2012 IEEE 12th International Conference on Data Mining

Towards Annotating Media Contents through Social Diffusion Analysis

Tong Xu[1], Dong Liu[2], Enhong Chen[1], Huanhuan Cao[2], Jilei Tian[2]
[1]*School of Computer Science and Technology, University of Science and Technology of China*
tongxu@mail.ustc.edu.cn, cheneh@ustc.edu.cn
[2]*Nokia Research Center, Beijing, China*
dong.e.liu@gmail.com, happia.cao@gmail.com, jilei.tian@nokia.com

***Abstract*—Recently, the boom of media contents on the Internet raises challenges in managing them effectively and thus requires automatic media annotation techniques. Motivated by the observation that media contents are usually shared frequently in online communities and thus have a lot of social diffusion records, we propose a novel media annotating approach depending on these social diffusion records instead of metadata. The basic assumption is that the social diffusion records reflect the common interests (CI) between users, which can be analyzed for generating annotations. With this assumption, we present a novel CI-based social diffusion model and translate the automatic annotating task into the CI-based diffusion maximization (CIDM) problem. Moreover, we propose to solve the CIDM problem through two optimization tasks, corresponding to the training and test stages in supervised learning. Extensive experiments on real-world data sets show that our approach can effectively generate high quality annotations, and thus demonstrate the capability of social diffusion analysis in annotating media.**

***Keywords*-Automatic annotation; common interests; diffusion maximization; social diffusion; social media.**

I. Introduction

Recent years have witnessed the boom of media contents on the Internet. On one hand, traditional media, like television or newspapers are accelerating digitized. On the other hand, thanks to the emergence of Web 2.0, various social media applications “allow the creation and exchange of user-generated content” [6]. Those huge amount of media contents raise significant challenges for efficient management and retrieval, e.g. keyword-based search on non-textual contents. Thus, automatic media annotation techniques are required. In this case, the annotation set is predefined by the content manager as *system-level annotations*, which target on describing the general properties of media contents.

From the machine learning perspective, the system-level automatic annotation can be casted as a (supervised) multi-label classification problem. In the literature, there are several sources to be used for automatic annotating. First, the metadata, usually provided by the publishers, is valuable, but “user generated contents” often lack high quality metadata. Second, the textual context, such as the title, description and comments. However, the context may not directly relate to the media, and even unavailable in some cases. Third, the personalized tags given by end users could be analyzed, but the ambiguity and irrelevance of personalized tags may influence the results [4], and also, processing the tags is not a easy work due to problems like cross-language translation [11]. Last but not least, some works learn annotations from the characteristics of the media content itself, e.g. low-level image/video features [12]. However, feature extraction from multimedia is computationally expensive and the learning effectiveness is restricted due to the gap between low-level features and high-level semantic annotations.

A distinct property of user generated media contents is that many of them are frequently shared in online communities such as social network services (SNS), which leading to a lot of *social diffusion records*. With this observation, in this paper, we propose a novel media annotation approach that adopts mining the social diffusion records, which are on hand or can be easily extracted from the web log of social network services. Our basic idea is motivated by a common phenomenon in our daily life that the diffusion of media contents usually reflects the *common interests* between sharers as well as the property of shared media. Researches have indicated that, on one hand, users tend to rely on their friends as “collector” and “filter” of interesting information rather than the other sources like system recommendations [10]; on the other hand, users with similar interests tend to build more tight social relationships, even constitute social communities, which in turn strengthens the influence from friend to friend [9].

Thus, we propose to use the common interests (CI) between users as the *intermediate* when learning social diffusion records to annotate. Note that common interests can be discovered from the media contents that users have shared. The more sharing concerns the same category, the stronger “common interest” is obvious. Reversely, users' sharing behaviors indicate that the shared media content meets their common interests. In order to describe such relationship, we present a novel *Common-interest-based (CI-based) diffusion* model. With this model, we need to solve two problems. First, how to learn the parameters to describe common interests. Second, how to use the model to annotate media contents. In this paper, borrowing the idea from traditional social diffusion analysis, we translate these problems into two optimization tasks, known as the *CI-based Diffusion Maximization* (CIDM) problem. For the model building, a diffusion maximization problem is formulated based on the set of annotated media contents with social diffusion records, and then solved to learn the “common interests”.

1550-4786/12 $26.00 © 2012 IEEE
DOI 10.1109/ICDM.2012.23

For the annotation, another diffusion maximization problem is formulated for the set of media contents to be annotated, and then solved to get the annotations. We conduct extensive experiments on several data sets and the experimental results clearly show that our approach can effectively generate high quality annotations for shared media contents, and thus demonstrate the capability of social diffusion analysis on media annotation.

The remainder of this paper is organized as follows. Section II describes our proposed CI-based diffusion model, and mathematically presents the CIDM problem. Followed in Section III, we present how to solve the two optimization tasks of the CIDM problem, respectively. Section IV presents the experimental results to validate the effectiveness of our approach. Section V discusses some related work, and finally, section VI concludes this paper.

II. Annotating Media Contents through CIDM

In this section, we will first introduce the CI-based diffusion model and then discuss how to annotate media contents through the common interest based diffusion maximization. And finally, we formally propose the formulation of the CIDM problem as two optimization tasks.

A. CI-based Diffusion Model

Social diffusion models are usually adopted to analyze the spread of information in social networks, e.g. the diffusion of media contents. A social network is depicted by a graph $G = \langle V, E\rangle$, where V is a set of nodes representing users, and E includes the edges or links between users. Each edge $\{e_{sr}|u_s, u_r \in V\}$ is augmented with a weight w_{sr}, which measures the probability of diffusion from u_s to u_r. On this graph, social diffusion can be simulated as stochastic processes. At any time, nodes are either *active*, which means they have the information, or *inactive*. Active nodes will propagate the information to their inactive neighbors with the probabilities defined on the corresponding edges. At the beginning of the process, some nodes have the information and these are known as "seed" nodes, i.e. activated initially. As time unfolds, more and more nodes become active.

As discussed in Section I, the diffusion of media contents between users indicates that the content more or less meets the common interests between users. Therefore, in our adapted common-interest-based diffusion model, $W = \{w_{sr}\}$, the edge weights in the graph G, will vary for different contents, and will relate to the "common-interest" parameters. Let T_i denote the set of annotations for the item i, $T_i \subset T$, where T is the complete set of pre-defined annotations. And $C = \{c_{sr}|u_s, u_r \in V\}$ denote the common interest parameters between pairs of users. Note that in this assumption, common interests are independently considered for each pair of users, and independent from different items. We assume w^i_{sr} (the superscript i indicates w_{sr} depends on the item i) can be approximated by $w^i_{sr} \approx corr(T_i, c_{sr})$ where $corr()$ denotes a correlation function between the item properties and the common interests. Details of the correlation function will be discussed in Section III-B.

B. CIDM based Media Annotation

Now let us turn to the social diffusion records of media contents. Still using the graph model, for the item i, its diffusion records can be summarized as one graph $G_i = \langle V_i, E_i\rangle$. Here V_i consists of the users who viewed the item, and E_i depicts how the item was diffused. Note that the edges in E_i indeed represent historical behaviors rather than "potential links", thus edges are directed and no circle exists, i.e. it is a *directed acyclic graph* (DAG).

As mentioned in Section II-A, social diffusion can be simulated as stochastic processes on a general social graph; Therefore, we try to "reproduce" the G_i's by social diffusion, rather than to predict how the diffusion may happen. Using our common-interest based diffusion model, where T_i and C are parameters, we can mathematically present the "reproducing" problem as follows: given the diffusion graph G_i, how to find out the proper C and T_i, so that G_i has the maximum likelihood, i.e.

$$(T_i^*, C^*) = \arg\max_{(T_i,C)} P(G_i|T_i, C), \tag{1}$$

where $P(G_i|T_i, C)$ can be represented as $P(\langle V_i, E_i\rangle\,|T_i, C)$.

We further propose to run the social diffusion simulation *only* on the graph G_i, i.e. diffusion may happen only on the fixed edges in E_i. This way, E_i is embedded into the diffusion processes. Thus, $\max P(G_i|T_i, C)$ can be simplified as $\max P(V_i|T_i, C)$, while the latter is virtually the traditional diffusion maximization task on the nodes V_i. Should be noted that "seed" nodes are no longer free variables here, while $P(V_i)$ indeed relies on three factors: G_i, T_i, and $C|_{E_i}$ denoting part of the common interest parameters C that are relevant to the edges in E_i.

C. CIDM Problem Formulation

We cast the automatic annotating task as a supervised learning problem and divide the CIDM problem into two optimization tasks, which correspond to the training and test stages in a typical supervised learning problem, respectively. For simplicity, the target function of the CIDM problem, i.e. the maximum expected diffusion, is denoted as $\max D(G_i, T_i, C|_{E_i})$.

Problem Formulation:

1) Training Stage – given a set of annotated media contents and their social diffusion graphs, i.e., a set of $\langle G_i, T_i\rangle$ which denotes the diffusion graph and annotations for item i, where $G_i = \langle V_i, E_i\rangle$, and $i \in I_a$ is a training sample, the problem is how to find a common interest parameter set C^* so as to maximize the diffusion of training samples, which can be formulated as follows:

$$C^* = \arg\max_{C} \sum_{i \in I_a} D(G_i, T_i, C|_{E_i}). \tag{2}$$

2) Test Stage – given a set of media contents to be annotated, and their social diffusion graphs $\{G_i\}$ where $i \in I_u$ is an item to be annotated, and $C|_{E_i}$ learnt from the training stage, the problem is how to find an annotation set T_i for each item i to maximize the likelihood of G_i, which can be formulated as follows:

$$T_i^* = \arg\max_{T_i} D(G_i, T_i, C|_{E_i}), \forall i \in I_u. \quad (3)$$

Indeed, both training and test stages utilize our CI-based diffusion model, but the optimization objectives and the control variables are different. In the training stage, we regard the common interest parameters C as the control variables to maximize the expected diffusion, and during this process, the proper C^* will be learnt. In the test stage, we utilize the common interests learnt during the training stage, and regard the annotations T_i as control variables to maximize the expected diffusion, then the solution T_i^* that leads to the maximum diffusion is the answer for annotation.

III. CI-based Diffusion Maximization

In our approach, both the optimization tasks in the training and test stages are dependent on common interest based diffusion simulation. In this section, we first explain how to simulate CI-based diffusion, and then introduce the details of the two optimization tasks, respectively.

A. Simulating Social Diffusion

As mentioned in Section II-B, diffusion simulation in our model is the same as in traditional models. Actually, the diffusion simulation in our approach can adopt any generic social diffusion model. For computation simplicity, here we adopt the state-of-the-art Steady State Spread (SSS) model [1] for diffusion simulation as an instance.

Generally, the activation probability of one node will be determined by two factors: 1) the statuses of its neighbors, and 2) diffusion probabilities from neighbors. Thus, the SSS model intuitively simulates the diffusion as a step-by-step evolving procedure, and the expectation of activation probability of node u_r at step k is

$$P_r(k) = 1 - \prod_{u_s \in N(r)} \left(1 - w_{sr} P_s(k-1)\right), \quad (4)$$

where w_{sr} and $P_s(k-1)$ correspond to the two factors mentioned above: w_{sr} denotes the diffusion probability from u_s to u_r, and $P_s(k-1)$ denotes the expectation of activation probability of u_s at the previous step; $N(r)$ includes the neighbors of node u_r. At step 0, those "seed" nodes have $P_r(0)$ set to 1, while all the others have 0. And then, step-by-step iterations will be executed.

With the SSS model defining the expectation of activation probability, the expected diffusion $D(G_i, T_i, C|_{E_i})$ in Section II-C can be defined as

$$D(G_i, T_i, C|_{E_i}) = \lim_{k\to\infty} \sum_{u_r \in \overline{V_i}} P_{r,i}(k), \quad (5)$$

where $\overline{V_i}$ includes all the nodes but those "seed" nodes in the graph G_i, and $P_{r,i}(k)$ denotes the activation probability that the user u_r has accessed the item i at the k-th step. Since all the G_i's are DAGs, it is easy to prove that according to the SSS model all the $P_{r,i}(k)$ will keep unchanged after limited steps, thus the convergence of $D(G_i, T_i, C|_{E_i})$ is ensured.

B. Introducing Common Interest into Diffusion Simulation

According to Eq. (4), activation probability $P_r(k)$ relies on the weights $\{w_{sr}\}$. Thus, we need to design a method for calculating w_{sr}, i.e., a proper $corr(T_i, c_{sr})$ function mentioned in Section II-A.

First, we need to determine the form of common interest parameter c_{sr}. Usually, multiple common interests may exist between users at different levels. Further, since we aim to leverage common interests to annotate the shared media contents, predefined annotations set T can be used to describe the abstract concept of "interest". Therefore, we propose to use a $|T|$-dimensional vector $\{c_{sr}^z\}$ to represent the "common interests" between users u_s and u_r, where c_{sr}^z indicates the level of the z-th ($1 \le z \le |T|$) "interest", or indeed the z-th predefined annotation. For normalization, we have a constraint that $\sum_z c_{sr}^z = 1$.

Also, the item-wise annotations T_i can be represented as a $|T|$-dimensional vector $\{t_i^z\}$ as well, in which each element t_i^z indicates whether the z-th predefined annotation is selected ($t_i^z = 1$) or not ($t_i^z = 0$). Based on these two $|T|$-dimensional vectors, we treat the edge weights on the graphs, or the diffusion probabilities, as an accumulation of multiple interests as:

$$w_{sr}^i = 1 - \prod_{z=1}^{|T|} (1 - c_{sr}^z \cdot t_i^z). \quad (6)$$

C. Learning Common Interests

By simulating CI-based diffusion we can now address the optimization problem of Eq. (2). If we simply expand Eq. (2) with Eqs. (4)–(6), the optimization problem will be very complex and difficult to solve since there are so many diffusion graphs and each contains thousands of edges. To simplify the optimization problem, we observed that each $D(G_i, T_i, C|_{E_i})$ in Eq. (2) is a sum of several $P_{r,i}(k = \infty)$ according to Eq. (5). Also, according to Eq. (4), if all the other variables are fixed, $P_{r,i}(k)$ is monotonically non-decreasing with respect to w_{sr}^i. Thus, the original optimization problem could be simplified as to maximize all the w_{sr}^i, which has a trivial solution $w_{sr}^i \equiv 1, \forall s, r, i : e_{sr} \in E_i$. Then, we relax the original problem as $\min \sum_{s,r,i: e_{sr} \in E_i, i \in I_a} (1 - w_{sr}^i)$, where I_a denotes the training sample set. Combined with Eq. (6) for weight calculation, we utilize the gradient descent method to deal with the optimization task as:

$$F = \sum_{s,r,i:e_{sr}\in E_i, i\in I_a} \left(\prod_{z=1}^{|T|}(1-c_{sr}^z \cdot t_z^i)\right)$$
$$c_{sr}^z(l+1) = c_{sr}^z(l) - \lambda \frac{\partial F}{\partial c_{sr}^z} \tag{7}$$
$$= c_{sr}^z(l) + \lambda \sum_{s,r,i:e_{sr}\in E_i, i\in I_a} \left(t_z^i \cdot \prod_{y\neq z}(1-c_{sr}^y(l)\cdot t_y^i)\right)$$

where l indicates the round of iterations. After each step, c_{sr} should be re-normalized.

It is worth noting that according to the above algorithm, we can calculate c_{sr} if and only if e_{sr} appears in $\{E_i\}$, which means at least one training item has been diffused from u_s to u_r. However, in practice, especially in the "cold-start" period, training samples are not enough so that $\{c_{sr}\}$ may be sparse. There are several alternative approaches to solve the problem, currently, we simply define $c_{sr}^z = 1/|T|, \forall z$ as a remedy for those edges that are needed but not covered in training samples.

D. Annotation Selection for Maximizing Diffusion

After learning common interest parameters from training samples, we can focus on the optimization task of the test stage, i.e., Eq. (3), which is indeed similar to a typical diffusion maximization problem, whereas the annotation set replaces "seed" node set as control variables. Theoretically, all the non-empty subset of T, i.e., $2^{|T|}-1$ kinds of combinations, can be enumerated as input to find out the maximal diffusion. But in practice, although the pre-defined annotation set T is usually not too large, the computational complexity of exhaustive search is still too expensive.

Therefore, when the efficiency is considered as an important factor, we can adopt a greedy algorithm, similar to the one to find the optimal seed nodes in the traditional diffusion maximization problem [1]. To be specific, we can test all the unselected annotations, and put the annotation that leads to the maximum incremental diffusion into the selected annotation set. This process will repeat until enough annotations are selected in order.

IV. Experimental Results

In this section, we report the results of extensive experiments to verify the effectiveness of our approach, then some further discussion is presented.

A. Experimental Setup

Raw data. We perform our experiments on two real-world data sets extracted from Douban.com, one of the most famous Chinese SNS which allows users to contribute comments on movies, books, and music[1]. Totally, we extract view logs of 107,164 users. For Douban Movie data set, 89,667 contents are extracted, and the number is 475,820 for Douban Book. For each data set, 2,500 media contents are randomly selected and labeled for experiments.

Diffusion records extraction. In typical online SNS, if one user "share" (e.g. mark a movie as "have seen it before" in Douban) some media contents, all the friends (or "followers", i.e. unidirectional friends such as in Twitter or Douban) will be notified, then they can continue the sharing to more friends. Thus, we extract diffusion trace based on the social relationship and the time of sharing, in a similar way to some previous work (e.g. [2]). Specifically, if u_s shares the media content i, and later on, u_r, a follower of u_s shares the same content, we extract a diffusion record of i from u_s to u_r, i.e., the edge e_{sr} exists in the diffusion graph G_i. With this method, we extract on average 1393.82 diffusion records for each item in the Douban Movie data set, and 167.18 for Douban Book.

Parameters. In our approach, three parameters are needed, namely, iteration stopping threshold and the step length for the gradient descent algorithm in training stage, and the iteration stopping threshold for the SSS model in test stage. In our experiments we empirically set the iteration stopping threshold and the step length for the gradient descent algorithm to be 0.0001 and 0.01, respectively. For the SSS model, since the iteration will onverge after limited steps, the stopping threshold is set to 0.

Evaluation metrics. To measure the effectiveness of our approach, precision (P) and recall (R) are calculated for the top-K annotation results, where P means the ratio of correct annotations in the test result, and R means the ratio of retrieved annotations in the ground truth. To reflect the trade-off between precision and recall, F-measure is also introduced as $F = \frac{2\times P\times R}{P+R}$.

Besides, the mean reciprocal rank (MRR) is also used for evaluation. As one item may have multiple annotations, we define the MRR-N as $\frac{1}{N}\sum_{i=1}^{N}\frac{1}{L_i}$, where N is the number of retrieved correct annotations, and $L_i, i = 1, 2, \ldots, N$ represents the rank of each correct annotation. This measure evaluates the ranking quality of test result.

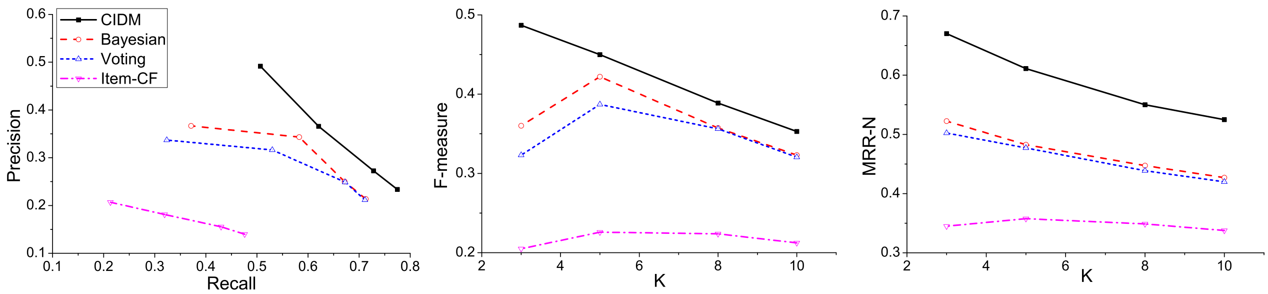
B. Baseline Approaches

Since we focus on the media annotation task that does not rely on metadata and/or textual description, traditional annotating approaches are not directly applicable for comparison. Here we select three baselines that rely on the same data source as our approach.

1) *Voting of users' preferences*: This approach is based on the assumption that the aggregated preferences of all viewers may indicate the feature of media content, where preferences are represented by annotations extracted from the individual view logs. Then, viewers vote with their preference, and the top-ranked annotations are selected.

2) *Item-based CF*: It is adapted from the basic item-based collaborative filtering (CF), with the assumption that "similar items attract similar users". In our experiments, the

[1] http://en.wikipedia.org/wiki/Douban/ (on Wikipedia)



(a) (b) (c)

Figure 1. Comparison results of different annotation methods on the Douban Movie data set, with different values of K: (a) precision vs. recall (both for top-K), (b) F-measure w.r.t. K, and (c) MRR-N w.r.t. K.

non-negative matrix factorization (NMF) technique [14] is utilized to represent items as vectors over latent factors by decomposing the original user-item matrix, and the similarity between items is calculated as the distance between the latent factor vectors. Then, given an item to annotate, training samples vote with their labels, and similarity is treated as weight, then the top-ranked annotations are selected.

3) *Diffusion based Bayesian approach*: It is a naive Bayesian approach based on the diffusion records as extracted in our CIDM approach. Assuming that all the pairwise diffusion records are independent, the annotating task can be translated into a maximum a posteriori probability (MAP) estimation problem as

$$\begin{aligned} P(t \mid i) &\propto \sum_{s,r:e_{sr}\in E_i} P(t \mid \langle u_s, u_r\rangle) \\ &= \sum_{s,r:e_{sr}\in E_i} \frac{P(\langle u_s, u_r\rangle \mid t)P(t)}{\sum_t P(\langle u_s, u_r\rangle \mid t)P(t)}, \end{aligned} \tag{8}$$

Here the likelihood $P(\langle u_s, u_r\rangle \mid t)$, which means the probability that u_s shares an item with annotation t to u_r, can be estimated as $P(\langle u_s, u_r\rangle \mid t) = \frac{N_t(u_s,u_r)}{N(u_s,u_r)}$, where $N(u_s, u_r)$ indicates the total number of diffusion records from u_s to u_r in the training data, and $N_t(u_s, u_r)$ indicates the number of diffusion records on training items that have annotation t. Besides, $P(t)$ can be obtained from the global frequency of annotation t in the training data.

C. Experimental Results

We first evaluate our CIDM approach as well as the three baselines on the Douban Movie data set. A standard 5-fold training-test is performed, and the average results of the 5 tests are reported with different values of K for more detailed comparison. As the average number of annotations of each movie is about 3 according to ground truth, and 95% of movies contain less than 5 annotations, thus, we set the value of K from 3, 5, 8, up to 10. The results are shown in Figure 1. We can clearly see that our approach consistently outperforms the baselines. Moreover, the improvement are more significant under smaller K. Note that in practice, a small K is preferred for the purpose of content management, then our CIDM approach seems more promising.

Table I
TWO CASES OF ANNOTATING RESULTS

Name	*Final Destination*	*We Were Soldiers*
Annotations	**America, Suspense, Horror**	**America, War, Action, History**
CIDM	**America**, **Horror**, **Suspense** Comedy, Love, China Action, Hongkong	**America**, **Action**, Classic Comedy, Sci-Fi, Drama **War**, Love
Bayesian	**America**, Comedy, Love **Horror**, Drama, **Suspense** China, Hongkong	**America**, Comedy, Love **Action**, Drama, Classic Hongkong, Sci-Fi
Voting	**America**, Comedy, Love Classic, Drama, Hongkong Action, Animation	**America**, Drama, Comedy Classic, **Action**, Love Hongkong, Sci-Fi
Item-CF	Love, Comedy, TV Series Japan, Hongkong, Animation Drama, **Horror**	Comedy, Love, Hongkong Japan, TV Series, Animation Drama, Suspense

Then, we further evaluate our CIDM approach and the three baselines on the data set extracted from Douban Book. Again, standard 5-fold training-test is performed, and we achieve similar results with the Douban Movie data set that our approach consistently outperforms the baselines under different K values, and for smaller K, the improvement of our approach is more significant. Due to the limitation of space, the figures of experimental results are omitted.

D. Case Study & Discussion

To better understand the experimental results, two exemplar movies are selected from the Douban Movie data set, which are shown in Table I. For each movie, its ground-truth annotations are first listed out, followed by the annotating results given by different approaches, where top $K = 8$ annotations are shown in order.

Obviously, in both cases our CIDM approach generates better annotations compared to baselines, since almost all the ground-truth annotations are retrieved and ranked at prior positions by the CIDM approach. When looking into the results, we could clearly find that our CIDM approach is able to not only find the frequent annotations such as “America”, but also find less frequent (and correct) ones like “Horror” or even rare (and correct) ones like “War”. On the contrary, the Bayesian and Voting approaches are good at finding out the frequent annotations, but are less capable in finding less frequent ones, or rank them at posterior positions (such as

"Horror" and "Suspense" in the result of *Final Destination* by the Bayesian approach). Moreover, since comedy movies or love stories are quite popular in the Douban movie data set, the baseline approaches (Bayesian, Voting, Item-CF) are always prone to assign these hot annotations, but this is not correct in the shown cases.. We may conclude that our CIDM approach is less influenced by varying frequency of annotations, thanks to the detection of common interests among minorities.

V. Related Work

A. Media Annotation

The most related work to ours is media annotation techniques. Based on the utilization of user factors, the work on automatic annotation could be roughly divided into two categories: generalized annotating and personalized annotating, while the former is more relevant to our work. As introduced in Section I, rich textual descriptions [16] or characteristics of multimedia [18] could all be utilized for automatic annotating. Besides, personalized tags are also used, for instance the tagging graph based approach [20].

For personalized annotating, the task is translated as to recommend the proper open-vocabulary tag to certain user. Usually, connections between user, item and tags are formed as tuple structure, and then methods like adapted PageRank [5] or tensor factorization [13] are utilized to "generate" the new connections. Some other works ask end users to input some "seed tags", and then other tags are recommended based on the correlation like "co-occurrence" frequency [15].

B. Social Diffusion Analysis

Another category of related work is social diffusion analysis. Though some work studies social diffusion in implicit networks (e.g. [19]), where potential links exist between any pair of nodes, most researches in this field focus on social diffusion in explicit networks, i.e. the links between nodes are given. Among them, the Independent Cascade (IC) model [7], following the assumption that influences from different nodes are mutually independent, is one of the most widely studied models. Other simulation models are also proposed, such as the maximum influence arborescence (MIA) model [3], the shortest-path based approximated model in [8], and the iteration-based Steady State Spread (SSS) model [1] that we adopted in our approach. With introducing topical factors, Tang *et al.* proposed a Topical Affinity Propagation (TAP) model [17] to represent the topic-level social influence on large networks.

VI. Conclusion

In this paper, we focused on automatic annotating task for media contents based on social diffusion analysis. As the "common interests" between users influence social diffusion behaviors, the annotating task was converted to the CI-based diffusion maximization (CIDM) problem. Then, we formulated the CIDM problem as two optimization tasks and proposed the solutions with CI-based diffusion simulation. The experimental results show that our approach can effectively generate high quality annotations and outperforms baseline approaches with a remarkable margin.

Acknowledgment

The work described in this paper was supported by grants from Natural Science Foundation of China (Grant No. 61073110), The National Major Special Science & Technology Projects (Grant No. 2011ZX04016-071), and Research Fund for the Doctoral Program of Higher Education of China (20113402110024), the Key Program of National Natural Science Foundation of China (Grant No. 60933013). Finally, thanks to the help from Nokia Research Center, Beijing.

References

[1] C. Aggarwal, A. Khan, and X. Yan, "On influential node discovery in dynamic social networks," in *SDM '11*, 2011, pp. 522–533.

[2] M. Cha, A. Mislove, and K. Gummadi, "A measurement-driven analysis of information propagation in the flickr social network," in *WWW '09*, 2009, pp. 721–730.

[3] W. Chen, C. Wang, and Y. Wang, "Scalable influence maximization for prevalent viral marketing in large-scale social networks," in *ACM SIGKDD '10*, 2010, pp. 1029–1038.

[4] J. Gemmell, M. Ramezani, T. Schimoler, L. Christiansen, and B. Mobasher, "The impact of ambiguity and redundancy on tag recommendation in folksonomies," in *ACM RecSys '09*, 2009, pp. 45–52.

[5] R. Jäschke, L. Marinho, A. Hotho, L. Schmidt-Thieme, and G. Stumme, "Tag recommendations in folksonomies," in *PKDD '07*, 2007, pp. 506–514.

[6] A. Kaplan and M. Haenlein, "Users of the world, unite! the challenges and opportunities of social media," *Business horizons*, vol. 53, no. 1, pp. 59–68, 2010.

[7] D. Kempe, J. Kleinberg, and E. Tardos, "Maximizing the spread of influence through a social network," in *ACM SIGKDD '03*, 2003, pp. 137–146.

[8] M. Kimura and K. Saito, "Tractable models for information diffusion in social networks," in *PKDD'06*, 2006, pp. 259–271.

[9] K. Lerman, "Social information processing in news aggregation," *IEEE Internet Computing*, vol. 11, no. 6, pp. 16–28, 2007.

[10] K. Lewis, M. Gonzalez, and J. Kaufman, "Social selection and peer influence in an online social network," *PNAS*, vol. 109, no. 1, pp. 68–72, 2012.

[11] T. Noh, S. Park, H. Yoon, S. Lee, and S. Park, "An automatic translation of tags for multimedia contents using folksonomy networks," in *ACM SIGIR '09*, 2009, pp. 492–499.

[12] G. Qi, X. Hua, Y. Rui, J. Tang, T. Mei, and H. Zhang, "Correlative multi-label video annotation," in *Multimedia '07*, 2007, pp. 17–26.

[13] S. Rendle, L. Marinho, A. Nanopoulos, and L. Schmidt-Thieme, "Learning optimal ranking with tensor factorization for tag recommendation," in *ACM SIGKDD '09*, 2009, pp. 727–736.

[14] H. Seung and D. Lee, "Algorithms for non-negative matrix factorization," *Advances in Neural Information Processing Systems*, vol. 13, pp. 556–562, 2001.

[15] B. Sigurbjörnsson and R. van Zwol, "Flickr tag recommendation based on collective knowledge," in *WWW '08*, 2008, pp. 327–336.

[16] Y. Song, Z. Zhuang, H. Li, Q. Zhao, J. Li, W. Lee, and C. L. Giles, "Real-time automatic tag recommendation," in *ACM SIGIR '08*, 2008, pp. 515–522.

[17] J. Tang, J. Sun, C. Wang, and Z. Yang, "Social influence analysis in large-scale networks," in *ACM SIGKDD '09*, 2009, pp. 807–816.

[18] L. Wu, L. Yang, N. Yu, and X. Hua, "Learning to tag," in *WWW '09*, 2009, pp. 361–370.

[19] J. Yang and J. Leskovec, "Modeling information diffusion in implicit networks," in *ICDM '10*, 2010, pp. 599–608.

[20] Z. Yin, R. Li, Q. Mei, and J. Han, "Exploring social tagging graph for web object classification," in *ACM SIGKDD '09*, 2009, pp. 957–966.