f(v_t^a), \log f(v_t^b)) - \log f(v_t^a, v_t^b)}{\log G - \min(\log f(v_t^a), \log f(v_t^b))}, \quad (24)$$

where  $f(v_t^a)$  and  $f(v_t^b)$  are the numbers of photos containing tag  $v_t^a$  and  $v_t^b$ , respectively.  $f(v_t^a, v_t^b)$  is the number of photos containing both tag  $v_t^a$  and  $v_t^b$ . These numbers can be obtained by performing search on Flickr website using the tags as keywords. The initial tag similarity between tag  $v_t^a$  and  $v_t^b$  is then given by

$$w_t^{ab(0)} = \exp \left\{ -\frac{Td(v_t^a, v_t^b)}{2\sigma_2^2} \right\}. \quad (25)$$

Similar to  $\sigma_1$ ,  $\sigma_2$  is also estimated by the median value of all the tag distances.

## 5.2 Inter-links

As shown in Eq.(20), there are three types of inter-links, i.e., from tag graph to user graph ( $\mathbf{\Pi}_{t \sim u}$ ), from photo graph to tag graph ( $\mathbf{\Pi}_{p \sim t}$ ), and from user graph to photo graph ( $\mathbf{\Pi}_{u \sim p}$ ). Accordingly, three link matrices are built as follows.

- $\mathbf{\Pi}_{t \sim u} \equiv [\pi_{t \sim u}^{ab}]_{(Z_t \times Z_u)}$ : measuring the link strength of entities in tag and user graph. It is defined as:

$$\pi_{t \sim u}^{ab} = \begin{cases} 1/n_{v_t^a} & v_u^b \in v_t^a \\ 0 & else \end{cases}, \quad (26)$$

where  $n_{v_t^a}$  denotes the number of users who provide tag  $v_t^a$  to a photo,  $v_u^b \in v_t^a$  means user  $v_u^b$  annotates a photo with tag  $v_t^a$ .

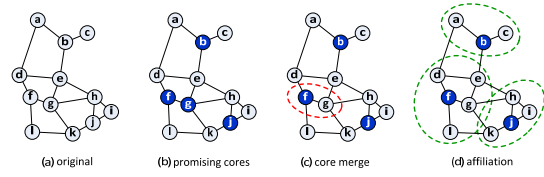


Figure 2: The illustration for building sub-nets.

- $\mathbf{\Pi}_{p \sim t} \equiv [\pi_{p \sim t}^{ab}]_{(Z_p \times Z_t)}$ : measuring the link strength of entities in photo and tag graph. It is given by:

$$\pi_{p \sim t}^{ab} = \begin{cases} 1/n_{v_p^a} & v_t^b \in v_p^a \\ 0 & else \end{cases}, \quad (27)$$

where  $n_{v_p^a}$  denotes the tag number of visual photo  $v_p^a$ , and  $v_t^b \in v_p^a$  means tag  $v_t^b$  belongs to photo  $v_p^a$ .

- $\mathbf{\Pi}_{u \sim p} \equiv [\pi_{u \sim p}^{ab}]_{(Z_u \times Z_p)}$ : measuring the link strength of entities in user and photo graph. Similarly, we have:

$$\pi_{u \sim p}^{ab} = \begin{cases} 1/n_{v_u^a} & v_p^b \in v_u^a \\ 0 & else \end{cases}, \quad (28)$$

where  $n_{v_u^a}$  denotes the photo number shared by user  $v_u^a$ , and  $v_p^b \in v_u^a$  means photo  $v_p^b$  is shared by user  $v_u^a$ .

## 5.3 Implementations

In practice, it is not necessary to run the multi-level graph propagation as a whole in the social network. Instead, the network itself is usually formed by disconnected sub-nets, and the propagation can be conducted separately on each sub-net. In our implementation, two users are connected if one of the following conditions is met: (1) the maximal initial tag similarity is above an average value; (2) their photos are taken in nearby locations; (3) they comment each other’s photos or comment on a photo contributed by a third party. This could result in a number of sub-nets. To make the finding of sub-nets computationally efficient, we first detect the “promising cores” as seeds and then grow the sub-nets from the seeds, by connecting other users fulfilling the aforementioned three conditions. Fig.2 illustrates the procedure of forming social sub-nets.

The “promising core” of a sub-net, intuitively, is the user who has a large number of photos, attracts significant photo views and comments from others, while not connecting to the cores of other sub-nets. Referred to as Fig. 2, we select cores as the users with a minimum of 500 photos, accumulate at least 2,000 comments, and share no more than 20% of similar tags with others. Then the rest of users are selected as the promising cores(Fig. 2(b)). Nearby cores as in Fig. 2(c) may be merged if they have high initial user similarity. Finally, users are connected to the cores based on their initial user similarities. Based on this implementation, taking our experiment as an example, we detect 101 sub-nets from 40,170 Flickr users. Typically, a sub-net contains about 400 users. It is also worth noting that only tags with high frequency will be used in the propagation since tags provided by Flickr users are often imprecise, even containing form and spelling form [10]. We empirically select top 1,000 tags.

## 6. EXPERIMENTS

In this section, we systematically evaluate the effectiveness of the proposed unified formulation and solution to en-

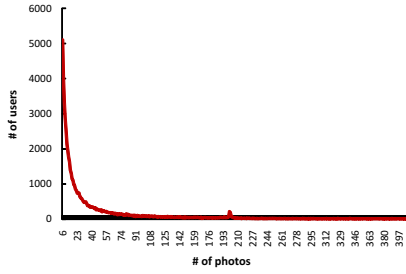


Figure 3: The statistics of photo number owned by users (users owned less than five photos are removed).

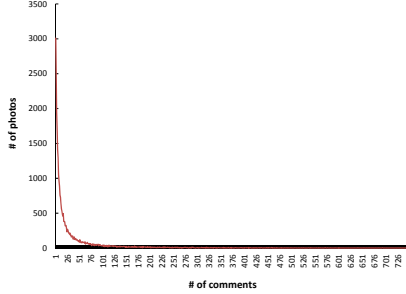


Figure 4: The statistics of comment number given by photos.

tity search in social media community, and report comparative results on three popular entity search tasks, i.e., friend suggestion, image tagging and personalized image search.

## 6.1 Dataset

We empirically verify the merit of entity search on a dataset crawled from Flickr [4]. The data from Flickr consists of five main elements: users, uploaded photos, the tags associated with each photo, geo-location of each photo, and users who comment on the photo (called “cmt users” for short). We randomly collect 165,558 users and the corresponding photos, tags and comments by using Flickr API. For our experiment, we remove the users who have less than 5 photos, and then the dataset includes 40,170 users (not includes the cmt users), owning about 9 million photos.

Fig.3 shows the number of photos owned by users. The long tail indicates most of users share similar number of photos, and very few users own the extremely large number of photos. In average, a user owns around 200 photos in our experiment. Fig.4 shows the number of comments on photos. The long tail phenomenon also exists, which indicates although the social interaction is complex, the comments on each photo are not too many (13 comments per photo in average). Fig.5 illustrates the tag frequency, where the noise (misspelling words, combination of words and affix variation) and some special words form the long tail.

## 6.2 Exp-I: Friend Suggestion

Friend suggestion aims to search a list of users, which have similar interests to the given user. Obviously, the entity type of both query and searching pool is “user.” Therefore, in this task, the user similarity is mainly explored and a ranking list of users can be obtained according to Eq.(19). We compare

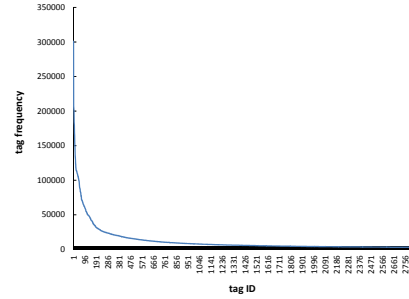


Figure 5: The tag frequency distribution.

our unified framework based friend suggestion (UFS) with the following five methods in both subjective and objective experiments:

- Initial user similarity (IUS). The friends are ranked based on the initial user similarities, which are directly determined by user interaction in photo sharing activities as mentioned in Section 5.1.
- Tag correlations (Tag). A tag correlation measured by the number of common tags on the users’ photos. The higher the number of common tags, the more likely to be friends of the two users.
- Visual relationships (Visual). A visual relationship estimated according to the visual similarity of representative photos of users.
- Fusion of user, tag and visual similarity (Fusion). A linear fusion of IUS, Tag and Visual.
- Context-based friend suggestion (Context) [22]. It aims to leverage multiple contexts, including user-contributed photos, their associated tags and geo-locations, as well as user behaviors like viewing and commenting.

### 6.2.1 Subjective User Study

To evaluate the proposed friend suggestion strategy, we randomly selected 1,000 users as the query users. The similarity of each user pair was manually labeled by three subjects on a scale of 1-3: (1) dissimilar (totally different interest), (2) similar (somewhat relevant interest), and (3) strongly similar (almost same interest). The ground truth similar degree of each pair is the median scale of three evaluations. Considering the pairs with high similarity have especially important for most of applications, we adopted the Normalized Discounted Cumulative Gain ( $NDCG$ ) as the performance metric. The  $NDCG$  score at the depth  $d$  in the ranked lists is defined by:

$$NDCG@d = Z_d \sum_{j=1}^d \frac{2^{r^j} - 1}{\log(1 + j)} \quad (29)$$

where  $r^j$  is the rating of the  $j$ -th pair,  $Z_d$  is a normalization constant and is chosen so that a perfect ranking’s  $NDCG@d$  value is 1. After computing the  $NDCG$  measure of each query user, we average them to obtain an overall performance evaluation on friend suggestion.

Fig.6 shows the experimental results. Overall, the results across different depth of  $NDCG$  consistently indicate that UFS achieves a performance boost compared to IUS, Tag

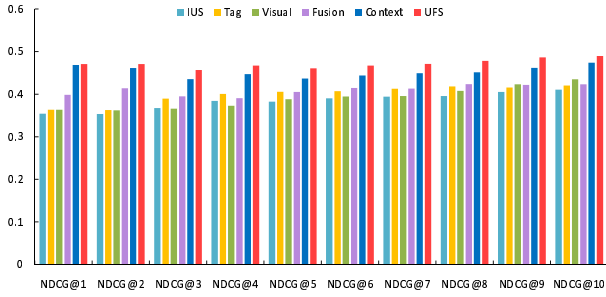


Figure 6: Subjective evaluation on friend suggestion.

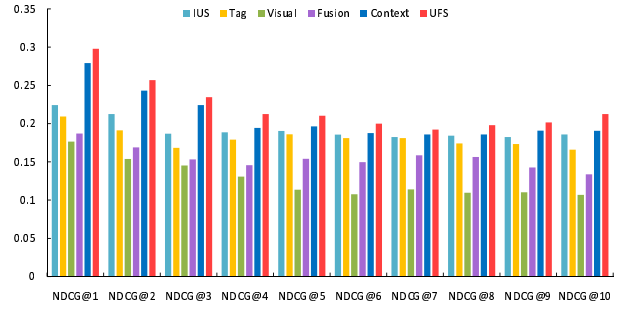


Figure 8: Objective evaluation on friend suggestion.

	User Id	Representative photos
Query users	8519617@N03	
Suggested friend list	8789171@N07	
	11771426@N03	
	14307969@N03	
	21362899@N08	
	21512060@N06	

Figure 7: The representative photos of suggested friends for the user Id “8519617@N03.”

and Visual which using single entity. Furthermore, in contrast to the linear fusion and context-based method, the superiority of UFS demonstrates that allowing entity interaction can lead to better performance gain. As indicated by our results, the friend suggestion based on our unified entity search framework can better express the user relationship by reinforcing the mutual exchange of information across multiple entities over an integrated graph model.

Fig.7 shows the top five users’ representative photos to a query user according to the final user intra-relations. From the representative photos of each user, we can see that these photos are similar with each other. These users are likely to have similar interest and thus are potential candidates for friend suggestion.

### 6.2.2 Objective Evaluation of Friend Suggestion

The objective evaluation was conducted according to the *contacts* function in Flickr site, which is organized by users themselves. If a user exists in the other user’s contact list, we view the two users as friends. We first filtered out the users who have no contacts in Flickr site or in our user set. After filtering, we obtained 2,975 users and took them as the query users. Similar to the subjective user study, the objective comparison as shown in Fig.8 also indicates the strong performance of our proposed UFS in comparison to others.

## 6.3 Exp-II: Image Tagging

Image tagging is to assign a set of tags to a given image to describe the image content. In our unified entity search framework, the query type is “image” and the type of entities to be searched is “tag.” By exploring the affinity matrix  $\mathbf{W}_t$ , and link matrix  $\Pi_{p \sim t}$ , the tag ranking list can be obtained by Eq.(19).

We compare our unified framework based image tagging (UIT) with the following two baseline methods.

- Initial image tags in Flickr site (IIT): The tags originally provided by the image owners.
- Tagging ranking (TR). A Random walk is performed over a tag similarity graph to estimate the relevance score of tags for a given image [10].

To demonstrate the effectiveness of the image tagging methods, we randomly select 200 images from our collected Flickr image dataset to perform image tagging. We also invite people to label the tags on a 1-3 scales: (1) “irrelevant,” (2) “relevant,” and (3) “excellent,” and adopt *NDCG* as the performance metric. After computing the *NDCG* score of each image’s tag list, we average them to obtain an overall performance evaluation of the image tagging method. Fig.9 shows the experimental results. We can see our proposed unified solution achieves much better performance compared to the other baseline algorithms at different depths of *NDCG*. The improvements demonstrate that UIT can facilitate the relationship and interaction of different entities through mutual reinforcement and all the relationships can be refined.

Fig.10 further illustrates some exemplary images with the top five tags generated by three image tagging methods. We can clearly see that our recommended tags can better characterize the image content.

## 6.4 Exp-III: Personalized Image Search

Given the large and growing importance of search engines, personalized search has the potential to significantly improve searching experience [15]. Compared with the common image search, in personalized image search, the user-specific information is considered to distinguish the exact intentions of textual queries and rank the images. Formally in our framework, the query types are “user” and “textual tags,” while the entities in searching pool are “images.” To obtain the image ranking list, we estimate the affinity matrix based on the photo affinity matrix  $\mathbf{W}_p$ , and link matrixes  $\Pi_{u \sim p}$  and  $\Pi_{p \sim t}$ . According to Eq.(19), we can obtain the image ranking list.



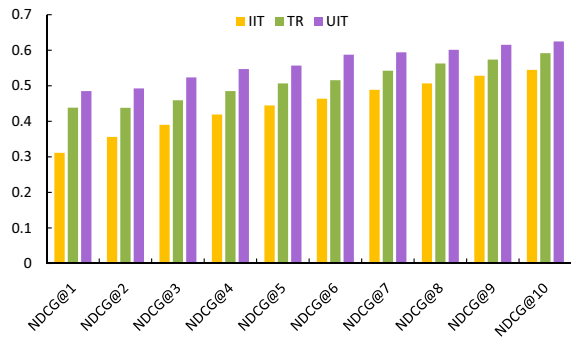


Figure 9: Performance comparison of image tagging methods.

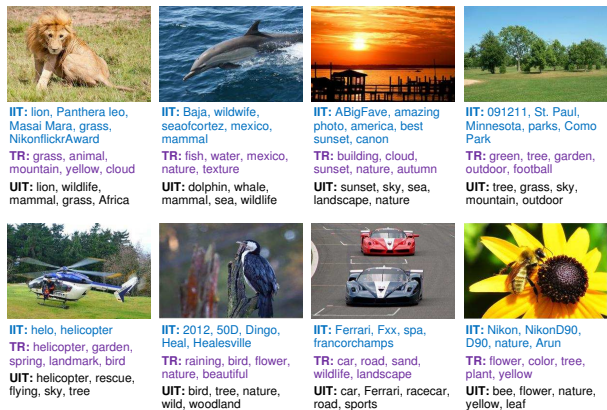


Figure 10: Example results of image tagging.

We select the following three state-of-the-art models as the baselines and compare them with our unified framework based method (UPIS).

- Non-personalized (Non). Use the common text-based image search mainly based on *tf-idf* functions [16].
- Topic-based (Topic). It aims to explore folksonomy for personalized search [19].
- Preference-based (Pref). Perform personalized image search by predicting user interests-based preference [12].

Similar to [12], we use *comment*-based evaluation approach, where the images attract comments from user  $u$  are treated as relevant when  $u$  issues queries. We build two test scenarios for the comment-based evaluation: 1) 50 randomly selected users who gave comments to 10-30 photos, denoted as Dataset\_10-30, and 2) 20 users who gave comments to more than 100 photos, denoted as Dataset\_100. 15 tags frequently appearing in the tags of commented images are selected as the test queries. The metric of *mMAP* [15] is utilized to evaluate the performance and the results are demonstrated in Fig.12. From the results, we have the following observations. (1) Our proposed unified solution achieves better performance compared to the other three baseline algorithms. This clearly demonstrates the effectiveness of our proposed unified solution in the personalized image search task. (2) All *mMAP* values are not very high. This phenomenon reflects the problem of comment-based evaluation

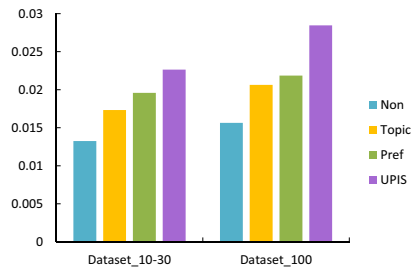


Figure 12: Performance comparison of personalized image search for different methods.

scheme: the commented images are considered relevant for all the test queries. As no query information is involved, for those queries non-relevant with the topic of the commented images, the performance tends to be low. (3) Comparing between the two test scenarios, the average performance of Data\_100 also improves over Data\_10-30. One possible reason for the improvement is that those users having more comments are active users who are likely to also attend more interest groups and tag more images.

Fig.11 displays exemplary search results for the query “sunrise.” The top eight non-personalized results and the personalized results of three users (User A, User B and User C) are shown. We can see that by simultaneously considering query relevance and user information, the unified framework based personalized image search could capture the user’s preference under certain topics. From the results, we can see that the top search results for user A mainly focus on sunrise at the seaside, while for user B and user C, the top search results are often co-exist with cloud and plants. For the baseline method (non-personalized) which separate query relevance and user preference, most of the search results are relevant to query, yet may hard to interpret a clear search intent.

## 6.5 Complexity Analysis

The complexity of our approach includes offline and on-line two parts. The offline multi-level graph construction on 40,170 users with their photos, tags and comments can be finished in ten hours on five servers, each having 4GB memories and running four threads. Our online entity search algorithm is extremely efficient. The complexity of the algorithm is  $O(|V| \lg |V| + |E|)$ , where  $|V|$  and  $|E|$  represent the number of vertices and edges, respectively. Take 200 images to perform image tagging for example, our algorithm takes less than 60.2 seconds on a regular PC (Intel dual-core 3.33GHz CPU and 4 GB RAM) to complete the whole process. In other words, tagging one image only takes 0.3 seconds. It can be even faster if executed on a more powerful machine with parallel computing capability. Clearly, the speed is efficient and provides almost instant response.

## 7. CONCLUSIONS

In this paper, we have shown how to integrate various entities in a multi-level graph model, which includes the content from images, tags and geo-locations, as well as the rich context derived from user behaviors. Specifically, the relations between users, images and tags are fully exploited by reinforcing the mutual exchange and propagation of information relevancy across different entities. Upon the model, entity

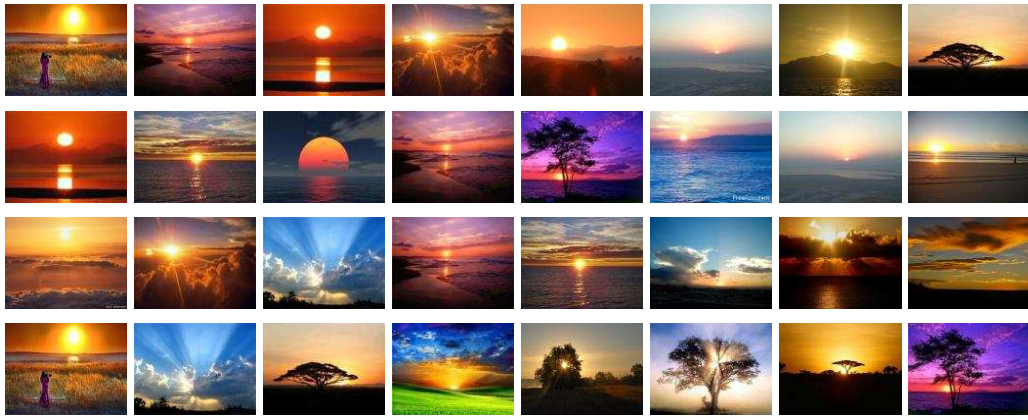


Figure 11: Example of non-personalized (top) and personalized (second, third and fourth for User A, User B and User C respectively) search results for query “sunrise.”

search is explicitly formulated within a unified framework, in which friend suggestion, image tagging and personalized image search can be implemented in one single way. Future work includes the extension to multiple media sharing sites, such as Facebook and LinkedIn. It would be a promising topic to investigate whether these entities are consistent across different social media sharing sites and whether specific factors may impact the way users create, share, use and annotate content.

## 8. ACKNOWLEDGMENTS

This work was supported in part by a grant from the RGC of the Hong Kong SAR, China (CityU 119610), a grant from the National Natural Science Foundation of China (#61272290), and a grant from Microsoft Research Asia Windows Phone Academic Program (FY12-RES-OPP-107).

## 9. REFERENCES

- [1] H. M. Chen, M. H. Chang, P. C. Chang, M. C. Tien, W. H. Hsu, and J. L. Wu. Sheepdog-group and tag recommendation for flickr photos by automatic search-based learning. In *Proceedings of ACM Multimedia*, 2008.
- [2] R. L. Cilibrasi and P. M. Vitanyi. The google similarity distance. *IEEE Trans. on Knowledge and Data Engineering*, 19(3):370–383, 2007.
- [3] D. Easley and J. Kleinberg. *Networks and Crowds and Markets: Reasoning About a Highly Connected World*. Cambridge University Press, 2010.
- [4] Flickr. <http://www.flickr.com/>.
- [5] ImageClef. <http://imageclef.org/>.
- [6] Y. Jing and S. Baluja. Pagerank for product image search. In *Proceedings of International WWW Conference*, 2008.
- [7] J. Li and J. Z. Wang. Real-time computerized annotation of pictures. *IEEE Trans. on Pattern Analysis and Machine Intelligence*, 2008.
- [8] Q. Li, Y. Zheng, X. Xie, Y. Chen, W. Liu, and W.-Y. Ma. Mining user similarity based on location history. In *Proceedings of ACM GIS*, 2008.
- [9] Y.-R. Lin, J. Sun, P. Castro, R. Konuru, H. Sundaram, and A. Kelliher. Metafac: Community discovery via relational hypergraph factorization. In *Proceedings of ACM KDD*, 2009.
- [10] D. Liu, X.-S. Hua, L. Yang, M. Wang, and H.-J. Zhang. Tag ranking. In *Proceedings of International WWW Conference*, 2009.
- [11] D. Lowe. Object recognition with informative features and linear classification. In *Proceedings of ICCV*, 2003.
- [12] D. Lu and Q. Li. Personalized search on flickr based on searcher’s preference prediction. In *Proceedings of International WWW Conference*, 2011.
- [13] R.-A. Negoescu, B. Adams, D. Phung, S. Venkatesh, and D. Gatica-Perez. Flickr hypergroups. In *Proceedings of ACM Multimedia*, 2009.
- [14] M. Roth, A. David, D. Deutscher, G. Flysher, I. Horn, A. Leichtberg, N. Leiser, Y. Matias, and R. Merom. Suggesting friends using the implicit social graph. In *Proceedings of ACM KDD*, 2010.
- [15] J. Sang, C. Xu, and D. Lu. Learn to personalized image search from the photo sharing websites. *IEEE Trans. on Multimedia*, 14(4), 2012.
- [16] J. Sivic and A. Zisserman. Video google: A text retrieval approach to object matching in videos. In *Proceedings of IEEE International Conference on Computer Vision*, 2003.
- [17] M. Sniedovich. Dijkstra’s algorithm revisited: the dynamic programming connexion. *Control and Cybernetics*, 35(3), 2006.
- [18] C. Wilson, A. Sala, K. P. N. Puttaswamy, and B. Y. Zhao. Beyond social graphs: User interactions in online social networks and their implications. *ACM Trans. on the Web*, 2012.
- [19] S. Xu, S. Bao, B. Fei, Z. Su, and Y. Yu. Exploring folksonomy for personalized search. In *Proceedings of ACM SIGIR*, 2008.
- [20] Yahoo. <http://people.yahoo.com/>.
- [21] T. Yao, T. Mei, and C.-W. Ngo. Co-reranking by mutual reinforcement for image search. In *ACM CIVR*, 2010.
- [22] T. Yao, C.-W. Ngo, and T. Mei. Context-based friend suggestion in online photo-sharing community. In *Proceedings of ACM Multimedia*, 2011.
- [23] T. Yao, C.-W. Ngo, and T. Mei. Circular reranking for visual search. *IEEE Trans. on Image Processing*, 22(4), April 2013.
- [24] N. Zheng, Q. Li, S. Liao, and L. Zhang. Flickr group recommendation based on tensor decomposition. In *Proceedings of ACM SIGIR*, 2010.
- [25] D. Zhou, O. Bousquet, T. Lal, J. Weston, and B. Schölkopf. Learning with local and global consistency. In *Proceedings of NIPS*, 2003.
- [26] J. Zhuang, T. Mei, S. Hoi, X.-S. Hua, and S. Li. Modeling social strength in social media community via kernel-based learning. In *Proceedings of ACM Multimedia*, 2011.