Hybrid Social Media Network

Dong Liu
Dept. of Electrical Engineering
Columbia University
New York, NY 10027, USA
dongliu@ee.columbia.edu

Guangnan Ye
Dept. of Electrical Engineering
Columbia University
New York, NY 10027, USA
yegn@ee.columbia.edu

Ching-Ting Chen
Dept. of Electrical Engineering
Columbia University
New York, NY 10027, USA
cc3016@columbia.edu

Shuicheng Yan
Dept. of ECE
National Univ. of Singapore
Singapore, 117576, Singapore
eleyans@nus.edu.sg

Shih-Fu Chang
Dept. of Electrical Engineering
Columbia University
New York, NY 10027, USA
sfchang@ee.columbia.edu

ABSTRACT

Analysis and recommendation of multimedia information can be greatly improved if we know the interactions between the content, user, and concept, which can be easily observed from the social media networks. However, there are many heterogeneous entities and relations in such networks, making it difficult to fully represent and exploit the diverse array of information. In this paper, we develop a *hybrid social media network*, through which the heterogeneous entities and relations are seamlessly integrated and a joint inference procedure across the heterogeneous entities and relations can be developed. The network can be used to generate personalized information recommendation in response to specific targets of interests, e.g., personalized multimedia albums, target advertisement and friend/topic recommendation. In the proposed network, each node denotes an entity and the multiple edges between nodes characterize the diverse relations between the entities (e.g., friends, similar contents, related concepts, favorites, tags, etc). Given a query from a user indicating his/her information needs, a propagation over the hybrid social media network is employed to infer the utility scores of all the entities in the network while learning the edge selection function to activate only a sparse subset of relevant edges, such that the query information can be best propagated along the activated paths. Driven by the intuition that much redundancy exists among the diverse relations, we have developed a robust optimization framework based on several sparsity principles. We show significant performance gains of the proposed method over the state of the art in multimedia retrieval and recommendation using data crawled from social media sites. To the best of our knowledge, this is the first model supporting not only aggregation but also judicious selection of heterogeneous relations in the social media networks.

Categories and Subject Descriptors

H.3.1 [**Information Storage and Retrieval**]: Content Analysis and Indexing

General Terms

Algorithms, Experimentation, Performance

Keywords

Hybrid Social Media Network, Edge Selective Information Propagation, Multimedia Retrieval and Recommendation.

1. INTRODUCTION

The modern information era is marked by the explosively growing networks of people, information and the rich interactions between them as demonstrated in the prevalent use of many well known social media networks, such as Flickr, Youtube and Facebook. Rather than simply searching for and passively consuming media content, users in such networks actively create and exchange rich multimedia content (including videos, images, music, blogs and so on) for social interactions [2], which brings greatly expanded experiences for end users.

Notably, the social media networks are characterized by the heterogeneous entities and relations. Take Facebook as an example, where the user, concept and multimedia object coexist in the network with various types of relations. Users may share multiple relations such as friendship, group membership, message communication and so on. Multimedia objects may also have multiple relations, each of which reflects the similarity in certain features or attributes (e.g., color, shape, attributes, or metadata). In addition, each user can post comments or rating for a multimedia object, bookmark it as favorite, or even assign notes to certain regions in the object, which forms multiple relations between the user and the multimedia object. The richness of entities and relations provides a comprehensive description of the contextual information, facilitating the design of novel multimedia applications with new user experiences.

However, the richness and diversity of the entities and relations mentioned above bring new challenges in leveraging such information for multimedia analysis. First, there are heterogeneous entities and relations in the social media network, in which some relations are multi-faceted (e.g., relations between two images could be measured based on
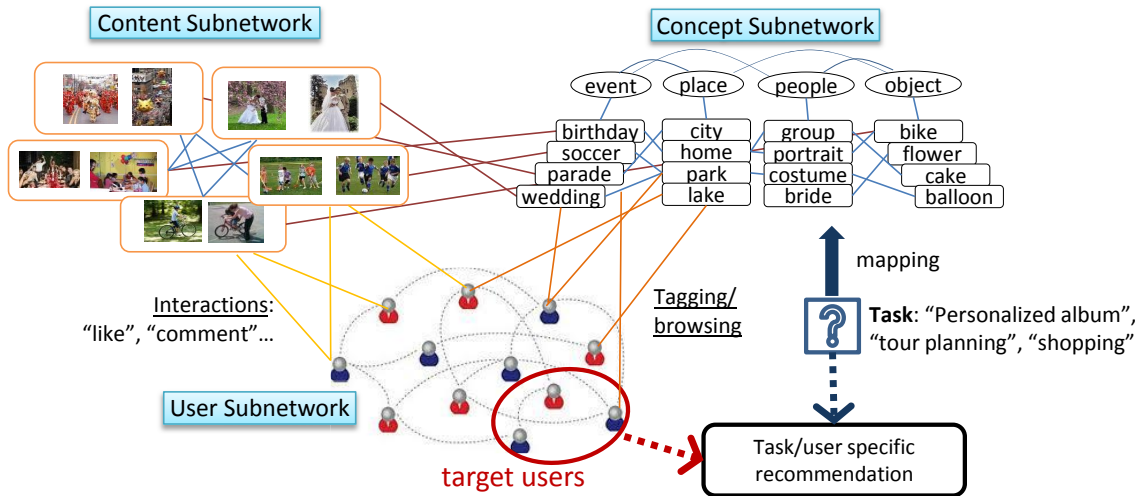
Figure 1: Different from the conventional network models that treat nodes in a homogeneous way, the proposed hybrid social media network preserves the heterogeneous types of entities and relations from subnetworks of user, content, and concept. Given a query from a user or task, the inference process dynamically selects the most informative links and relation types based on a sparsity constrained optimization model, and propagates the information along the activated paths only. The results can be used to recommend information in a personalized way for applications like personalized album generation, friend/topic recommendation, and target ads.

similarities of different raw features as well as mid-level attributes) while some relations are beyond pairwise (e.g. a user annotates a photo with a tag, which forms a three-way relation among the three entities), making it difficult to develop a unified model to capture such heterogeneous data types and relations. Second, to exploit the rich entities and relations, straightforward application of traditional methods (like information propagation [23] and spectral clustering [13, 6]) designed for homogeneous networks is no longer adequate. We need to investigate novel solutions that can not only fuse complex relations but also prudently select most useful information to improve the overall performance in information propagation.

In this work, we propose a novel *hybrid social media network* model, through which entities of heterogeneous types and their relations can be seamlessly integrated and new inference techniques can be developed. The hybrid social media network is illustrated in Figure 1. Specifically, the hybrid network spans several homogeneous subnetworks (made of content, people and concept) and the rich multi-faceted and possibly high-order connections both within and between subnetworks. It differs from the conventional networks by incorporating the much richer types of entities and relations. For example, users may form relations based on friendship, author-subscriber, or group membership as discussed earlier. In the concept subnetwork, there may be rich statistical and ontological relationships. Between the subnetworks, users may view, like, dislike, or recommend content, while individual photos, videos, albums, or channels may be assigned various concepts (about scenes, objects, events, or aesthetics). In addition, users can assign concepts to the specific content, forming high-order relations among the three subnetworks. The proposed representation over the hybrid networks provides great flexibility for encoding various types of complex relations at different levels.

The proposed hybrid network can be used as an effective system for information propagation and recommenda-

tion. The specific targets of interest (about user, concept, or content) can be first formulated as a query vector with the vector elements corresponding to the target nodes set to 1 and the rest 0. For example, a task of "recommend media content about architecture and fashion of New York for user A" can be formulated as a query vector specifying the corresponding entities as the targets. The goal is to propagate the information contained in the query vector to other nodes in the network, through the large diverse pool of relation links. After the propagation process is complete, each node in the network will be associated with a predicted utility score, which can then be used to rank and recommend the most relevant information in response to the initial targets.

We investigate how to determine the best mechanism for propagating information over the network and propose a solution based on edge-selective optimizations, which activates only a sparse subset of relations for information propagation. Specifically, we define the propagation process as a Markov random walk in which the transition probability between each node pair can be calculated by the weighted combination of only the selected relations along the edges. The consistency between the predicted utility scores and the initial scores of the query provides a natural supervision step. To achieve effective edge selection, we use the $\ell_{2,1}$-norm induced group sparsity to determine a subset of relation types while using ℓ_1-norm induced sparsity to reduce the total number of edges selected and eliminate the potential redundancy that may exist among the diverse relations. We formulate the approach as a sparsity regularized objective and derive an optimization process that can simultaneously estimate the node utility scores and the edge selective indicators.

The use of the sparsity principles can be justified in several ways. First, given the large array of relations captured in the network, it is natural to expect information redundancy. For example, the similarities between images may be redundantly manifested in several features like edges, local features, and textures. Likewise, contact relations between

users may strongly bias the existence of favorite relations or the use of similar tags in content annotation. Second, judicious selections of a small subset of relations may greatly reduce the computational complexity of the information propagation process. We will confirm the validity of the above rationales using empirical experiment results to be presented in Section 6.3.

We will demonstrate experimentally the proposed hybrid network can achieve significant performance gains when evaluated over various use scenarios such as personalized content recommendation, topic/friend discovery, and target ad, etc. We will also show clear evidences confirming the capabilities of the model for dynamically selecting the most useful relations in answering each query.

2. RELATED WORK

Some recent work [7, 21] focused on aggregating the rich variety of information in the social network, through which the performance of information retrieval can be improved. However, these works typically aggregate information using a simple weighted average and then directly utilize the aggregation results for final analysis. Such an averaging operation blindly converts the hybrid relations into a homogeneous type and cannot adapt the edge combination weights for each query dynamically. In contrast, we emphasize the core challenge addressed in our work is how to harness the heterogeneous nodes and relations in solving the associated optimization problem. Our main contribution is the explicit modeling and optimal selection of diverse edges in the hybrid network based on a sound sparsity principle. The edge selection technique is key, which simultaneously exploits a large pool of heterogeneous edges (instead of blindly aggregating all the edges as in the prior work) and explicitly determines a subset of edges for each query dynamically (the selected edges will be assigned with different weights reflecting the importance of each edge for each specific query). Solving the dynamic edge selection problem is crucial in successful modeling of hybrid networks, but has not been addressed in the literature so far.

Graph-based ranking has shown effective performance in propagating information in sparsely connected networks. One representative work along this direction is the manifold ranking method proposed by Zhou *et al.* [23], which ranks the samples based on the intrinsic manifold structure among the samples. Given a query vector that indicates the relevant nodes in the graph (treated as initial positive labels), the objective is to estimate a function that predicts the relevance score of each node in the rest of the graph. The optimization process finds the optimal solution that fits the initial query and is smooth over the graph. He *et al.* [9] employed such ranking method for content-based image retrieval and obtained good performance. The other popular method is the graph based random walk, where the stationary probability of the random walk process is used as the ranking scores of the nodes. Such ranking method has been widely used in video search reranking [11], social media tag ranking [15], etc. However, all these methods are limited to homogeneous networks with nodes of the same type, which are inadequate for modeling the heterogeneous entities and relations.

Some recent work also used graphs to model the multiple types of entities and relations. The recent works in [1, 24] proposed to build multiple graphs, each of which models a distinct feature type, but the approach was to aggregate the diverse types of relations into a single composite type, which is fundamentally different from our proposed hybrid network model using an explicit joint learning procedure over heterogeneous entities and edges. Notably, the recently proposed hypergraph [3] is also a generalization of the traditional graph, where a hyperedge can connect any number of nodes and thus models the high-order relations among the entities. In [4], the authors employed the hypergraph to combine the social media information and music content, and achieved good performance on mucic recommendation. The work in [8] employed the hypergraph to model the multi-typed objects including documents, tags and users, based on which a personalized tag recommendation task is cast into a graph based ranking problem. Despite the capability in modeling heterogeneous entities and relations, such graph-based models aim at describing the high-order relationship among a collection of entities, but cannot model the multi-faceted yet diverse relations among entities. Specifically, none of the prior work addressed the issue of "selecting" the subset of most informative relations. Instead, all of them went on to develop "fusion" or "aggregation" of a potentially over-populated pool of relations – likely suffering from the problem of information overload.

The heterogeneous information network [10, 12] in the data mining community also attempts to model the relations between heterogeneous entities. However, the modeling is performed by connecting any two entities with a single kind of relation, which actually falls in the single edge graph and thus differs from our proposal here. In particular, the authors in [16] introduced a homogeneous multi-edge graph model to describe multiple parallel relations between the multi-label images, where an edge connects two similar local regions in different images. However, it can only handle the homogeneous entities (local regions) and relations (pairwise regional similarities), and hence cannot be applied into the heterogeneous scenarios.

3. USE SCENARIOS OF THE HYBRID SOCIAL MEDIA NETWORK

Our hybrid network can be used as an effective system for information propagation and recommendation. We present a few sample use cases below to exemplify the recommendation process. Additional cases can also be easily constructed.

Personalized recommendation of contents, friends, and topics: Given a specific user and his/her interest (e.g., "NYC" + "The statue of liberty") as the target query (i.e., setting the nodes corresponding to the users and the concepts to 1 in the initial query vector), the hybrid network is used to propagate the information throughout the network and predict the utility scores for all other nodes. After the propagation, the highest ranked images can be used to generate a personalized multimedia album relevant to the specific interest. The nodes corresponding to users receiving the highest utility scores can be used as the candidates for recommending friends who may share similar interests. Additionally, other concepts (like "art" and "jewelry") may be discovered as recommended topics of interest to the user.

Target media ads: We can also set a certain set of media objects or topical concepts as the initial query and propagate such information through the network. The final predicted utility scores associated with the user nodes in the network can be used to identify the best group of users

Table 1: Statistics of the entities and relations.

Entities	Notation	Avg. #
Tags	$\mathcal{C}$	100
Users	$\mathcal{U}$	778
Photos	$\mathcal{O}$	7,046

Relation Types	Order	Avg. #
user-image favorite	second	467
user-user favorite	second	1,322
user-user contact	second	2,146
user-image comment	second	4,203
concept-concept co-occurrence	second	5,231
user-image ownership	second	7,046
image-image SIFT similarity	second	21,138
image-image mid-level similarity	second	21,138
user-concept ownership	second	23,115
concept-image tagging	second	23,115
user-image-tag tagging	third	23,026

that may be interested in receiving such media objects or topics. Media objects and their associated products (e.g., music, art, or food) can be pushed to such identified user groups. Note the explicit use of heterogeneous relations (e.g., image content similarity, ontological concept relations) provides unique benefits that will allow recommendation results to go beyond what has been done in the collaborative filtering community which often focuses on user preference rating data only.

4. CONSTRUCTION OF HYBRID SOCIAL MEDIA NETWORK

In this section, we will construct a sample hybrid social media network by crawling the real-world social images from Flickr. We first introduce the details on the crawling process and then explain how to extract the observable relations to construct the hybrid social media network.

4.1 Social Media Dataset

To construct a sample hybrid social media network for modeling the heterogeneous entities and relations, we collected social media data from the popular social media website Flickr. Specifically, we crawled 10 photo groups including "Beijing", "New York", "Taipei", "IKEA", "Universal Studio", "Fashion", "Korean Food", "Sneaker", "Painted", and "Wildlife", each of which is a self-organized photo sharing community with common interests. We chose these groups to cover a diverse set of user groups, ranging from cities, tourist sites, products, and life styles. We downloaded all the photos, users and tags in each group. To ensure that each tag corresponds to a meaningful semantic concept, a filtering process was performed to remove the meaningless tags. We ranked the tags based on their occurring frequencies, and selected the top 100 tags as the concept set of each group (excluding the stop words and other meaningless words)[1]. The final set of users, photos, and concepts are used as the entity nodes of the constructed network. For the heterogeneous relations between the entities, we crawled the observable relations between the different entities in Flickr and obtained 8 types of relations, which include *user-user contact*

relation, *user-user favorite relation*, *user-image ownership relation*, *user-image comment relation*, *user-image favorite relation*, *user-concept ownership relation*, *concept-image tagging relation* and *user-image-tag tagging relation*. Besides these relations from Flickr, we also constructed the *concept-concept co-occurrence relation* and two kinds of *image-image similarity relations* to exemplify the relations between the corresponding entities, resulting in 11 types of relations in total. Table 1 shows the statistics of the entities and relations averaged across the 10 groups.

4.2 Hybrid Network Construction

For each photo group, a hybrid social media network can be built. Denote by $\mathcal{O} = \{o_1, \ldots, o_{|\mathcal{O}|}\}$, $\mathcal{C} = \{c_1, \ldots, c_{|\mathcal{C}|}\}$ and $\mathcal{U} = \{u_1, \ldots, u_{|\mathcal{U}|}\}$ the entities of multimedia objects, concepts and users in a group respectively, where $|\cdot|$ denotes the cardinality of a set. We treat each entity as one node and establish the relational edges (both pair-wise and high-order) and their associated weights as follows:

User-user contact: For users u_i and u_j, we establish an edge if one user is listed in the contact list of the other, and the weight w_{ij} is set as 1.

User-user favorite: An edge is built to connect u_i and u_j if one user favorites the images owned by the other user. The edge weight is estimated as $w_{ij} = t_{ij}/\max(t_{ij})$, where t_{ij} is total number of favorite images between the two users, and $\max(t_{ij})$ denotes the maximum number of image favorites between any two users.

User-image ownership/favorite/comment: For the first two relations, if a user owns or favorites one image, there will be an edge with weight set as 1. For the comment relation, an edge is built to connect a user and an image, and the weight is determined as the number of comments the user has provided for the image divided by the maximum comment number between any pair of user and image.

User-concept ownership: If concept c_j has appeared in the photos of u_i, an edge will be built between u_i and c_j and the weight is set as 1.

Concept-image tagging: We establish an edge if one image is annotated with one tag and the corresponding edge weight is set as 1.

User-image-concept tagging: If a user has tagged one image with one concept, there will be a hyperedge connecting these three entities and the weight is 1.

Concept-concept co-occurrence: Given tag c_i and c_j, we estimate their statistical relation as $w_{ij} = f(c_i, c_j)/|\mathcal{O}|$, where $f(c_i, c_j)$ denotes the number of images in the group that are simultaneously tagged by c_i and c_j while $|\mathcal{O}|$ denotes the total number of images in the network. Currently, we use co-occurrence relations among concepts to facilitate the hybrid network study. Our framework allows incorporation of other relations like lexical and ontological ones.

Image-image similarity: We use two kinds of features to estimate the image similarity, resulting in two types of relations. One feature is the SIFT Bag-of-Words (BoW) feature [14], and for each image we extract 128-dimensional SIFT feature from the keypoints detected by two kinds of detectors: Difference of Gaussian and Hessian Affine [18]. Then, k-means method is applied to group the SIFT features into 5,000 clusters. Finally, each image is represented by a 5,000-dimensional BoW feature. The second feature is the mid-level Classeme feature [22], which is a 2,659-dimensional feature vector corresponding to the classifier

---

[1]Though advanced techniques can be used to extract higher-level semantic concepts from tags, in this paper it is not our focus and the terms "tag" and "concept" are used equally.

output scores for $2,659$ semantic concepts. In this work, we choose the nearest neighbor graph to construct the image-image similarity subnetwork. For each image o_i, we find its K nearest neighbors and establish an edge between o_i and its neighbors. The similarity between images is defined as

$$w_{ij} = \begin{cases} \exp(-\frac{d(o_i,o_j)}{\sigma}), & \text{if } j \in \mathcal{N}_K(i) \text{ or } i \in \mathcal{N}_K(j), \\ 0, & \text{otherwise,} \end{cases} \quad (1)$$

where $\mathcal{N}_K(j)$ denotes the index set of the K nearest neighbors of sample o_j (we set $K = 3$), $d(o_i, o_j)$ denotes the distance between two samples (We use χ^2 distance for SIFT BoW and use Euclidean distance for Classeme), σ is the radius parameter of the Gaussian function, which is set as the mean value of all pairwise distances between the samples.

5. INFORMATION PROPAGATION OVER HYBRID SOCIAL MEDIA NETWORK

In this section, we present the proposed information propagation method over the hybrid network. We first present the formulation based on the multi-level sparsity and then introduce an efficient procedure for the optimization.

5.1 Problem Formulation

Denote by r the total number of relation types amongst the multimedia objects, concepts and users. Let $G = (V, E)$ denote a hybrid social media network with the set of nodes V and the set of edges E. The node set is a finite set $V = \{v_1, \ldots, v_n\} = \{o_1, \ldots, o_{|\mathcal{O}|}, c_1, \ldots, c_{|\mathcal{C}|}, u_1, \ldots, u_{|\mathcal{U}|}\}$ with $n = |\mathcal{O}| + |\mathcal{C}| + |\mathcal{U}|$. The edge set is a family of multi-edge set $E = \{\mathcal{E}_{ij} \bigcup \mathcal{H}_{ij}\}$ amongst V where $\mathcal{E}_{ij} = \{e_{ij}^1, \ldots, e_{ij}^{n_{ij}}\}$ denotes the multiple edges between node v_i and v_j with the number of edges $n_{ij} = |\mathcal{E}_{ij}|$ and each e_{ij}^k characterizes one type of relations between the two nodes. Furthermore, $\mathcal{H}_{ij} = \{h_{ij}^1, \ldots, h_{ij}^{m_{ij}}\}$ is the set of hyperedges that are associated with v_i and v_j, where each h_{ij}^k denotes one type of high-order relation composed of v_i, v_j and other nodes in V, m_{ij} is the total number of hyperedges that contain v_i and v_j. To ease the representation, we harmonize the edges and hyperedges associated with v_i and v_j as $\mathcal{S}_{ij} = \{s_{ij}^1, \ldots, s_{ij}^{n_{ij}}, s_{ij}^{n_{ij}+1}, \ldots, s_{ij}^{n_{ij}+m_{ij}}\}$, where each s_{ij}^k denotes an edge ($k = 1, \ldots, n_{ij}$) or a hyperedge ($k = n_{ij}+1, \ldots, n_{ij}+m_{ij}$). Furthermore, we use w_{ij}^k to denote the edge affinity weight associated with edge s_{ij}^k.

Given the initial query nodes, the task is to propagate the initial information contained in the query through the network to predict the utility scores for each node in the entire network. Let $\mathbf{f} = [f_1, \ldots, f_n]^\top$ denote the predicted utility score vector for all the nodes. We define a query vector $\mathbf{y} = [y_1, \ldots, y_n]^\top$ in which $y_i = 1$ if $v_i \in V$ is in the query and $y_i = 0$ otherwise. For example, for the aforementioned query "recommend content about architecture and fashion of New York to user A", y_i for the nodes "architecture", "fashion", "New York", and "user A" will be set to 1.

The information propagation over the hybrid social media network is to infer the utility score of each node based on the hybrid network as well as the label information conveyed by the query. However, we are facing challenges particularly in view of the large number of relation links of heterogeneous types. It is intuitive to assume that when propagating information for each query, many of the relations are redundant and only a small subset of relation types should be used. Therefore, the edges should be selected so that only the most appropriate ones are retained as the suitable path for information propagation. To realize the edge selective information propagation, we resort to sparsity induced regularizer to couple the edge selection and information propagation together. Given an edge/hyperedge s_{ij}^k and its associated affinity weight w_{ij}^k, we establish an edge strength function $\phi(w_{ij}^k) = \alpha_{ij}^k w_{ij}^k$ to indicate the importance of the edge in the propagation, where $\alpha_{ij}^k \geq 0$ is a selective parameter controlling whether the edge is selected. An edge is strong if its affinity value is high and it is selected. The selective parameter plays a key role in our framework – only edges assigned non-zero values are selected for information propagation. Note w_{ij}'s capture the raw relations computed offline from the data, while α_{ij}'s are learned for each query dynamically online. It is important to note that we only associate a single selective parameter for a multi-way hyperedge, which means that all the node pairs involved in a hyperedge will share the same edge selective parameter. In this way, the high-order hyper-relation can be well preserved in our model.

To handle each type of relations differently, we also group the selective parameters for edges belonging to the same type (e.g., all edges corresponding to color similarities, user-friendship, user-content favorites, etc) together and denote each type as $\boldsymbol{\alpha}_g$, with g indicating the type. Denote by $\boldsymbol{\alpha}$ the vector comprising of all selective parameters. Then we formulate the edge selective optimization problem as minimizing the following objective function:

$$\Omega(\mathbf{f}, \{\phi(w_{ij}^k)\}) + \frac{\lambda}{2}\|\mathbf{f} - \mathbf{y}\|_2^2 + \gamma(\sum_{g=1}^{r} \|\boldsymbol{\alpha}_g\|_2 + \|\boldsymbol{\alpha}\|_1), \quad (2)$$

where $\lambda, \gamma > 0$ are two trade-off parameters among the three competing terms. The first term Ω is a regularization term responsible for the information propagation. The second term is the label fitness constraint enforcing that a good propagation should not deviate too much from the initial label assignment of the query nodes. The third term is sparse penalty on the edges that explicitly achieves the sparse edge subset selection at both the group and individual edges. The ℓ_{21}-norm in the last term ensures the sparse selection principle is applied to the group level, namely, only a small number of relation types will be selected and the ℓ_1-norm is used to ensure only a small set of individual edges are selected. This is quite intuitive: for instance, for friend recommendation, it is likely that the image similarity edges will not contribute much and the optimization process may result in the complete deselection of all relation edges of this type. In the case when all types of relations are useful, selection of edges at the individual levels is still useful for minimizing the redundancy. Note that the relation selection process is optimized for each query dynamically instead of using a fixed set of selection rules for different queries.

In this work, we establish the connection between the edge strength function $\phi(w_{ij}^k)$ and the node utility score f_i within the term Ω based on a Markov random walk process. Denote by Q_{ij} the transition probability from node v_i to node v_j, which can be calculated by,

$$Q_{ij} = \begin{cases} \frac{\exp(\sum_k \alpha_{ij}^k w_{ij}^k)}{\sum_j \exp(\sum_k \alpha_{ij}^k w_{ij}^k)}, & \text{if } \mathcal{S}_{ij} \in E, \\ 0, & \text{otherwise.} \end{cases} \quad (3)$$

According to Markov random walk process, the stationary distribution vector $\mathbf{f}$ is the solution of the following eigenvector equation:

$$\mathbf{f} = \mathbf{Q}^\top \mathbf{f}. \tag{4}$$

Note that the above equation unifies the utility scores of both labeled and unlabeled nodes in the network, which achieves the information propagation along the network structure. Recall that an edge is strong if its affinity value is high and it is selected after the propagation process. A strong edge can help propagate scores from one node to its neighboring nodes through the random walk framework in Eq. (4). This is enforced by minimizing $\|\mathbf{f} - \mathbf{Q}^\top \mathbf{f}\|_2^2$ – two nodes connected by selected strong edges will have similar scores after the propagation. In this way, Eq. (4) establishes the connection between node utility scores and the edge selective parameters via the transition matrix $\mathbf{Q}$, which estimates the node scores based on the edge strengths and thus offers the capability for selecting the most appropriate edges during propagation. Note other functional forms, such as graph Laplacian, may be considered as alternatives to the random walk formulation above. Here, we adopt this approach due to its amenity to the optimization process to be presented below. Finally, the objective function can be written as:

$$
\begin{aligned}
F(\mathbf{f}, \boldsymbol{\alpha}) &= \frac{1}{2}\|\mathbf{f} - \mathbf{Q}^\top \mathbf{f}\|_2^2 \\
&+ \frac{\lambda}{2}\|\mathbf{f} - \mathbf{y}\|_2^2 + \gamma\left(\sum_{g=1}^r \|\boldsymbol{\alpha}_g\|_2 + \|\boldsymbol{\alpha}\|_1\right). \quad (5)
\end{aligned}
$$

By minimizing the objective function, the node utility scores and the edge selective parameters can be derived as:

$$
\begin{aligned}
\min_{\mathbf{f}, \boldsymbol{\alpha}} \quad & F(\mathbf{f}, \boldsymbol{\alpha}), \\
\text{s.t.} \quad & \mathbf{f} \geq 0, \ \boldsymbol{\alpha} \geq 0. \tag{6}
\end{aligned}
$$

The above objective function is non-convex, and thus can only achieve a local optimum. In the next subsection, we will use a gradient based optimizer to solve it.

5.2 Optimization Procedure

In Eq. (5), the gradient of $\mathbf{f}$ can be directly calculated. However, the gradient of $\boldsymbol{\alpha}$ cannot be calculated due to the non-smoothness of the ℓ_{21}-norm and ℓ_1-norm regularizer. In this subsection, we will show that, using the dual norm, these sparsity terms can be approached by a smoothing approximation such that the gradient can be calculated. Then we will employ the gradient based method for optimization.

5.2.1 Smoothing Approximation

Since the dual norm of ℓ_2-norm is still ℓ_2-norm, we can write $\sum_{g=1}^r \|\boldsymbol{\alpha}_g\|_2$ in Eq. (5) as:

$$\sum_{g=1}^r \|\boldsymbol{\alpha}_g\|_2 = \sum_{g=1}^r \max_{\|\boldsymbol{\beta}_g\|_2 \leq 1} \langle \boldsymbol{\beta}_g, \boldsymbol{\alpha}_g \rangle = \max_{\boldsymbol{\beta} \in \mathcal{B}} \boldsymbol{\beta}^\top C \boldsymbol{\alpha}, \tag{7}$$

where $\boldsymbol{\beta}_g$ is a vector of auxiliary variables associated with $\boldsymbol{\alpha}_g$, and $\langle \cdot, \cdot \rangle$ denotes the inner product operator. Let $\boldsymbol{\beta} = [\boldsymbol{\beta}_1^\top, \ldots, \boldsymbol{\beta}_r^\top]^\top$ and its domain is $\mathcal{B} = \{\boldsymbol{\beta} | \|\boldsymbol{\beta}_g\|_2 \leq 1, g = 1, \ldots, r\}$. Denote by n_g the number of edges in the gth relation $\mathcal{G}_g$, and let $J = \sum_{g=1}^r n_g$ the total number of edges in the hybrid social media graph. Then $C \in \mathbb{R}^{\sum_{g=1}^r n_g \times J}$ is a matrix whose rows are indexed by all pairs of $(i, g) \in$

$\{(i, g) | i \in \mathcal{G}_g, i \in \{1, \ldots, J\}\}$, and the columns are indexed by $j \in \{1, \ldots, J\}$. Each element of C is defined as:

$$C_{(i,g),j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{otherwise.} \end{cases} \tag{8}$$

Based on Nesterov's smoothing approximation method [20], we construct a smooth approximate of $\sum_{g=1}^r \|\boldsymbol{\alpha}_g\|_2$ as:

$$h_\mu(\boldsymbol{\alpha}) = \max_{\boldsymbol{\beta} \in \mathcal{B}} \langle \boldsymbol{\beta}, C\boldsymbol{\alpha} \rangle - \frac{\mu}{2}\|\boldsymbol{\beta}\|_2^2, \tag{9}$$

where μ is a positive smoothness parameter to control the accuracy of the approximate. For a fixed $\boldsymbol{\alpha}$, the optimal solution of Eq. (9) can be defined as:

$$\boldsymbol{\beta}_g(\boldsymbol{\alpha}_g) = S_2\left(\frac{\boldsymbol{\alpha}_g}{\mu}\right), \tag{10}$$

where S_2 is the projection operator which projects any vector $\mathbf{x}$ to the ℓ_2-ball:

$$S_2(\mathbf{x}) = \begin{cases} \frac{\mathbf{x}}{\|\mathbf{x}\|_2}, & \|\mathbf{x}\|_2 > 1, \\ \mathbf{x}, & \|\mathbf{x}\|_2 \leq 1. \end{cases} \tag{11}$$

On the other hand, the dual of ℓ_1-norm is ℓ_∞-norm, the $\|\boldsymbol{\alpha}\|_1$ can be approximated by the following smooth function:

$$l_\mu(\boldsymbol{\alpha}) = \max_{\|\mathbf{u}\|_\infty \leq 1} \langle \mathbf{u}, \boldsymbol{\alpha} \rangle - \frac{\mu}{2}\|\mathbf{u}\|_2^2, \tag{12}$$

where the optimal auxiliary variable $\mathbf{u}(\boldsymbol{\alpha})$ can be defined as:

$$\mathbf{u}(\boldsymbol{\alpha}) = S_\infty\left(\frac{\boldsymbol{\alpha}}{\mu}\right), \tag{13}$$

where S_∞ is the projection operator which projects a value to the ℓ_∞-ball:

$$S_\infty(x) = \begin{cases} x, & -1 \leq x \leq 1, \\ 1, & x > 1, \\ -1, & x < -1. \end{cases} \tag{14}$$

Based on the above two approximations, we arrive at the following smoothed objective function:

$$F_\mu(\mathbf{f}, \boldsymbol{\alpha}) = \frac{1}{2}\|\mathbf{f} - \mathbf{Q}^\top \mathbf{f}\|_2^2 + \frac{\lambda}{2}\|\mathbf{f} - \mathbf{y}\|_2^2 + \gamma \Pi_\mu(\boldsymbol{\alpha}), \tag{15}$$

where $\Pi_\mu(\boldsymbol{\alpha})$ is given by

$$\Pi_\mu(\boldsymbol{\alpha}) = h_\mu(\boldsymbol{\alpha}) + l_\mu(\boldsymbol{\alpha}). \tag{16}$$

Apparently, the original sparsity term in Eq. (5) can be written as $\Pi_0(\boldsymbol{\alpha})$ with $\mu = 0$. We have the following theorem on the accuracy of the approximation:

THEOREM 1. *The maximum gap between $\Pi_0(\boldsymbol{\alpha})$ and $\Pi_\mu(\boldsymbol{\alpha})$ is $\mu(r + J)/2$, i.e.,*

$$\Pi_\mu(\boldsymbol{\alpha}) \leq \Pi_0(\boldsymbol{\alpha}) \leq \Pi_\mu(\boldsymbol{\alpha}) + \mu(r + J)/2. \tag{17}$$

It shows that a smaller μ will result in a more precise approximation, which makes the solution of Eq. (15) sufficiently close to the optimal solution of Eq. (5). In our implementation, we set μ to be 10^{-4}. The following theorem shows that $\Pi_\mu(\boldsymbol{\alpha})$ is smooth w.r.t $\boldsymbol{\alpha}$ with a simple form of gradient.

THEOREM 2. *For any $\mu > 0$, the approximation function $\Pi_\mu(\boldsymbol{\alpha})$ is convex and continuously differentiable function of $\boldsymbol{\alpha}$ and its gradient is*

$$\nabla \Pi_\mu(\boldsymbol{\alpha}) = C^\top \boldsymbol{\beta}(\boldsymbol{\alpha}) + \mathbf{u}(\boldsymbol{\alpha}). \tag{18}$$

The proofs of the above two theorems are straightforward by following the proofs in [5].

Algorithm 1 Gradient descent optimization for information propagation over hybrid social media network

---

1: **Input:** $\{w_{ij}^k\}$, $\mathbf{y} \in \{0,1\}^n$, $\mathbf{f}^0 \in \mathbb{R}^n$, $\boldsymbol{\alpha} \in \mathbf{R}^J$, λ and γ.
2: **Initialization:** Set $t = 0$, initialize $\mathbf{f}^t = \mathbf{1}$ and $\boldsymbol{\alpha}^t = \mathbf{1}$.
3: **repeat**
4:     Fix $\boldsymbol{\alpha}^t$, and then employ PR conjugate gradient algorithm to estimate $\mathbf{f}^{t+1}$ based on $\partial_{\mathbf{f}} F_\mu(\mathbf{f}^t, \boldsymbol{\alpha}^t)$.
5:     Force the negative entries in $\mathbf{f}^{t+1}$ to 0.
6:     Fix $\mathbf{f}^{t+1}$, and employ PR conjugate gradient algorithm to estimate $\boldsymbol{\alpha}^{t+1}$ based on $\partial_{\boldsymbol{\alpha}} F_\mu(\mathbf{f}^{t+1}, \boldsymbol{\alpha}^t)$.
7:     Force the negative entries in $\boldsymbol{\alpha}^{t+1}$ to 0.
8:     $t = t + 1$.
9: **until** convergence.
10: **Output:** The optimized $\mathbf{f}^*$ and $\boldsymbol{\alpha}^*$.

5.2.2 Smoothing Gradient Descent

It is apparent that the smooth objective function $F_\mu(\mathbf{f}, \boldsymbol{\alpha})$ is differentiable w.r.t $\{f_i\}$ and $\{\alpha_{ij}^q\}$ as follows:

$$
\begin{aligned}
\partial_{f_i} F_\mu &= \left(f_i - \sum_k f_k Q_{ki}\right) + \sum_{j \in \mathcal{N}_i}\left(f_j - \sum_k f_k Q_{kj}\right)(-Q_{ij}) \\
&+ \lambda(f_i - y_i), \qquad\qquad\qquad (19) \\
\partial_{\alpha_{ij}^q} F_\mu &= \left(f_i - \sum_k f_k Q_{ki}\right)\left(-f_j \frac{\partial Q_{ji}}{\partial \alpha_{ij}^q}\right) \\
&+ \left(f_j - \sum_k f_k Q_{kj}\right)\left(-f_i \frac{\partial Q_{ij}}{\partial \alpha_{ij}^q}\right) + \gamma \nabla \Pi_\mu(\alpha_{ij}^q), \,\,(20)
\end{aligned}
$$

where $\mathcal{N}_i$ in Eq. (19) denotes the indices of the nodes in the hybrid network connecting to node v_i. The derivation of $\partial Q_{ji}/\partial \alpha_{ij}^q$ and $\partial Q_{ij}/\partial \alpha_{ij}^q$ in Eq. (20) can be calculated by,

$$
\begin{aligned}
\frac{\partial Q_{ji}}{\partial \alpha_{ij}^q} &= \frac{w_{ij}^q \exp(\sum_k \alpha_{ij}^k w_{ij}^k)}{\sum_p \exp(\sum_k \alpha_{jp}^k w_{jp}^k)} - \frac{w_{ij}^q\big(\exp(\sum_k \alpha_{ij}^k w_{ij}^k)\big)^2}{\big(\sum_p \exp(\sum_k \alpha_{jp}^k w_{jp}^k)\big)^2}, \\
\frac{\partial Q_{ij}}{\partial \alpha_{ij}^q} &= \frac{w_{ij}^q \exp(\sum_k \alpha_{ij}^k w_{ij}^k)}{\sum_p \exp(\sum_k \alpha_{ip}^k w_{ip}^k)} - \frac{w_{ij}^q\big(\exp(\sum_k \alpha_{ij}^k w_{ij}^k)\big)^2}{\big(\sum_p \exp(\sum_k \alpha_{ip}^k w_{ip}^k)\big)^2}.
\end{aligned}
$$

Based on the gradients, we can iteratively update $\mathbf{f}$ and $\boldsymbol{\alpha}$ by gradient descent method. Our implementation uses the Polack-Ribière (PR) conjugate gradient method [19], which is efficient for solving large-scale problems due to its simplicity and low storage. Given a function and its partial derivatives, the solver iteratively improves the estimates of the variables, converging to a local optimum. After obtaining the converged solution for $\mathbf{f}$ or $\boldsymbol{\alpha}$ in each iteration, we further enforce the negative entities in these vectors to be 0 such that the nonnegative constraints can be maintained. The entire optimization procedure is shown in Algorithm 1.

5.2.3 Algorithmic Analysis

Note that each gradient descent iteration in Algorithm 1 will monotonically decrease the objective function, and hence the algorithm will approach a local optimum. Figure 2 shows the convergence process of the iterative optimization which is captured in our later experiment. As seen, the objective function converges after 6 iterations, and thus the convergence speed is fast.

The time complexity of Algorithm 1 is $\mathcal{O}\big(l(n+J)\big)$, where l, n and J are respectively iteration number, node number and edge number. It is linear in the number of entities and relations. In the image ranking experiment of a
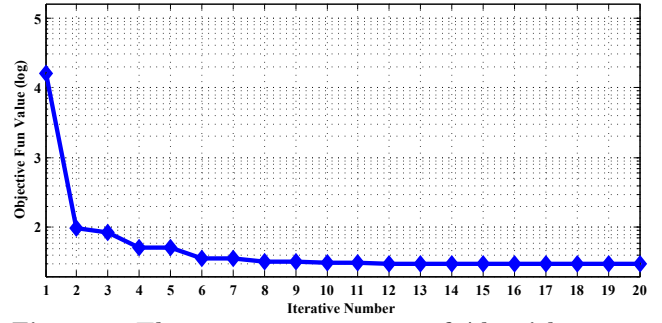

Figure 2: The convergence curve of Algorithm 1 on the IKEA group when performing image ranking for a Task-II query (see Section 6).

Task-II query on IKEA group (see Section 6) implemented on the MATLAB platform on an Intel XeonX5660 workstation with 3.2 GHz CPU and 18 GB memory, Algorithm 1 can be finished within 1.2 minutes. Of course, since the objective function is not convex, some care must be taken to avoid the local minima. However, we empirically find that the convergence behavior is pretty stable (see Figure 2) and is not sensitive to the initialization. To verify this, we test different initialization options including all-one vector and random vector for initializing $\mathbf{f}$ and $\boldsymbol{\alpha}$ when performing Task-II experiment. We found small differences in the accuracy, but the all-one initialization converges about 3 times faster. This provides some empirical justifications that using all-one vector as the initialization is a good choice for the optimization, through other initial values can also be used if prior knowledge about the score distribution is available.

6. EXPERIMENTS

In this section, we evaluate the effectiveness of the proposed information propagation algorithm over the hybrid network. We first present the experimental setup and then introduce the performance comparison and analysis.

6.1 Experiment Setup

As discussed in Section 4, we construct a hybrid network for each photo-sharing social group. The following three tasks will be used for performance evaluation. (1) **Task-I**: Given a user in the network as query, rank all entities in the network and then show the top ranked concepts/images/users as the customized recommendation. (2) **Task-II**: Given a user plus few concepts as query, rank all entities in the network and then use the top ranked concepts/images/users for recommendation. (3) **Task-III**: Given an image plus its associated concepts, rank the users to find the potential users who may have interests in receiving such media objects and topics.

For each of the above tasks, we compare the performances of the following four methods. (1) Edge Averaging based Information Propagation (EAIP). In this method, we average the multiple edge weights between each node pair to obtain a fused weight. This is actually the most common method for early fusion of heterogeneous graphs. Some works like [7, 21] extended the fusion to simple weighted averaging. However, it's unclear how optimal fusion weights can be determined and thus its impact on performance is uncertain. After the edge fusion, we employ the graph ranking method in [23] to obtain the ranking scores. The method in [23] also includes optimization terms based on graph smoothness and label
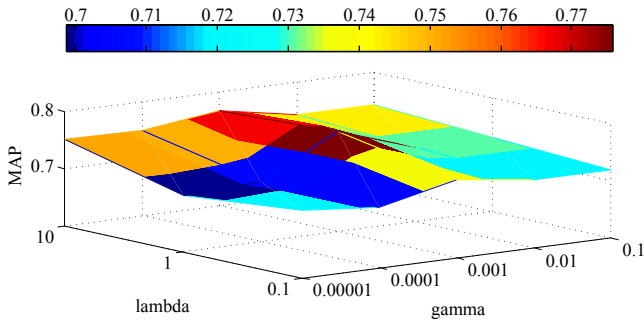
Figure 3: MAP of our method as a function of the combination of parameter λ and γ. This figure is generated when performing image ranking for a Task-II query on the IKEA group (see Section 6).

fitness. But it did not involve edge selection and sparsity penalties. (2) Edge Selective Information Propagation (ES-IP). We only use ℓ_1-norm regularizer to guide the individual edge selection while removing the $\ell_{2,1}$-norm regularizer from the objective function. (3) Relation Selective Information Propagation (RSIP), where we only keep the $\ell_{2,1}$-norm regularizer while removing the ℓ_1-norm regularizer. This essentially selects the relations at the macro level – relations of the same type will tend to be selected or deselected altogether instead of individually. (4) Our proposed Edge and Relation Selective Information Propagation (ERSIP).

To evaluate the performances of different methods mentioned above, we define the ground truths for each task as follows. (1) For Task-I where the query is a user, the relevant images are all the images that the user owns, likes or comments on; the relevant concepts are all the concepts of the relevant images, and the relevant users are the users who have contact, comment or favorite relations with respect to the query user (as explained in Section 4.2). (2) For Task-II where the query is a user plus a few concepts, the relevant images are those which are not only relevant to the user but also are annotated with the query concepts. (3) For Task-III where the query is an image plus its associated concepts, the relevant users are defined as those who have comment, own or favorite relations with respect to the query image.

To avoid the data scarcity of the evaluation, for Task-I and Task-II, we only select the users who have own 70 or more images as the query user. For Task-II, we employ the most frequent 3 tags appearing in the images of the query user as the query concepts which will be combined with the user as a combined query. Finally, for Task-III, we select the images which have 5 or more relevant users (as defined in the previous paragraph) as the query image and then employ its concepts as the query concepts. This results in totals of 197, 197, and 221 distinct queries for task I, II, and III respectively. In each query task, we randomly select 50% of edges which have connections with the relevant entities as the test set for performance evaluation. To achieve this, we remove these edges in the hybrid network, making them invisible in the information propagation. Given a ranking method, we expect the entities connected by these removed relations are still ranked at relatively higher positions.

In this work, we employ the Average Precision (AP) as the evaluation metric. Notably, each query of Task-I and Task-II is composed of three subtasks including image ranking, user ranking and concept ranking. For each subtask, we calculate AP for each query and then calculate the Mean Average

Precision (MAP) across all the queries for each social group. Finally, we average all MAP's across the 10 social groups and employ the Average MAP as the final evaluation metric of the subtask. For parameter setting, we use the cross validation to determine the appropriate parameter values for each method. Specifically, we vary the values of λ and γ on the grid of $\{10^{-1}, 10^0, 10^1\}$ and $\{10^{-5}, 10^{-4}, \ldots, 10^{-1}\}$ respectively, and then run 2-fold cross validation (parameter setting on the training set and performance evaluation on the held-out test set separately). The same parameter setting procedure is used for all other methods compared. Figure 3 shows the performance sensitivity under various parameter combinations for our method. It showed a relatively stable performance, within a variation range of 10%.

6.2 Performance Comparison

In this subsection, we compare the performance of different information propagation methods over the hybrid social media network. Figure 4(a)–4(c) show the performance (average MAP computed at the full length) for all the methods in comparison over three tasks respectively. From the results, we have the following observations: (1) The proposed ER-SIP method consistently beats all the other baseline methods for every task by a large margin, which demonstrates its effectiveness in propagating information across heterogeneous entities and relations. Moreover, it shows significant performance improvements on all social groups (except one) of different interest, ranging from geolocations, tourist sites, products, or life styles. It confirms the general benefits of the purposed hybrid social media network. (2) The edge selective methods including ERSIP, ESIP and RSIP outperform the EAIP on most of the tasks. This is due to the fact that the former methods are able to remove some noisy or redundant relations and edges by enforcing sparsity on the edge selection indicators while the latter simply fuses all the information in an aggregated manner without comparing their individual contributions. (3) The final proposed ERSIP method performs significantly better than the ESIP and RSIP methods. This implies that simultaneously selecting the relation types and the individual edges at multiple levels does improve the robustness of information propagation. Moreover, although ESIP enforces to select the fewest number of edges during the propagation, its performance is not promising. The reason may be that it ignores the correlation of the edges in the same relation type, and hence cannot sufficiently discover the most informative edges for information propagation. Similarly, RSIP is only able to distinguish the significance of each relation type, but may fail to discover the most useful edges for the query task. In Figure 4(c), we show the performance comparisons for each of the 10 social groups. Figure 5(a)–5(c) show the average MAP of our method at different return depth. As seen, the average MAPs are pretty high when the depth varies from 10 to 100 where all results are higher than 0.7. This clearly demonstrates that our method is able to rank the most relevant entities at top positions in the ranking list. Figure 6 shows the image ranking result of a Task-II query.

6.3 Effect of Edge Selection

The proposed information propagation method over the hybrid network is able to dynamically pick the best subset of edges for customized recommendation in response to the specific needs of different users, tasks and queries. Although
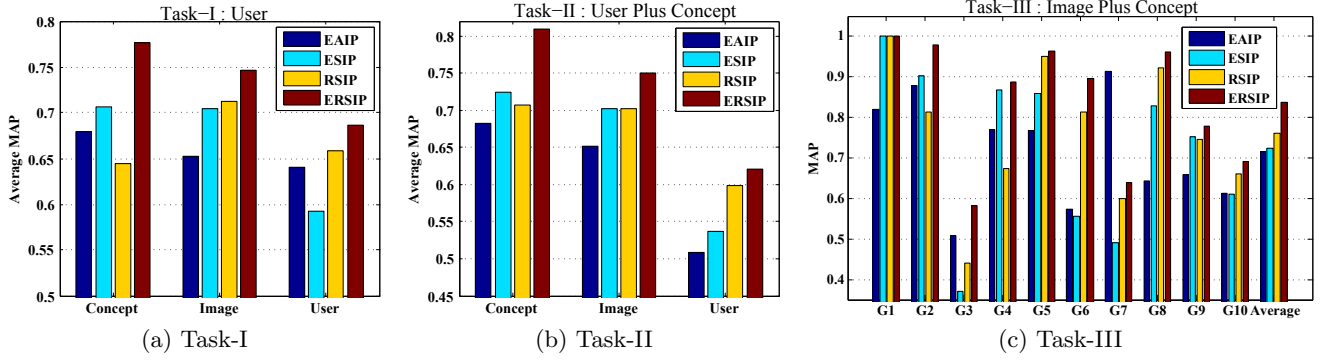
Figure 4: Performance of different methods over the three tasks. (a) and (b) show the average MAP comparison on the three subtasks of Task I, II respectively. (c) shows the MAP comparison for each social group, where the 10 groups from left to right in the horizontal axis are respectively (1)"Beijing", (2)"Fashion", (3)"IKEA", (4)"Korean food", (5)"New York", (6)"Painted", (7)"Sneaker", (8)"Taipei", (9)"Universal studio" and (10)"Wildlife", and the final one is the average MAP. This figure is best viewed in color.
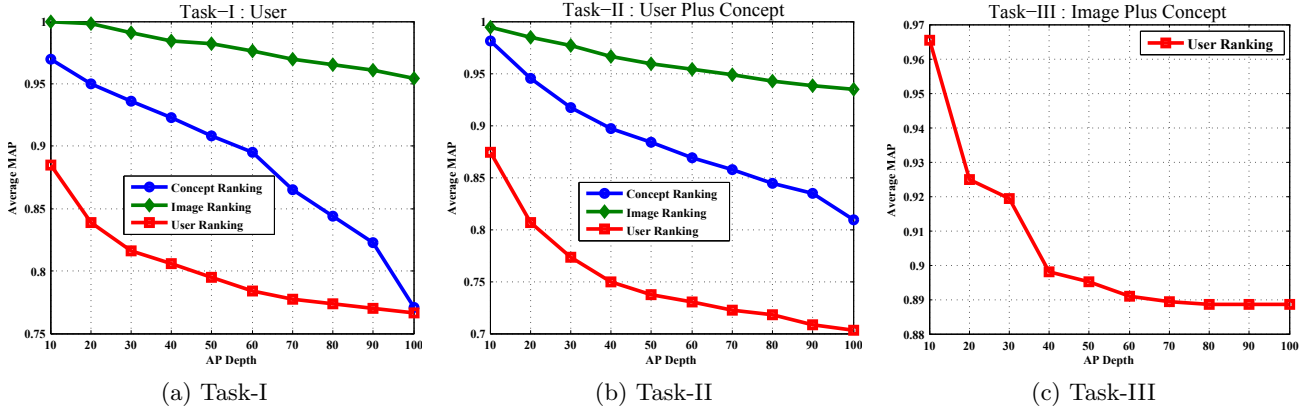

Figure 5: Average MAP of our method at variant return depths over the three tasks.

the experiment results in Section 6.2 have verified the benefits of edge selection, we perform analysis to gain further insights on how relations are selected for different queries. To this end, for each query, we calculate the average edge strength in each relation type (i.e., the sum of all $w_{ij}^k \alpha_{ij}^k$ in a relation type divided by the number of edges in this relation) and then average across all queries as an estimate of the contributions of each specific relation type. We calculate such information in the three tasks individually and plot the results in Figure 7(a)–7(c). As shown in Figure 7(a), if the query is user, the relations associated with the user will have more influence while other relations that are not associated with user tend to be less important. For example, the average edge strengths from the two types of image-image similarity relations are very tiny for user related queries, indicating that edges in these two relations are not informative for the query. Similarly, from Figure 7(b) and Figure 7(c) we can see that, if the query is composed of concept or image, the relations associated with these entities tend to be more important compared with those relations that do not contain the query entity. From these results obtained from different query tasks, we can see that the edge selection strategy is able to identify the relations that are closely related to the query, which further verifies the power of the edge selection.

7. CONCLUSIONS

We introduced a novel hybrid social media network, upon which the the heterogeneous types of entities, relations and the interactions among users, multimedia objects and concepts can be seamlessly integrated within a unified framework. To fully explore the diverse types of entities and relations in the hybrid network, we developed an edge selective information propagation method which can adaptively select a sparse subset of the most informative relations from the massively connected network so that the information can be best propagated through the activated paths in the network. Given a target user and a specific task, our system maps the task to relevant concepts and formulates the query. Then the query information is propagated throughout the selected edges so that the utility scores of other nodes can be predicted. The predictions can be used to recommend information in a task-specific personalized way for various applications like personalized album generation, friend/topic recommendation, target ad, etc. Although we evaluated the proposed model on a prototype based on Flickr, the proposed methods are general. They can be applied to enhance search and recommendation functions of other complex systems involving hybrid data types and relations.

For future work, we will investigate the application of the hybrid social media network in the task-specific social media community discovery which automatically groups related entities into a number of community clusters based on the diverse types of relations. Besides, we will also study efficient methods for scaling the proposed optimization framework to a large network.

Figure 6: Top 30 ranked images in response to a query composed of a specific user and three concepts. Images with green borders have half of their relations removed in the network. They are still successfully retrieved due to the effective information propagation through optimally selected paths in the hybrid social media network. This figure is best viewed in color.

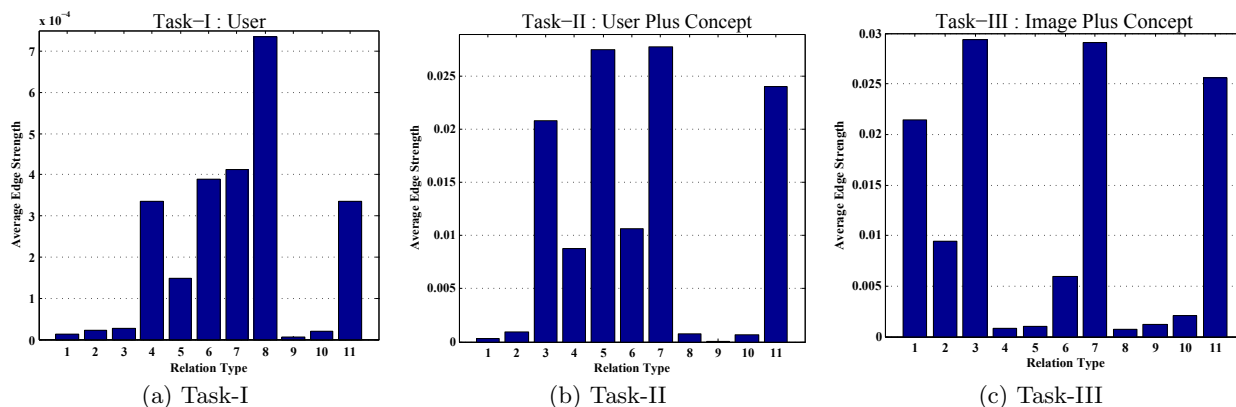(a) Task-I   (b) Task-II   (c) Task-III

Figure 7: Average edge strength in each type of relations over the three different tasks, where the 11 relations from left to right in the horizontal axis are respectively (1) *image-image SIFT similarity,* **(2)** *image-image mid-level similarity,* **(3)** *concept-image tagging relation,* **(4)** *user-concept ownership relation,* **(5)** *user-image comment relation,* **(6)** *user-image favorite relation,* **(7)** *user-image ownership relation,* **(8)** *user-user contact relation,* **(9)** *user-user favorite relation,* **(10)** *concept-concept co-occurrence relation,* **(11)** *user-image-concept tagging relation.*

8. ACKNOWLEDGEMENT

9. REFERENCES

[1] A. Argyriou, M. Herbster, and M. Pontil. Combining graph laplacians for semi-supervised learning. In *NIPS*, 2005.

[2] M. Ames and N. Naaman. Why we tag: motivations for annotation in mobile and online media. In *CHI*, 2007.

[3] S. Agarwal, K. Branson, and S. Belongie. Higher order learning with graphs. In *ICML*, 2006.

[4] J. Bu, S. Tan, C. Chen, C. Wang, H. Wu, L. Zhang, and X. He. Music recommendation by unified hypergraph : Combining social media information and music content. In *MM*, 2010.

[5] X. Chen, Q. Lin, S. Kim, J. Carbonell and E. Xing. Smoothing proximal gradient method for general structured sparse learning. In *UAI*, 2011.

[6] I. Dhillon. Co-clustering documents and words using bipartite graph partitioning. In *KDD*, 2001.

[7] I. Guy, etc. Harvesting with SONAR: the value of aggregating social network information. In *CHI*, 2008.

[8] Z. Guan, J. Bu, Q. Mei, C. Chen, and C. Wang. Personalized tag recommendation using graph-based ranking on multi-type interrelated objects. In *SIGIR*, 2009.

[9] J. He, M. Li, H. Zhang, H. Tong, and C. Zhang. Manifold-ranking based image retrieval. In *MM*, 2004.

[10] J. Han, Y. Sun, X. Yan, and P. Yu. Mining heterogeneous information networks. In *KDD*, 2010.

[11] W. Hsu, L. Kennedy, and S. Chang. Video search reranking through random walk over document-level context graph. In *MM*, 2007.

[12] M. Ji, J. Han, and M. Danilevsky. Ranking-based classification of heterogeneous information networks. In *KDD*, 2011.

[13] A. Langville and C. Meyer. A survey of eignvector methods for web information retrieval. *SIAM Review*, 2005.

[14] D. Lowe. Distinctive image features from scale-invariant keypoints. *IJCV*, 2004.

[15] D. Liu, X. Hua, L. Yang, M. Wang, and H. Zhang. Tag ranking. In *WWW*, 2009.

[16] D. Liu, S. Yan, Y. Rui, and H. Zhang. Unified tag analysis with multi-edge graph. In *MM*, 2010.

[17] J. Liu, W. Lai, X. Hua, Y. Huang, and S. Li. Video search re-ranking via multi-graph propagation. In *MM*, 2007.

[18] K. Mikolajczyk and C. Schmid. Scale and affine invariant interest point detectors. *IJCV*, 2004.

[19] J. Nocedal, etc. Numerical optimization. *Springer*, 1999.

[20] Y. Nesterov. Smooth minimization of non-smooth functions. *Mathematical Programming*, 2005.

[21] I. Ronen, etc. Social networks and discovery in the enterprise (SaND). In *SIGIR*, 2009.

[22] L. Torresani, M. Szummer, and A. Fitzgibbon. Efficient object category recognition using classemes. In *ECCV*, 2010.

[23] D. Zhou, etc. Ranking on data manifolds. In *NIPS*, 2003.

[24] D. Zhou, etc. Learning multiple graphs for document recommendations. In *WWW*, 2008.