Topic-Level Opinion Influence Model (TOIM): An Investigation Using Tencent Microblogging[1]

Daifeng Li[1,*], Ying Ding[2], Xin Shuai[2], Guozheng Sun[4], Jie Tang[1], Zhipeng Luo[5], Jingwei Zhang[6], Tamy Chambers[2]

1 Dept. of Computer Science and Technology, Tsinghua University, Beijing, China

2 School of Informatics and Computing, Indiana University, Bloomington, IN, USA

3 Tencent Company, Beijing, China

4 Beijing University of Aeronautics and Astronautics, Beijing, China

5 Dept. of Electronic Engineering, Tsinghua University, Beijing, China

daifli_3000@163.com, {dingying, tischt, xshuai}@indiana.edu, gordon.gzsun@gmail.com, jery.tang@gmail.com, patrick.luo2009@gmail.com, iceboal@gmail.com.

***Corresponding Author:** Daifeng Li, daifli_3000@163.com

Abstract- **Text mining has been widely used in multiple types of user-generated data to infer user opinion, but its application to microblogging turns out to be difficult, since text messages are short and noisy, providing limited information about user opinion. Given that microblogging users communicate each other to form a social network, we hypothesize that user opinion is influenced by its neighbors in the network. In this paper, we infer user opinion on a topic by combining two factors: the user's historical opinion about relevant topics and opinion influence from his/her neighbors. We thus build a topic-level opinion influence model (TOIM) by integrating both topic factor and opinion influence factor into a unified probabilistic model. We evaluate our model in one of the largest microblogging sites in China, Tencent Weibo and the experiments show that TOIM outperforms baseline methods in opinion inference accuracy. Moreover, incorporating indirect influence further improves inference recall and f1-measure. Finally, we demonstrate some useful applications of TOIM in analyzing users' behaviors in Tencent Weibo.**

1. INTRODUCTION

Online social media, including microblogging, contain rich information about user opinion. This information is valuable to public services and initiatives such as marketing and political campaigns. For example, by understanding users' real time sentiment towards new events related to stock market and make correct prediction (There always exists a time-delay for users' information towards a certain event), investors are able to make better decisions(Zhang, 2010; Bollen, 2011; Mao, 2012). Several text mining and natural language processing techniques have been used to identify opinion within formal, well-written text. However, messages from microblogging are generally short and written informally. In addition, most users only follow others and seldom post new messages. It is therefore difficult to mine user opinion based solely on the user posts. Recent researches show that incorporating topic model into opinion mining can help to identify users' opinion distributions on different topics (Mei, 2007; Lin, 2009). Additionally, social influence can mitigate the data sparsity problem and help to improve the inference utilizing users' relationships (Guerra, 2011; Tan, 2011). Thus different from existing opinion mining and prediction algorithms, we mainly consider incorporating topic level opinion influence into our model. According to recent researches, incorporating topic model into opinion mining can help to identify users' opinion distributions on different topics (Mei, 2007; Lin, 2009). In another aspect, social influence can help to improve the inference of users' relationships, especially for the data sparsity problem (For example, no sufficient data for analyzing users' behaviors) (Guerra, 2011; Tan, 2011). Thus different from existing opinion mining and prediction algorithms, we mainly consider incorporating topic level opinion influence into our model.

[1] This paper is an extension of a short paper presented at CIKM 2012 Conference [1]

There exist a few highly active users (5% of all Tencent Weibo users), who frequently post messages to express their opinions around different topics and discuss them with their social network neighbors showing either agreement or disagreement attitude. In this paper, we focus on above users and aim to put the task of inferring their opinion into the context of a social network, where they exchange ideas with one another and influence each other's opinion. Particularly, such opinion influence is topic-specific, since users may have diverse opinions about different topics. Based on this we propose a Topic-Level Opinion Influence Model (TOIM), which integrates both topic factor and opinion influence into a unified probabilistic model. A user's individual messages and his/her communication records with neighbors are combined by TOIM to infer user opinion towards a specific object related to a discussion topic. We testify TOIM on Tencent Weibo, one of the biggest Chinese microblogging sites, and the results show that TOIM can infer user opinion more accurately upon the topic level in the social network, than some common baseline methods, if the user is active and frequently communicate with his/her neighbors. We also demonstrate some typical applications of TOIM in analyzing specific users' behaviors, attitudes, and influence on Tencent Weibo.

2. RELATED WORK

2.1 Sentiment Analyses and Opinion Mining

Online discussions surrounding specific entities (e.g., events or people) often include a mixture of features/topics related to that entity from different perspective. Pang, et al. (2002) studied the problem of classifying documents using overall sentiment identified through machine-learning methods. Hu, et al. (2004) mined opinion features from customers' online reviews. Liu, et al. (2010) analyzed document-level sentiment by first extracting the comment target, then predicting the polarity of opinions of the target. In recent researches, sentiment analysis is widely applied in social media. Gerald, et al. (2013) analyzed users' behaviors in different social medias, such as Twitter, Facebook, Amazon and etc, and provided frameworks and solutions to different social media to obtain optimized sentimental analysis results. Moghaddam, et al. (2013) applied different models, such as LDA, HMM, CRF, to realize opinion mining in online reviews, they used tweetfeel, Google Products, Stock Sonar and etc as testing data set to evaluate different methods. Vaileios, et al. (2013) provided a structured multi-task regularization method to infer users' voting intention on Twitter. Leon, et al. (2013) demonstrated the importance of making approaches specific to the microblog-genre, they used Twitter, Facebook and etc as example, to make a more accurate semantic annotation. Opinion detection and prediction algorithms have also been widely applied in market analysis, where Bollen et al. (2011) utilized the public mood mined from Twitter to predict the stock market. Gruhl, et al. (2005) predicted book sales by analyzing online chat. Mishne et al. (2009) analyzed blogger sentiment to predict movie sales. Liu et al. (2007) also used sentiment information from blogs to predict sales performance. However, a common difference of all those works with ours is that their proposed approaches only extract the overall sentiment of a document, and they do not distinguish different subtopics within a document, nor analyze the sentiment of a subtopic. Mei et al. (2007) proposed the Topic-Sentiment Mixture (TSM) model, which could reveal latent topical facets in the combination of a Weblog collection, the subtopics in the results of an ad hoc query, and their associated sentiments; authors used OpinMind (OpinMind.com) to label positive and negative sentiment polarity for queries and topics to formulate training data set. They generated formulas $P(z_{w,d,j}, S)$ to express the sentiment probability S of a certain entity w related with document d under a certain topic *zj*, $P(w \mid \theta_j)$ to represent entity-topic distribution and $P(w|\theta_S)$ for entity-sentiment distribution; EM updating formulas are also built for learning all required parameters. Lin et al. (2009) proposed a joint sentiment/topic (JST) model based on LDA (Blei, 2003; Rosen-Zvi, 2004; Nallapati, 2008; Ramage, 2009; Zhai, 2011), which could detect topic and sentiment simultaneously; they improved the original LDA model by adding sentiment-topic distribution: when selecting a topic z, a sentiment label l is also assigned based on sentiment-topic distribution, after that, a word w is chosen based on both z and l (a sentiment dictionary is used for mapping l from word w). Both TSM and JST tried to model topic and sentiment at the same time(they share similar mechanism, thus in this paper, we only take JST as baseline), however, social influence was not considered. Most existing research has mainly focused on identifying the sentiment polarity of sentences, or detecting a person's opinion from his/her textual information (Riloff, 2003; Kim, 2004; Liu, 2010). The main idea of these researches is to first build grammar rules, the domain features for sentiment analysis, then apply machine-learning algorithms such as SVM, CRFs, or propagation algorithm to learn those rules. The rules are mainly about identification of word's sentiment polarity (Kamps, 2004; Takamura, 2005), identification of subjective and objective sentiment (Wiebe, 2006; Mihalce, 2007; Su, 2009), identification of sentiment objects (Guerra, 2011), identification of opinion holders' attitudes towards sentiment objects (Bollen, 2011; Zhai, 2011), sentiment rules of word combinations (Bloom, 2007), sentiment rules of the context (Song, 2006), and so forth.

2.2 Social Influence in Sentiment Analysis

Social influence is an important research topic in social network analysis. Social influence occurs when one's emotions, opinions, or behaviors are affected by others (qualities-of-a-leader.com, 2013). One main purpose of social influence analysis is to detect and evaluate the existence and context of specific forms of social influence (Anagnostopoulos, 2008). Anagnostopoulos et al. (2008) focused on identifying and understanding social influence. They apply a statistical analysis method to identify and measure whether social influence is a source of correlation between the actions of individuals with social ties. Domingos and Richardson (2001) investigated social influence in the customer network. They propose a model to identify customer's influence between each other in the customer network. They (2002) also built a probabilistic model to mine the spread of influence for viral marketing. Similar works are to maximize the spread of influence through a social network, for example, Kempe et al. (2003) considered that influence maximization is a NP-hard problem, and applied an approximate method to solve it. Tang et al. (2009) identified social influence on different topics and proposed Topical Affinity Propagation (TAP) to model topic-level social influence. Liu et al. (2010) designed a LDA-based Social Influence model to detect influential relationships among individuals. Crandall et al. (2008) developed techniques for identifying and modeling the interactions between social influence and selection using data from online communities.

The rise of social media, such as Facebook and Twitters, put sentiment analysis in the context of a social network. For example, Tan e. al. (2011) utilized social connection to improve sentiment-classification based on the intuition that connected users are more likely to share similar opinions; they tested the model on Twitter dataset to identify who will support or against a political celebrity. Gurrea et al. (2011) applied transfer learning to utilize users' communication features to build opinion agreement graphs, thus inferring user opinion about each aspect of an event, for example, they use their model to analyze public's sentiment during a football game based on real time data from Twitter. However, they did not further define a series of topics related to particular entities or examine opinions on different topics. Furthermore, in the research on influence propagation, Liu et al. (2011) did suggest Conservation and Non-conservation methods for mining indirect influence in different heterogeneous networks, however, they focused mainly on user behaviors such as citing, following, and replying.

Those researches provide us a new perspective for investigating social opinion from the topic level. We further apply topic-level social influence to capture user opinion on different topics in heterogeneous social networks.

Simultaneously modeling social network structures, user behaviors, and user opinion preferences into a unified model, allows user opinion to be captured to a good degree of accuracy.

3. PROBLEM DEFINITION

Similar to Twitter, Tencent users can post messages of up to 140 Chinese characters and follow other users to read their messages. Two mechanisms are provided to facilitate users' interaction, i.e. ***forward*** (similar to retweet in Twitter) and ***comment*** (users can comment on messages by replying the original messages or mentioning authors of the original messages). We do not consider the ***forward*** interaction in our paper, which generally contains little information about the forwarding user's own opinion. Instead, we mainly focus on detecting opinion influence from users' ***comments,*** which contain rich textual information about user opinion. Two types of objects (users and messages) and multiple types of relations (user posts/comments on message, user replies to another user) constitute a heterogeneous network built on Tencent Weibo. A typical scenario of Tencent Weibo is introduced in Figure 1 to explain our model. In Tencent Weibo, we name all kinds of messages as "weibo" uniformly.

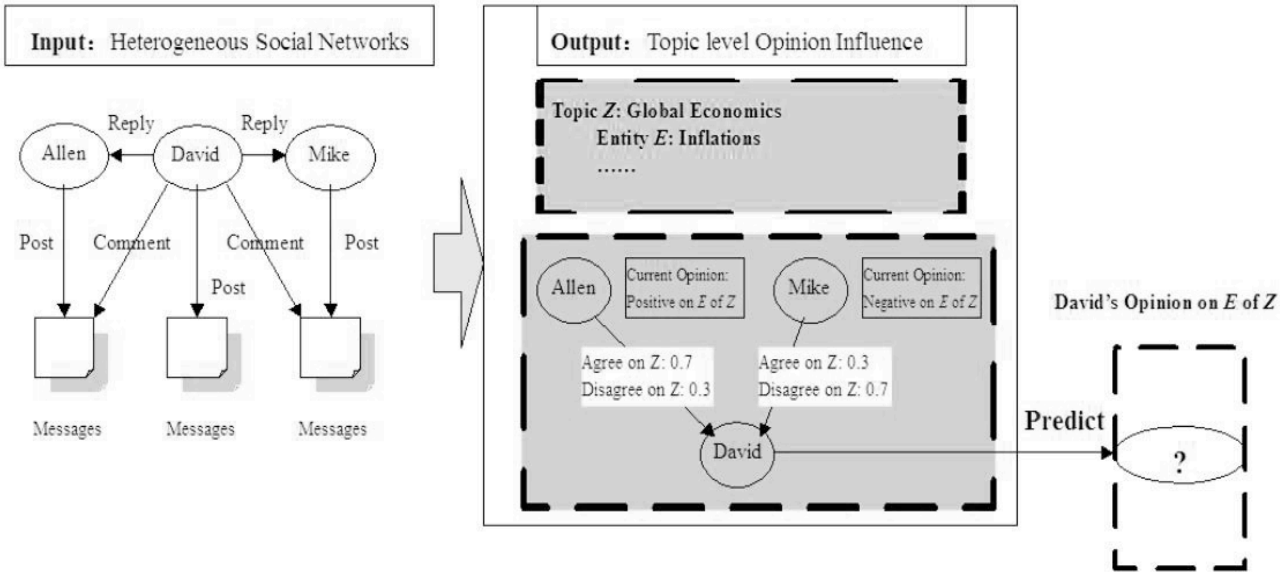


Figure 1: Motivating Example. This Figure raises the question that can we provide David's opinion towards "inflation" based on his historical behaviors and his communication records with his neighbors, such as Allen and Mike. "Topic" is considered as a latent variable for better understanding users' behaviors.

In Figure 1, David comments on both Allen and Mike's messages and replies to both of them on the topic about inflation of the **global economics**. Given the topological and textual information, we generate a topic opinion influence network, with David at the center influenced by Allen and Mike. the historical communication records among David, Allen and Mike are taken into consideration to calculate the opinion influential relationship about the topic of global economics among. Specifically, with regards to the topic of global economics, we statistically count how many times they agree with each other, and how many times they disagree with each other on that topic, based on which an opinion influence value (agree/disagree probability) can be calculated between David, Allen and Mike. To be specific, their frequency of agreement (seen as positive influence) will be high if they share common opinion preferences; otherwise, their frequency of disagreement (seen as negative influence) will be low. Finally, if Allen and Mike have provided their opinions on inflation of Global Economic, then David's opinion on the same object related Global Economics can be inferred by jointly considering his own opinion preference and opinion influence from Allen and Mike.

Based on the above observation, the problem of inferring a user's opinion regarding a certain topic can be solved by jointly considering the users historical opinions and opinion influences from his/her neighbors. We propose TOIM (see Figure 2) to model that process, which is a multi-layer graphical model consisting of four groups of random variables. In the ***User Layer***, we use $U = \{u_1, u_2, ...u_V\}$ to denote all users in Tencent Weibo, who can comment on each other's messages. Since ***forward*** actions are considered as an important index to measure users' opinion agreement in many researches, but seldom researches are related with inferring users' influential relationships from analyzing the content of their communication records. In our research, we mainly focus on how to measure users' opinion influence based on their comment records(users do not simply repeat the original message to others, but make comments on it). Thus in our proposed model, in order to exclude the influence of forward behaviors, only behaviors related with *comment* are considered to connect users to form social networks. In the ***Message Layer***, we use $N = \{n_1, n_2, ...n_X\}$ to denote all entities or objects occurring in all messages represented by $M = \{m_1, m_2, ...m_D\}$. Some NLP and statistical techniques are used to detect opinions about $n_i (i \in [1, X])$. In the ***Topic Layer***, we use $T = \{t_1, t_2, ...t_K\}$ to denote all topics discussed by users. Those topics are latent variables, which cannot be observed or detected. In the ***Opinion Layer***, o_{ij} represents the opinion of user $u_i (i \in [1, V])$ about topic t_j. The main difference between TOIM and other opinion mining models is that the *opinion influence* is considered when inferring a user's opinion about a certain object (entity).

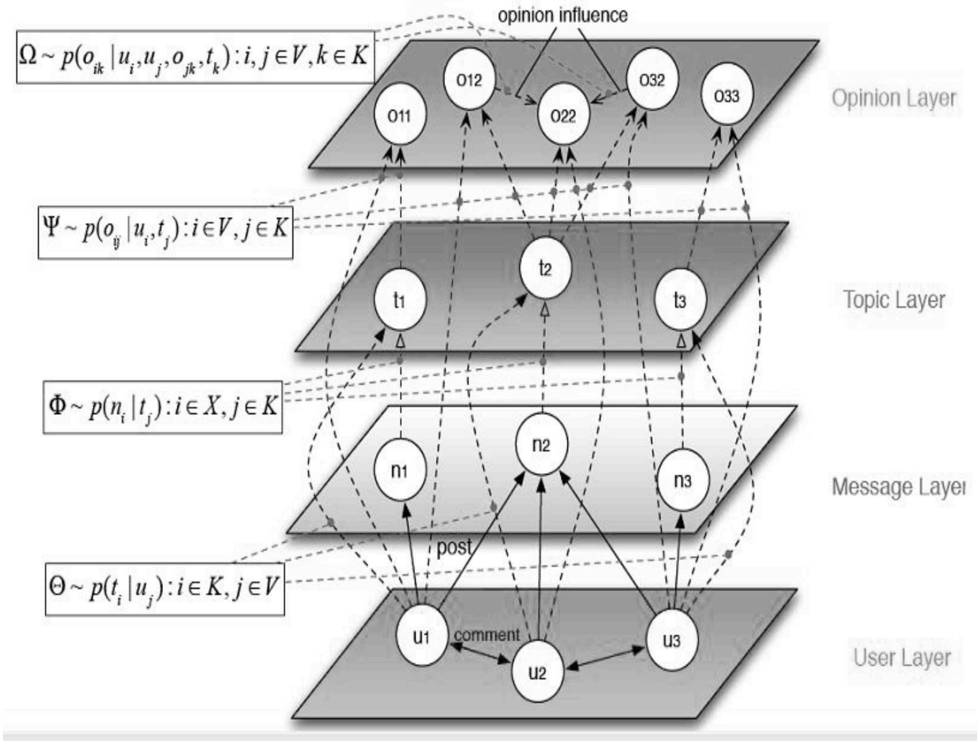


Figure 2: The framework of TOIM. There are four layers: user layer, message layer, topic layer and opinion layer. Users posted new messages and comment on others' messages. Some important entities occur in messages and belong to some topics. Given a new entity, some users' opinion about the entity is decided by both their historical opinion about the entity related topic and the opinion influence around this topic from his neighbors.

Four probabilistic matrices are defined to connect multiple variables from the four layers. Message Layer and Topic Layer are connected by $\Phi = \{\phi_{ij}\}_{K \times N}$, where ϕ_{ij} denotes the probability that object n_j belongs to topic t_i. User Layer and Topic Layer are connected by $\Theta = \{\theta_{ij}\}_{V \times K}$ where θ_{ij} denotes the probability that topic t_j is selected by user u_i. Opinion Layer, Topic Layer and User Layer are connected by two matrices. One is $\Psi = \{\psi_{i,ox}^{k}\}_{K \times V \times 3}$ where $\psi_{i,ox}^{k}$ denotes the probability that user u_i prefers opinion ox given topic t_k and $ox \in \{-1, 0, +1\}$ (-1 represents negative, +1 represents positive, 0 represents no opinion). The other is $\Omega = \{\omega_{i,j,agree}^{k}\}_{K \times V \times V \times 2} \cup \{\omega_{i,j,disagree}^{k}\}_{K \times V \times V \times 2}$, where $\omega_{i,j,agree}^{k}$ denotes the probability that u_i agrees with u_j's opinion and $\omega_{i,j,disagree}^{k}$ denotes the probability that u_i disagrees with u_j's opinion on topic t_k. Thus Ω represents the agreement/disagreement opinion influence. Besides, $S = \{s_{i,j,agree}^{k}\}_{K \times V \times V \times 2} \cup \{s_{i,j,disagree}^{k}\}_{K \times V \times V \times 2}$, where $s_{i,j,agree}^{k}$, $s_{i,j,disagree}^{k}$ are defined as the confidence/strength of $\omega_{i,j,agree}^{k}$ and $\omega_{i,j,disagree}^{k}$ respectively. Finally, In order to reduce the high dimensions of $V \times V$, we only consider 2,000 high active users and their followers are limited into 1,000 most active and related users. The details will be introduced in the next section. the opinion mining problem can be formulated as a mapping function as follows:

$$f(u_i, n_j, \Phi, \Psi, \Omega, S) \rightarrow o_{u_i}^{n_j} \tag{1}$$

where $o_{u_i}^{n_j}$ denotes the opinion of u_i about a query entity n_j.

The implementation of TOIM is composed of two phases. In the learning phase, Θ, Ψ, Ω and *S* are learned simultaneously utilizing statistical and NLP techniques. In the inference phase, given an entity n_q, the opinion about n_q is inferred using Φ, Ψ, Ω and *S*. For instance, in Figure 2, the opinion of u_2 about entity n_2 that belongs to topic t_2 is affected by two factors: u_2's historical opinion about t_2, and the opinion influence from u_2's neighbors u_1 and u_3. Although Θ is actually not used in opinion inference, it provides useful information about user topic preference and can be used for other topic-level opinion related analysis.

4. METHODOLOGY

In this section, we illustrate the details about how to learn the parameters of users' opinion influence, and make inference. Section 4.1 gives the computational equations of five parameters of TOIM(Θ, Φ, Ψ, Ω and S) separately and introduces the unified probability model to estimate the five parameters. Section 4.2 demonstrates how to use the learned parameters to infer a certain users' opinion taking opinion influence of his neighbors into consideration.

4.1 Learning

We formulize the calculation of five parameters of TOIM(Θ , Φ , Ψ , Ω and *S*) separately; and then propose the unified probability model to estimate them. The definitions of all notations used in this paper are summarized in Table 1:

Table 1. Notations

Notations	Definitions
U	The set of users in Tencent Weibo, assume $U = \{u_1, u_2, ...u_V\}$
M	The set of weibos in Tencent Weibo, assume $M = \{m_1, m_2, ...m_D\}$
N	The set of noun entities, assume $N = \{n_1, n_2, ...n_X\}$
T	The set of topics, assume $T = \{t_1, t_2, ...t_K\}$
ox	$ox \in \{-1, 0, +1\}$ (-1 represents negative, +1 represents positive, 0 represents no opinion)
o_{ij}	User u_i 's opinion distribution on topic t_j , $o_{ij} \in [-1, +1]$, where $o_{ij} < 0$ represents negative opinion, $o_{ij} > 0$ represents positive opinion; the bigger $\vert o_{ij} \vert$ means higher confidence.
$o_{u_i}^{n_j}$	User u_i 's opinion on entity n_j , $o_{u_i}^{n_j} \in [-1, 0, +1]$
θ_{xz}	User-Topic distribution, $x \in U$, $z \in T$
ϕ_{zn}	Topic-Entity distribution, $z \in T$, $n \in N$
$\psi_{i,+1}^k$	The probability of user u_i 's positive opinion on topic t_k
$\psi_{i,-1}^k$	The probability of user u_i 's negative opinion on topic t_k
$\omega_{i,j,agree}^k$	The probability of user u_i agrees with u_j 's opinion on topic t_k
$\omega_{i,j,disgree}^k$	The probability of user u_i disagrees with u_j 's opinion on topic t_k
$s_{i,j,agree}^k$	The confidence/strength of $\omega_{i,j,agree}^k$
$s_{i,j,disagree}^k$	The confidence/strength of $\omega_{i,j,disgree}^k$
Θ	The matrix of author-topic distributions. $\Theta = \{\theta_{xz}\}_{x \in U, z \in T}$
Φ	The matrix of topic-entity distributions. $\Phi = \{\phi_{zn}\}_{z \in T, n \in N}$
Ψ	A structure to record users' opinion preference on different topics; $\Psi = \{\psi_{i,j}^k\}_{K \times V \times 2}$
Ω	A structure to record users' opinion influence on different topics; $\Omega = \{\omega_{i,j,agree}^k\}_{K \times V \times V \times 2} \cup \{\omega_{i,j,disagree}^k\}_{K \times V \times V \times 2}$
S	A chain structure to record the confidence/strength of Ω ; $S = \{s_{i,j,agree}^k\}_{K \times V \times V \times 2} \cup \{s_{i,j,disagree}^k\}_{K \times V \times V \times 2}$

4.1.1 Opinion Detection

Opinion detection captures the opinion word *ow(ni)* for an entity *ni* and judges the polarity of *ow(ni)* in the context of the message where *ni* occurs. The basic method of opinion detection is proposed by several previous studies(Riloff, 2003; Kim, 2004; Kamps, 2004; Takamura, 2005; Liu, 2010), and the process can be summarized as follow: First, a parse tree[2] developed by FudanNLP group is constructed to exhibit the syntactic structure of a sentence and dependency relations between Chinese words. Second, *ni* and *ow(ni)* are identified using parse tree structure. Third, the polarity of *ow(ni)* is determined by searching the Chinese sentiment word lexicon provided by Tsinghua NLP group[3], which consists of 5,567 positive and 4,479 negative words. Finally, two grammar rules are applied to identify the sentimental relation: (1) whether there exists a negation word, like not or don't, and (2) whether there exists an adversative relation between *ni* and *ow(ni)*, like but or however. As an example, the parse tree for "This product is good and cheap" is shown in Figure 3. All words are organized into a tree structure according to their grammar dependency relationship, where the label under each word represents corresponding part of speech tag. Here the entity is "product" while the opinion words are "good" and "cheap". By searching the sentiment lexicon we conclude that the opinion is positive.

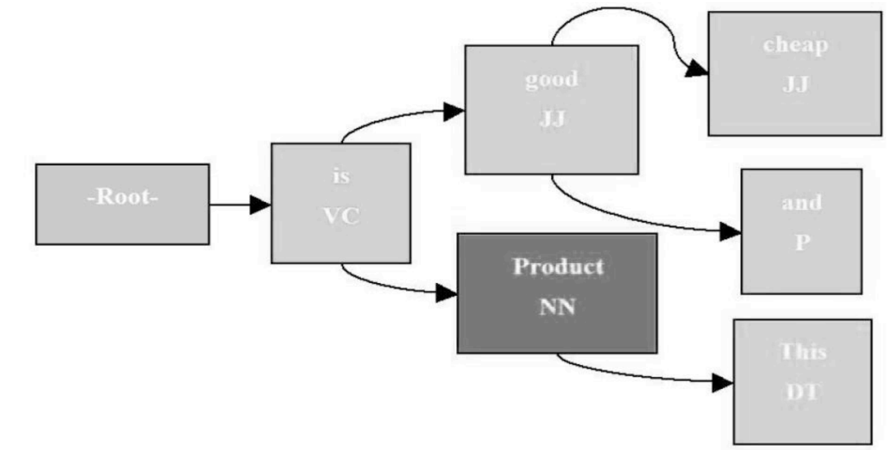


Figure 3: The parse tree structure of an example sentence. All words are organized into a tree structure according to their grammar dependency relationship, where the label under each word represents corresponding parts of speech.

The parse tree works well when the sentence is complete and grammatically correct. However, our preliminary study shows only a small portion (around 5% to 10%) of messages can be successfully parsed while most other messages are short and noisy, because the corresponding *ow(ni)* cannot be found. To overcome this sparseness of data, we detect the pairs of (*ni, ow(ni)*) both within a single message and among all messages. For each $n_i \in N$, we identify all adjective/verbs *wi* that co-occur with *ni* by scanning all messages and pick out the top 20 most frequent co-occurred words S = [$\omega_1,...\omega_{20}$]. Then *ow(ni)* can be determined using the following equation:

$$o\omega(n_i) = \arg\max_{\omega_j \in S} \frac{CO(n_i,\omega_j)}{DIS(n_i,\omega_j)} \quad (2)$$

where $CO(n_i,\omega_j)$ denotes the frequency of co-occurrence of n_i and ω_j and $DIS(n_i,\omega_j)$ denotes the average distance between n_i and ω_j.

4.1.2 User-Topic and Topic-Entity Distribution (Θ and Ψ)

2 http://code.google.com/p/fundannlp/download/list
3 http://nlp.csai.tsinghua.edu.cn/site2/index.php?

Gibbs sampling is used to estimate Θ and Ψ, with two prior hyperparameters α and β, respectively (Asmussen, 2007; Blei, 2006; Gilks, 1995; Walsh, 2004). We construct a Markov chain that converges to joint posterior distribution on random variables topic *z*, user *x* and noun *w*, which can then be used to infer Θ and Ψ (Griffiths, 2004). The transition between successive stats of Markov chain results from repeatedly drawing z from its distribution conditioned on all other variables. Assuming that *ui* posted a message and *uj* replied to *ui* by commenting on *ui*'s message. If the *lth* noun found in *ui*'s message is n_h, we can sample a topic for *ui* based on Equation 3.

$$P(z^l = t_k \mid x = u_i, \omega = n_h, z^{-1}) \propto \frac{C_{xz}^{-l} + \alpha}{\sum_{z \in T} C_{xz}^{-l} + K\alpha} \frac{C_{z\omega}^{-l} + \beta}{\sum_{\omega \in W_N} C_{z\omega}^{-l} + N\beta} \tag{3}$$

where $z^l = t_k$ denotes the assignment of the *l*th noun into topic t_k and z^{-1} denotes all topic assignments not including n_h. C_{xz}^{-l} and $C_{z\omega}^{-l}$ denote the number of times topic *z* is assigned to user *x*, and noun *w* is assigned to topic *z* respectively, not including the current assignment for the *l*th noun. For user u_j, if n_h also occurs in u_j's reply message, n_h is also assigned to topic t_k and t_k is assigned to user u_j. For all other nouns in u_j's replying message, the assignment of words and topics are calculated as shown in Equation 3. The final Θ and Ψ can be estimated by:

$$\theta_{xz} = \frac{C_{xz} + \alpha}{\sum_{z \in T} C_{xz} + K\alpha}, \phi_{zn} = \frac{C_{zn} + \beta}{\sum_{\omega \in N} C_{zn} + N\beta} \tag{4}$$

4.1.3 Topic-Level User Opinion Distribution (Ψ *)*

Topic-Level user opinion distribution characterizes the relative frequency of positive or negative opinions from one user about a certain topic. Similarly with studies of Mei (2007) and Lin (2009), we define two counters $C_{i,+1}^k$ and $C_{i,-1}^k$, *i* =1,…,*V*; *k* = 1,…,*K* to record the number of times user u_i express positive or negative opinions towards topic t_k by scanning all u_i's message. Then Ψ can be estimated as:

$$\psi_{i,+1}^k = \frac{C_{i,+1}^k}{C_{i,+1}^k + C_{i,-1}^k}, \psi_{i,-1}^k = \frac{C_{i,-1}^k}{C_{i,+1}^k + C_{i,-1}^k} \tag{5}$$

4.1.4 Topic-Level Opinion Influence (Ω *, S)*

The topic-level opinion influence quantifies how much two users agree or disagree on the same topic. To build the opinion influence relations among users, we mainly follow Liu's study(2010). In many cases, users discuss on the same topic but may focus on different entities. When we judge whether two users agree or disagree on a topic, we need to consider the sentimental relation between the entities they discuss within this topic. For instance, user u_i supports the US Government while user u_j dislikes Gaddafi. Both US Government and Gaddafi belong to the same politics topic, but exhibit an antagonistic relation. We can conclude that u_i and u_j agree with each other in politics. However, if u_i supports Obama while u_j supports Romney, we conclude that u_i and u_j disagree with each other in politics. To capture the sentimental relations between different pairs of entities that fall into the same topics, we run LDA algorithm on the training corpus and set the number of topics as 50. For each topic, we select the top 20 most frequently appearing nouns (entities) and construct pairs of entities. We then manually label the sentimental relation for each pair and obtain 2,104 labeled pairs. We use $OSR(n_i;n_j)$ to denotes the sentimental relations between n_i and n_j, with $OSR(n_i;n_j)$ = 1 representing coherent sentiment and $OSR(n_i;n_j)$=-1 representing antagonistic sentiment.

We also define two counters $C^k_{i,j,agree}$ and $C^k_{i,j,disagree}$ to record the number of times u_i and u_j agree or disagree on topic *k* by scanning all their communication messages. Specifically, if $o^{n_i}_{u_i} \cdot o^{n_j}_{u_j} \cdot OSR(n_i,n_j) > 0$; $n_i, n_j \in t_k$, u_i and u_j agree on topic t_k; otherwise, u_i and u_j disagree on topic t_k. Therefore, Ω can be estimated as:

$$\omega^k_{i,j,agree} = \frac{C^k_{i,j,agree}}{C^k_{i,j,agree} + C^k_{i,j,disagree}}, \omega^k_{i,j,disgree} = \frac{C^k_{i,j,disagree}}{C^k_{i,j,agree} + C^k_{i,j,disagree}}$$

(6)

In addition to the type of opinion influence, we also need to quantify the strength of opinion influence. For instance, if u_i and u_j only communicate once and agree with each other, the strength of influence is very low. For user u_i, we first assume that all neighbors who have discussed topic t_k with u_i constitute a set $ON(u_i,t_k)$, and then we calculate the strength of influence from $ON(u_i,t_k)$ on u_i:

$$s^k_{i,j,agree} = \frac{C^k_{i,j,agree}}{\max(\rho, \sum_{u_j \in ON(u_i,t_k)} C^k_{i,j,agree})}, s^k_{i,j,disagree} = \frac{C^k_{i,j,disagree}}{\max(\rho, \sum_{u_j \in ON(u_i,t_k)} C^k_{i,j,disagree})} \quad (7)$$

where ρ denotes a threshold of the minimum agreement/disagreement frequency.

The above detection of opinion influence works well when $OSR(n_i;n_j)$ is known. However, the values of

OSR$(n_i\,;n_j)$ are unknown for most pairs of $(n_i\,;n_j)$. Therefore, to quantify the opinion relation between u_i and u_j, other contextual information is utilized to generate the following Opinion Agreement Index (*OAI*) that can be used to quantify the opinion influence of u_j on u_i:

$$OAI(u_i,u_j) = a \times Influence(u_i) + b \times Tightness(u_i,u_j) + c \times similarity(u_i,u_j) \tag{8}$$

where $Influence(u_i)$ is the normalized function of u_i's followers, $Tightness(u_i,u_j)$ is the normalized function of the interaction (ie. comment) frequency between u_i and u_j, and $similarity(u_i,u_j)$ is the cosine similarity between θ_i and θ_j. a, b and c are assigned as 0.6, 0.3 and 0.1, respectively, based on empirical knowledge. $OAI(u_i,u_j)$ is generally normalized for u_j:

$$NOAI(u_i,u_j) = \frac{OAI(u_i,u_j)}{\sum_{u_j \in ON(u_i,t_k)} OAI(u_i,u_j)} \tag{9}$$

If u_j comments on u_i's one message and opinions influence can not be determined, then $NOAI(u_i,u_j)$ can be used to approximate $s^k_{i,j,agree}$.

4.1.5 Parameter Estimation

Five parameters of TOIM (Θ, Φ, Ψ, Ω and *S*), are estimated on a training corpus under a statistical sampling and counting framework that simulates users' communication process. Given a user u_i who wants to post a message, u_i first chooses a topic t_k from his/her topic distribution depending on Θ, and then selects an entity *nx* associated with t_k depending on Φ, and finally expresses his/her opinion $o^{n_x}_{u_i}$ towards *nx*. After that, Ψ is updated according to the value of $o^{n_x}_{u_i}$. Another user u_j replies to u_i by commenting on u_i's message. The similar random process applies to u_j and u_i tends to select the same topic t_k as u_i. Finally, Ω and *S* are updated based on both of $o^{n_x}_{u_i}$ and $o^{n_x}_{u_j}$. The learning process is listed in Algorithm 1.

Input: ***U, M***

Output: $\Theta, \Phi, \Psi, \Omega$ *and S*

Initiation: Iteration *Iter*

Pre1: Generate distinct word list *N*;

Pre2: Construct parse tree for $m_i \in M$;

Pre3: Calculate $o\omega(n_i)$ for $n_i \in N$, based on Equation 2;

Pre4: Calculate *OSR* (n_i, n_j), $n_i, n_j \in N$

Pre5: Calculate *NOAI*(u_i, u_j), $u_i, u_j \in U$ based on Equation 9;

Start:

For e=1 : *Iter* do

For m_i in *M* do

Find the user u_i who posted m_i; Find all comments CM_i on m_i;

For n_i in m_i do

Sample topic z_i based on Equation 3;

Detect u_i's opinion $o_{u_i}^{n_i}$ of n_i based on parse tree or Equation 2;

If $o_{u_i}^{n_i}$ ==+1, $C_{i,+1}^{z_i} += 1$; else $C_{i,-1}^{z_i} += 1$;

For m_j in CM_i do

Find the user u_j who posted m_j;

For n_j in m_j do

If $n_j == n_i$, set $z_j = z_i$; else sample topic z_j based on Equation 3;

Detect u_j's opinion $o_{u_j}^{n_j}$ of n_j based on parse tree or Equation 2;

If $o_{u_j}^{n_j} == +1$, $C_{\mathrm{j},+1}^{z_j} += 1$; else $C_{\mathrm{j},-1}^{z_j} += 1$;

If $z_i == z_j$ then

If $o_{u_i}^{n_i} \cdot o_{u_j}^{n_j} \cdot OSR(n_i, n_j) > 0$, $C_{i,j,agree}^{z_i} += 1$; else $C_{i,j,disagree}^{z_i} += 1$;

End

If $o_i == NULL$ or $o_j == NULL$ or $OSR(n_i, n_j)$==NULL then

Sample temp from $NOAI(u_i, u_j)$;

If $temp >= 0.5$, $C_{i,j,agree}^{z_i} += 1$; else $C_{i,j,disagree}^{z_i} += 1$;

End End End End End End

Calculate $\Theta, \Phi, \Psi, \Omega$ *and S* based on Equation 4,5,6 and 7.

Algorithm 1: Estimation of $\Theta, \Phi, \Psi, \Omega$ *and* S

4.2 Inference

Once the four parameters are learned through Algorithm 1, we infer the opinion of user u_i about a query object n_q, i.e. $o_{u_i}^{n_q}$. First, we find the topic t_k that is most probabilistically related to n_j under Φ. Second, we find all neighbors of u_i under topic t_k, i.e. $ON(u_i, t_k)$ and collect all messages between u_i and $u_m \in ON(u_i, t_k)$. Third, for each $u_m \in ON(u_i, t_k)$, we directly detect the opinion of u_m towards *nq* from the messages. Fourth, if opinion $o_{u_i}^{n_q}$ is not neutral, we sample the type of opinion influence (agree or disagree) of *um* on *ui* from Ω. Finally, we obtain the $o_{u_i}^{n_q}$ from a linear combination of the two types of opinions under a random sampling technique: *ui*'s historical opinion about topic t_k sampled from Ψ, and *ui*'s influenced opinion purely deducted from $u_j \in ON(u_i, t_k)$ and the corresponding type of influence. The exact inference process is shown in Algorithm 2. It is worth attention that TOIM may not be able to infer a user's opinion if the provided information is insufficient.

Input: Θ,Φ,Ψ,Ω *and* S, user u_i, object n_q, weight ω

Output: opinion $o_{u_i}^{n_q}$

Initiation: Iterations *Iter*, ω, *SWO* = **0**

For *e=1:Iter* **do**

Find the most probabilistically related topic t_k regarding n_q from ϕ_{kj};

For *user* u_m *in* $ON(u_i,t_k)$ do

If u_m's opinion $o_m^{n_q}$ is known **then**

Set *temp* = $\omega\times\phi_{i,+1}^{k}+(1-\omega)\times\omega_{i,m,agree}^{k}$; sample o_i^{temp} from temp;

If $o_i^{temp}==o_m^{n_q}$ **then**

SWO+ = $o_i^{temp}\, g_{i,m,agree}^{k}$;

Else

SWO+ = $o_i^{temp}\, g_{i,m,disagree}^{k}$;

End

End

End

End

If *SWO*>=0.5 **then**

$o_{u_i}^{n_q}$ +=1;

Else if $o_j^{new}<=-0.5$ **then**

$o_{u_i}^{n_q}-=1$;

Else

$o_{u_i}^{n_q}$ **is unknown;**

End

Algorithm 2: Opinion Inference

5. EXPERIMENT

5.1 Data Description

Similar to Twitter, Tencent Weibo allows users to post messages up to 140 Chinese characters. Users can broadcast (post) new messages, comment, reply or forward existing messages, and follow other users. The social networks we build to infer the topic-level opinion influence are based on comment or reply actions, where user u_i and u_j are connected by an edge if u_i (u_j) comments on or replies to u_j (u_i)'s messages.

We collect messages from Tencent Weibo between 1st, October, 2011 and 5th, January 2012, with an average of 30 to 60 million messages posted daily. We find that most frequent user action is "forwarding", followed by "commenting" and very few records are about "replying". This indicates that large number of Tencent Weibo users prefer to discuss with their friends and express their own ideas by commenting on messages around similar topic, which provides historical records for us to learn the topic-level opinion influence needed by TOIM. To evaluate TOIM, we select five hot entities that were discussed frequently in Tencent Weibo over this three-months period: *O1*: *Muammar Gaddafi*, *O2: The Flowers of War* (*a Chinese movie*), *O3: Chinese Economics*, *O4: School bus accident*, and *O5: College talk from the President of Peking University*. For each entity *Oi*, we search all relevant users who once mentioned *Oi* in their messages during the three-month period. Among all retrieved users, the top 2,000 most active users (the users who create a lot of "comments", "replies" and "mentions") were selected as test samples to infer their opinions about the corresponding query entity *Oi*. The reason for choosing the 2,000 most active users is that we can obtain plenty of textual information about their communication records, which can be utilized by TOIM to better infer their topical opinion.

Given the 2,000 selected users, we further generate four types of data. First, all messages posted by the 2,000 users and related to *Oi* are picked out and denoted as *D1*. We design a bag of keywords, which includes alias, synonym, error-correction terms and other important words, to describe each *Oi*. For example, movie "The Flower of War" has many alias in Chinese, and all alias are collected to identify right weibos; "School bus accident" is a very hot event, which arouse publics' attention towards public safety; we also use a combination of terms to detect weibos, which are related with entity "College talk from the President of Peking University", such as "The name of president" plus "College talk". We manually label all the 2,000 users' opinions towards the five entities based on their weibos; for example, if a user *A* held positive opinions towards "Chinese Economics", then we design the data format as below:

Opinion:+1   Username:A   Entity:Muammar Gaddafi   WeiboContent:"I think Chinese Economic will be better in the next year"   TimeStamp: 2011-1107-12-35

The data records users' message related with "Chinese Economics" and our labeled opinion information "+1" of user *A* towards "Chinese Economics". *D1* is considered as testing data, in our experiment, we want to use the learned model to predict user *A*'s opinion (+1 or -1). If the model is reasonable and well-trained, we will gain a high precision.

Second, all messages posted by the 2,000 users excluding *D1* are denoted as *D2*. The purpose for collecting *D2* is to learn users' opinion preferences according to their own historical records.

Third, the social networks of the 2,000 users are crawled via one-hop commenting and replying actions, where all related commenting and replying messages are extracted and denoted as *D3* (In order to reduce the high dimension of user-user matrix, we set the limitation as top 1,000 ranked followers for each of the 2,000 selected users, the number of communications between selected users and their neighbors is the main indicator to calculate the rank score).

Fourth, all messages of users from the social networks of the 2,000 users are also collected (Not includes the 2,000 users themselves) as *D4*. The learning of those users' behaviors can help to construct stronger influence relationships between current users and the selected 2,000 users.

Particularly, we use *D1* as the testing corpus to manually decide the related users' opinions about *Oi*. We use *D2*, *D3* plus *D4* as training corpus, to learn the topic-level user opinion distribution and topic-level opinion influence, respectively. The data format of training data are summarized as below:

#IDX"The ID of current weibo"
#!"Content of current weibo"
#@"Username of current weibo"
#T"TimeStamp"
#C"IDX"
#END

The data description can be seen in Table 2:

Table 2. Summary of Experimental Data

	# of post messages	# of reply and comment messages	# of users
Total	**2,350,372**	**959,918**	**145,327**
O_1	320,176	114,382	24,382
O_2	591,433	243,876	31,432
O_3	742,853	298,764	38,796
O_4	472,463	275,148	28,254
O_5	295,447	136,748	22,463

5.2 Results

We use two classic measures to evaluate the inference performance of TOIM, precision and recall, but slightly modify them for our case. Here recall is the percentage of users whose opinions can be inferred by TOIM from all users; precision is the percentage of correct opinion inferences by TOIM of all users whose opinions are detectable. In other words, recall indicates the capability of TOIM to detect user opinion given data sparsity, while precision indicates the capability of TOIM to correctly infer user opinion.

For comparison, three algorithms, Support Vector Machine (SVM), Conditional Random Field (CRF), and Joint Sentiment Topic (JST), are severed as the baseline methods. As common classical algorithms, SVM and CRF have been widely applied in sentiment analysis. JST is a probabilistic graphical model used to estimate user opinion preferences on various topics. Different from TOIM, all three algorithms do not take opinion influence into consideration. We adopt SVM-light[4] for SVM, adopt the code from Tang et al. (2009) for CRF and implement the work of Lin et al. (2009) for JST. For those baseline models, the inference of user opinions are mainly based on their historical opinions of their posted messages. We extract user name, key words, topics, entities and users' sentiment towards each entity and other information from each message as attributes, then apply SVM, CRF and

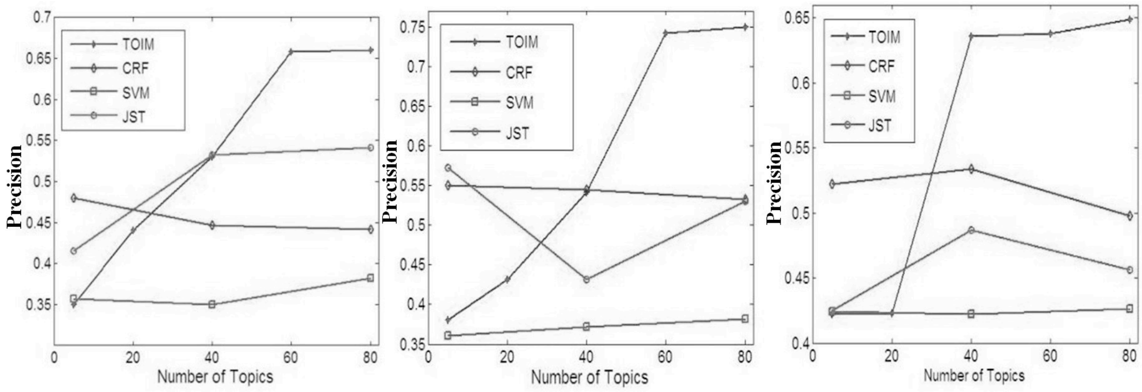
[4] http://svmlight.joachims.org/

JST to make inference. To generate the training and testing dataset, four attributes of each user are defined as: 1) username, 2) nouns with their weighted score, 3) qualifiers of the nouns and 4) topics ID related with those nouns. For example, a user *Xt* writes a micro-blog *mb*, then the input format should be: {Opinion (+1, 0, -1); Username:Weight0; Noun1:Weight1; Qualifier1:Qweight1; Noun2:Weight2; Qualifier2:Qweight2;…Topic ID:WeightZ}. Nouni (*i=1,2…*) is a word, which is mentioned in *mb*. KeyWord Extraction technology is used to compute the weight of each noun and their qualifiers according to their grammar position in *mb*: 1). The noun with highest score should be the most important core words. 2) User's attitude toward this noun could be considered as the input of "Opinion". 3) "Weight0" of "Username" is assigned as 1, while "WeightZ" of "Topic ID" is the score of "Topic ID" on *mb*. All the inferred results from each model are normalized from -1 to 1. The normalization methods are summarized as below:

a) The result value range of SVM is from -1 to +1, so there is no need to normalize the results of SVM.
b) For the results of TOIM, JST and CRF, because all results are probabilities, the value range of which is from 0 to +1. In order to map the value of {0, 1} to {-1, +1}, we design the formula as below:

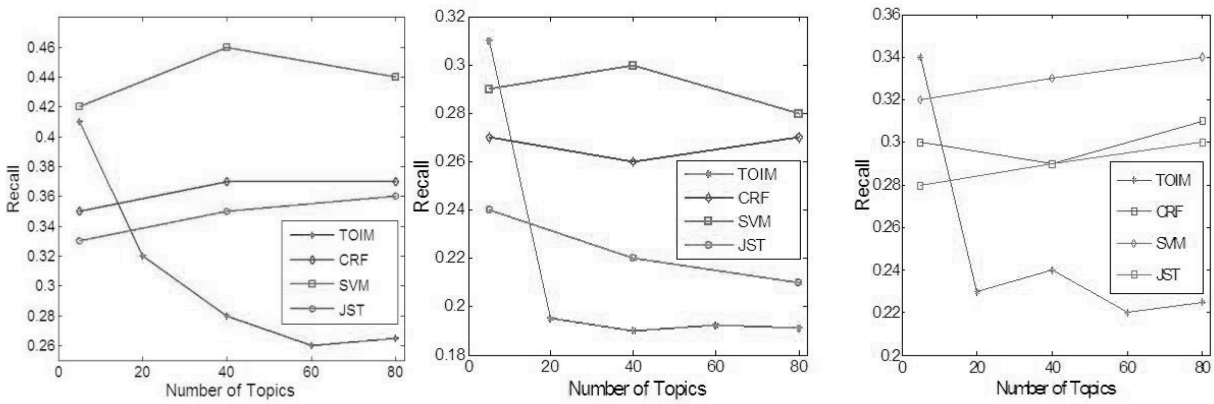
$$f(\mathrm{S},\mathrm{X},\mathrm{T},\mathrm{W}) = S \times P(\mathrm{S} \mid \mathrm{X},\mathrm{T},\mathrm{W}) \qquad \text{s.t. } \mathrm{S} \in \{-1, 0, +1\} \tag{10}$$

where *S* means sentimental polarities, +1 means positive sentiment, -1 means negative sentiment, while 0 means no sentiment preferences. The results can be seen in Figure 4 and Figure 5.



(a) O_1 (b) O_2 (c) O_3

Figure 4: The precision of opinion inference for (a) *O1*, (b) *O2*, and (c) *O3*, using four algorithms: SVM, CRF, JST and TOIM. As the number of topics increases, the precision of TOIM increases as well, since the topic-level opinion influence becomes more and more precise. By contrast, the precision of other three algorithms is not affected by topic numbers.



(a) O_1 (b) O_2 (c) O_3

Figure 5: The recall of opinion inference for (a) *O1*, (b) *O2*, and (c) *O3*, using four algorithms: SVM, CRF, JST and TOIM. As

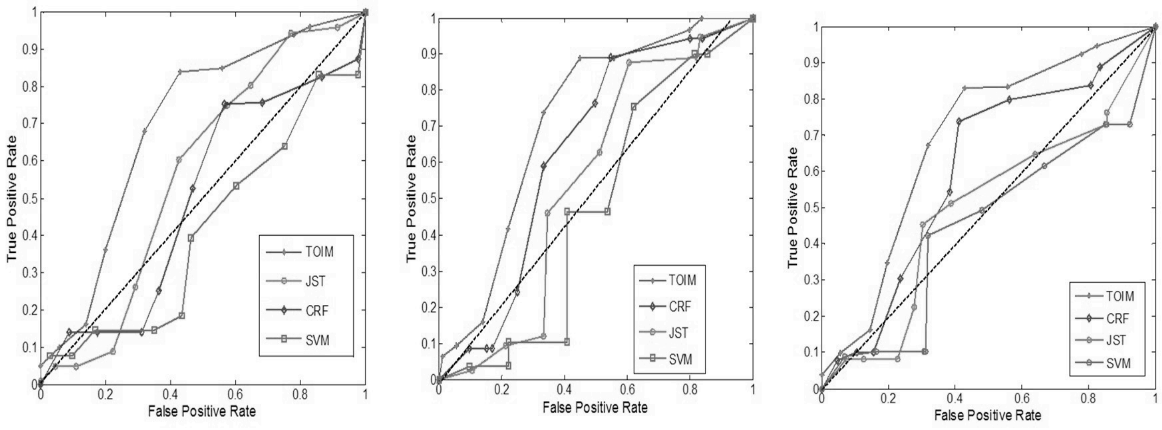
the number of topics increases, the recall of TOIM decreases as well, since the likelihood of finding two users' opinions on the same topic decreases. By contrast, the recall of other three algorithms is not affected by topic numbers.

Figure 4 shows the precision of opinion inference for entities *O1*, *O2*, *O3* using four algorithms including TOIM and three baseline methods, against different number of topics. We can see that when the number of topics (K) is small, the precision of TOIM is almost as low as SVM, and not as good as CRF and JST. However, as K increases, the precision of TOIM dramatically increases and surpasses the other three baselines methods, until reaches a plateau. By contrast, the other three methods do not show such trend with K. The reason is that the topic-level opinion influence is very sensitive to the number of topics. When K is small, many topics are mixed together, which makes the topic-level opinion influence vague and imprecise. For instance, topic “politics” and “military” may be mixed together when discussing opinion influence about “Libya”. However, as K increases, the distinct topics are separated and the opinion influence becomes more precise. By contrast, the other three algorithms, which do not consider the opinion influence, are obviously not affected by K, because we did not pre-define logical relationships between topics and other attributes as we did in TOIM.

Figure 5 shows the recall of opinion inference using four algorithms against different K values, based on entities O1, O2, O3. Again, the recall of TOIM is also sensitive to K and decreases as K increases. As mentioned, recall indicates the capability to detect user's opinion (regardless correct or incorrect) given the sparseness of the data. Therefore, when K is small, we have high probability to track the opinions of two connected users at the same topic and (when $K = 5$) easy to identify the opinion influence at the topic level. By contrast, when K is large, the chance that two connected users express their opinions at the same topic (when $K = 100$) is very small thus TOIM fails to detect the opinion influence at the topic level, leading to low recall. Similar to the precision, the other three algorithms without considering the opinion influence are not obviously affected by K.

The overall inference performance of TOIM exhibits low recall but higher precision values compared with other three baseline methods. Since text messages on microblogging are generally noisy and short, many messages do not show obvious sentimental polarization, and the communication records between users are sometimes too sparse to judge their opinion influence type. If the main goal of opinion inference is to identify a set of opinioned users from a large number of opinion-detectable users, then TOIM is not a good choice, since too many topic partitions will sparsify the data. However, if the main goal is to correctly classify the opinions of a small set of users, who are very active and frequently communicate with their neighbors, then TOIM is a good choice because topic-level opinion influence can improve the inference precision.

In order to better illustrate the performance of TOIM, we assign the threshold as {0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1}, and drew ROC(Receiver Operating Characteristic) for object O1, O2, O3, the figure can be seen as below:

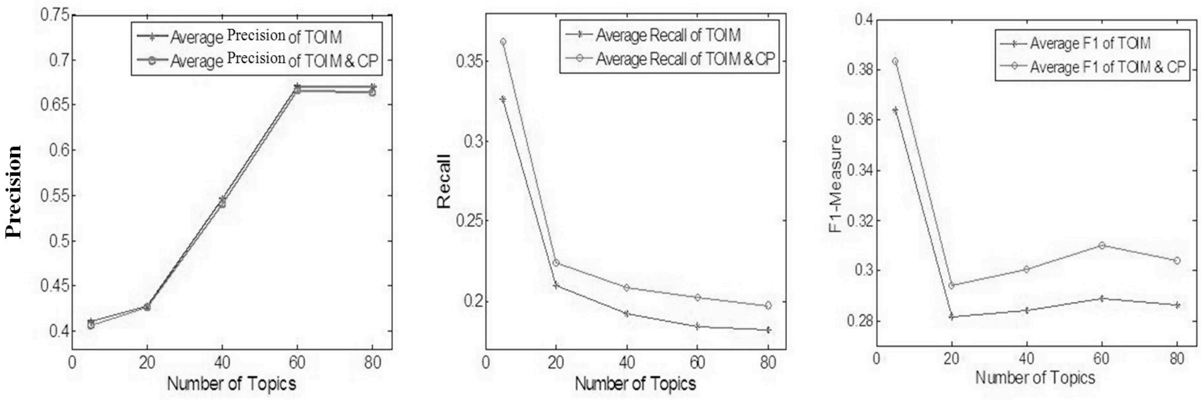


(a) *O1* (b) *O2* (c) *O3*

Figure 6: ROC of opinion inference for (a) *O1*, (b) *O2*, and (c) *O3*, using four algorithms: SVM, CRF, JST and TOIM. For all results, AUC of TOIM is biggest than other baseline.

In Figure 6, we plot ROC curves to further evaluate the performance of our proposed model. ROC curve is used to observe the performance of classifiers under different conditions. Particularly, AUC is the area under ROC curve, whose value is an important index to evaluate the performance of a certain classifier. We can see that the AUC of TOIM is significantly bigger than other methods, and its ROC curves are all above the diagonal line for three selected entities: *O1, O2* and *O3,* implying that TOIM is a better method to judge a user's sentiment towards a certain topic than other baseline methods.

According to Liu's studies (2010), incorporating indirect influence can further improve the performance of influence model. Indirect influence represents the influence relations between two users who do not have direct connections. For examples, user *A* and *B* do not have communication records before, the influence relationships between *A* and *B* may also be calculated from the third part, for example, if *A* has communication records with user *C*, and *C* has communication records with *B*, then we may use indirect influence to calculate the influence between *A* and *B*. One main benefit of incorporating indirect influence is that it can improve "recall", because more influential relationships can be detected for inferring users' potential opinion preferences. Based on Liu's work, we mainly used Conservative Propagation (*CP*) to calculate indirect influence. The performance of combining indirect influence can be seen in Figure 7:



(a) Average Precision (b) Average Recall (c) Average F1-Measure

Figure 7: Average Precision, Recall and F1-Measure combining TOIM with Indirect Influence. We use Conservative Propagation(CP) to calculate indirect influence. For each entity *Oi*, we calculate Precision *PA(Oi)*, Recall *RA(Oi)* and

F1-Measure *FA(Oi)* for TOIM, and Precision *PB(Oi)*, Recall *RB(Oi)* and F1-Measure *FB(Oi)* for TOIM with Indirect Influence. Then we calculate the average Precision, Recall and F1-measure for both TOIM and TOIM with Indirect Influence.

Figure 7 shows that incorporating indirect influence in TOIM not only can improve average precision, but also can improve average recall and f1-measure significantly. The reason is that indirect influence helps to build influence connections between two users who are not directly connected, thus provide more information regarding the users' opinions inferences.

6. CASE STUDY

As is mentioned above, the main feature of TOIM is to infer user's opinion by simultaneously considering topic and opinion influence. Consequently, we list several applications of TOIM in opinion mining around two aspects: opinion influence mining and topic-level opinion mining.

6.1 Opinion Influence Mining

6.1.1 Influence vs. Non-influence Opinion Mining

In Section 5.2, we show that TOIM outperforms other three baseline methods in opinion inference precision because TOIM considers the opinion influence deducted from the social networks. Here we select five representative users and infer their opinions towards the entity "Muammar Gaddafi" using both TOIM and CRF. Table 3 compares the opinion inference using TOIM and CRFs. For the five selected users, it is difficult to detect their opinions towards "Muammar Gaddafi" by only analyzing their personal messages using CRFs. By contrast, TOIM can leverage the opinions of their neighbors to find their opinion influence relationships and make better opinion inference.

Table 3: The comparison between TOIM and CRF in opinion inference for five selected users.

User (True Opinion)	Method	No. of Opinion Neighbors		Inferred Results
		Agree (Average Probability)	Disagree (Average Probability)	
Xgdd (Positive)	TOIM	5 neighbors (0.7236)	3 neighbors (0.3253)	Correct
	CRF	N/A	N/A	Incorrect
LyhLawer (Negative)	TOIM	6 neighbors (0.1324)	NULL	Correct
	CRF	N/A	N/A	Incorrect
Zhang (Positive)	TOIM	2 neighbors (0.0432)	4 neighbors (0.5853)	Correct
	CRF	N/A	N/A	Incorrect
HuChunhua (Negative)	TOIM	2 neighbors (0.0012)	2 neighbors (0.6872)	Correct
	CRF	N/A	N/A	Incorrect

Buffaloes (Positive)	TOIM	1 neighbor (1.0000)	3 neighbors (0.2346)	Correct
	CRF	N/A	N/A	Incorrect

As an example, we have detected the opinions of user Xgdd's eight highly influential neighbors and their opinion relationships with Xgdd. Specifically, five of them agree with Xgdd's opinion with an average probability of 0.7236 to have positive opinion towards "Muammar Gaddafi" while the rest three disagree with an average probability of 0.3253 to have positive opinion. So it is easy to infer that Xgdd may have a positive opinion, which is consistent with his real opinion ("Positive" after his user id in the first column). All of these can be integrated using Algorithm 2 to infer Xgdd's opinion towards "Muammar Gaddafi". The five examples indicate that, given sufficient information about users' neighbors' opinion and their corresponding opinion influence types and strengths, TOIM can detect their topic-level opinions more accurately than other baseline methods such as CRF.

6.1.2 Opinion Leader Recognition

The "opinion leaders" refer to those users whose opinions on some topic are largely supported by their followers. We select the five most popular topics: *College & Education* (A), *Daily Emotion*(B), *Chinese Economics*(C), *Economics & Tech*(D), and *International Political*(E). For each topic, we use a chain counter *CUTOxzw* to calculate the influence score of each user, which are the normalized values of *CUTOxzw* defined for each user u_x. *CUTOxzw* is consistently updated through the TOIM learning process: If we find another user who has the same opinion o_ω on topic t_z with user u_x, then we increase *CUTOxzw* by one. We select top three users ranked by influential scores and show them in Figure 8.

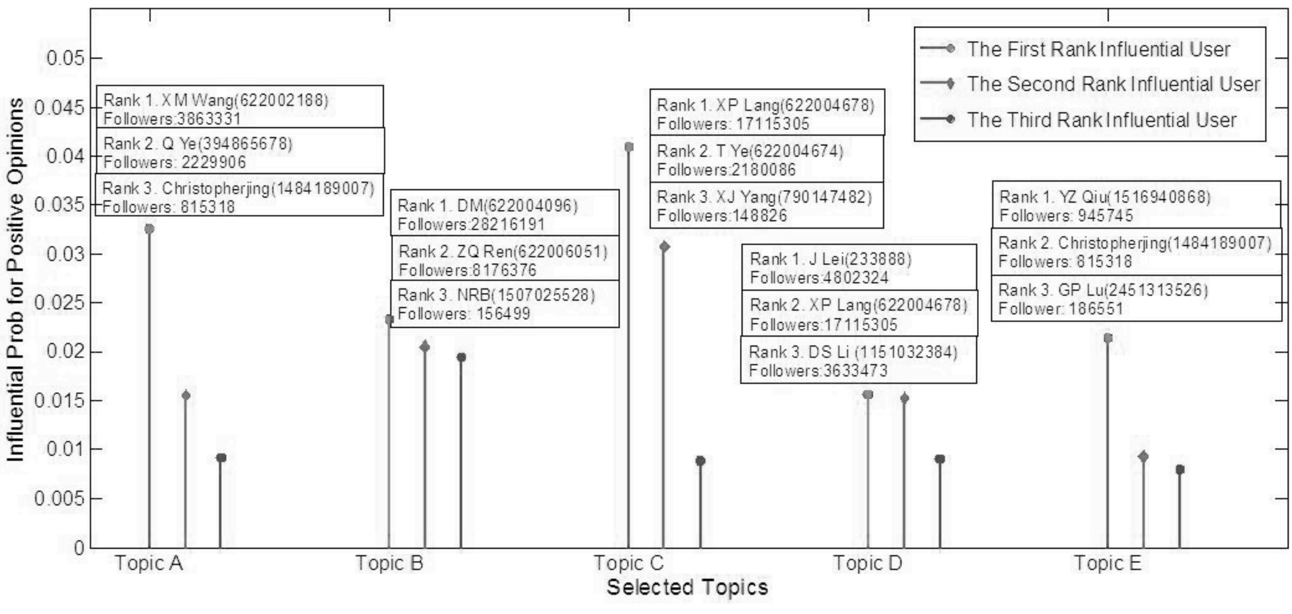


Figure 8: The top three most influential opinion leaders in five topics: Topic A-College & Education, Topic B-Daily Emotion, Topic C-Chinese Economics, Topic D-Economics & Tech, Topic E- International Political. The users are represented by stem lines with the length proportional to the influential degree.

We check the profiles of those opinion leaders and find their backgrounds were consistent with the corresponding topics. For instance, user *XM Wang* (622002188) is a famous expert in the education domain; user *XP Lang* (622004678) is a top notch economist; user *YZ Qiu* (1516940868) is a famous international journalist and professional in the international political fields. Then we further check their activities during the three-month period and find that they are all very active that their messages are forwarded and commented on by thousands of users. For instance, *YZ Qiu* often went to conflicting areas (i.e. Libya) to interview local residents, and published the latest news reports. He is also very popular among Tencent users and often participated in online communications with his neighbors. Opinion leaders may not have a huge number of followers (exceed 10 million), but are actually influential in a certain domain, where their messages are closely watched and commented by thousands of users.

6.1.3 Opinion Influence Network

The opinion influence network demonstrates the type and strength of opinion influence from a central node to all other nodes in the social network. Figure 8 shows the positive opinion influence network of *XP Lang*(622004678), a famous economist, who is located at the center. The lines from the center node to others represent positive opinion influence from *XP Lang* on those users on O3 (Chinese Economics), and the values on every line represent the influential degree. For example, the influential degree for node 815902964 is 0.53, indicating that when *XP Lang* expresses a positive opinion, the node will have a probability of 0.53 to agree with him. In addition, the values in the right box denote the frequencies of agreement with *XP Lang*. For example, UID-1516940868: 32(Times of Agree) means that, user 1516940868 agrees with *XP Lang*'s opinion on Chinese Economy for 32 times. Those values can be considered as a confidence for influential degree. If two users have agreed for many times, the influence inference will be more accurate.

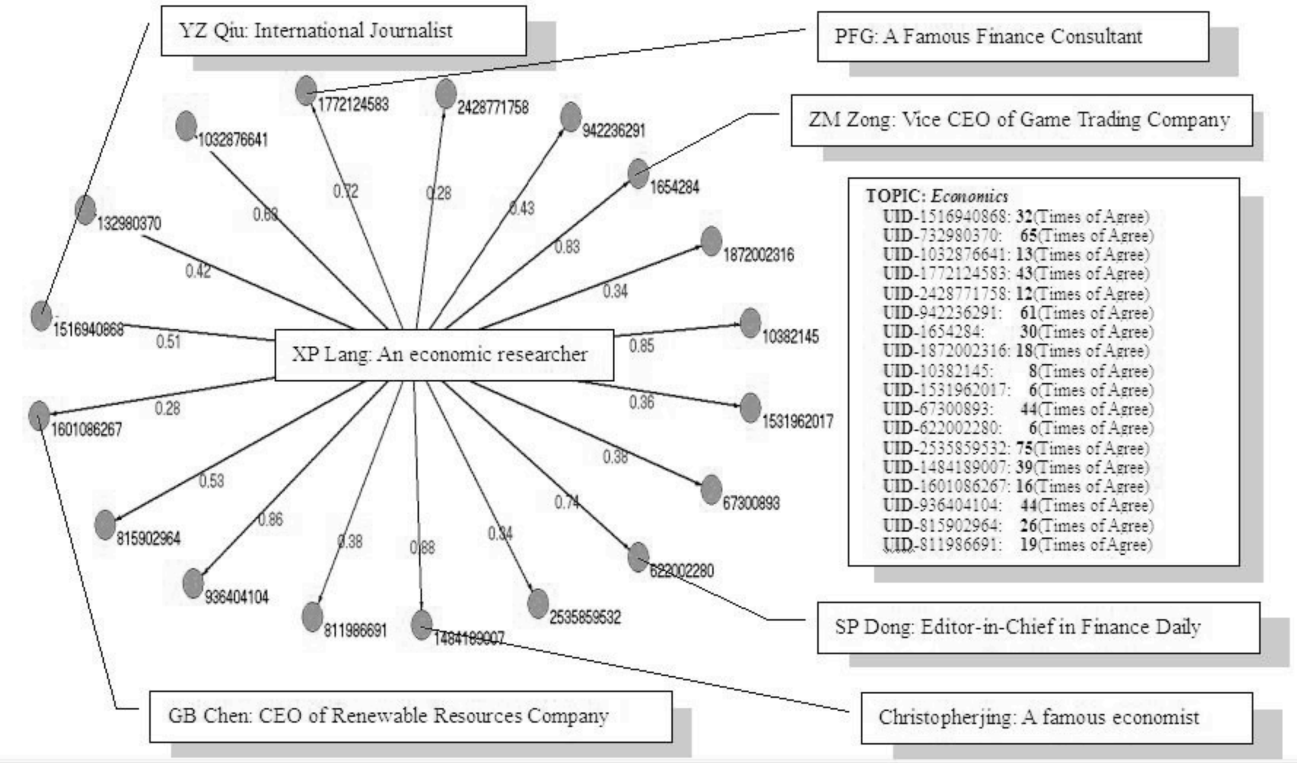


Figure 9: The direct opinion influence map of *XP Lang* on Chinese Economy. The central node is *XP Lang* and the other nodes are opinion neighbors, who frequently communicate with *XP Lang* and show agreement with him. Especially, some marked nodes with identification information are verified celebrities.

Especially, the opinion influence network can detect how influential users communicate with each other, which might provide more commercially valuable information than opinion communications involving ordinary users. The nodes marked with identification information are manually detected celebrities in the network. We find that those celebrities are interested in commenting on *XP Lang*'s opinion on Chinese Economy, and have high probabilities of agreement with his opinions.

6.2 Topic-Level Opinion Mining

6.2.1 Topic-Opinion Preference Detection for Celebrities

We select four user accounts already verified by Tencent Weibo, and detect their topical opinion preferences. The top three most discussed topics and their corresponding positive opinion probabilities, as well as some representative sentimental words under each topic are listed in Table 4.

Although the four users' average number of messages is around 50, they are still detected as the most influential persons in their topics, because according to TOIM, if a user's opinions are popular and discussed by many others, he/she will gain a high score of Influence. The third column lists the frequently discussed topics of each user, where the words in brackets are the expression of that topic, and the value behind each topic is the weight of users' opinion preferences. For example, for Topic 1, user *L Yang* (622007070) will have a probability of 0.7101 to pick a positive opinion, and 0.2899 to pick a negative opinion. The words in the sixth column behind each topic give the most representative word, which has been often used to describe the current user's opinions on certain topic. For example, University means that user *L Yang* (622007070) has ever held positive opinions toward that word on Topic 1. In order to detect the representative words for topics under different opinions, we define a chain counter *CUTOWxyzw* for each user u_x: $u_x \rightarrow t_y \rightarrow n_z \rightarrow o_w$. When user u_x expresses opinion o_w towards entity n_z under topic t_y, we increase *CUTOWxyzw* by one, during the learning process. Then we can use *CUTOWxyzw* to calculate the weight of each entity towards user u_x under certain topic t_y.

Table 4. User's opinion preference on different topics represented by different words. The superscript symbol "+" denotes positive opinion or agree, while "-" denotes negative opinion or disagree.

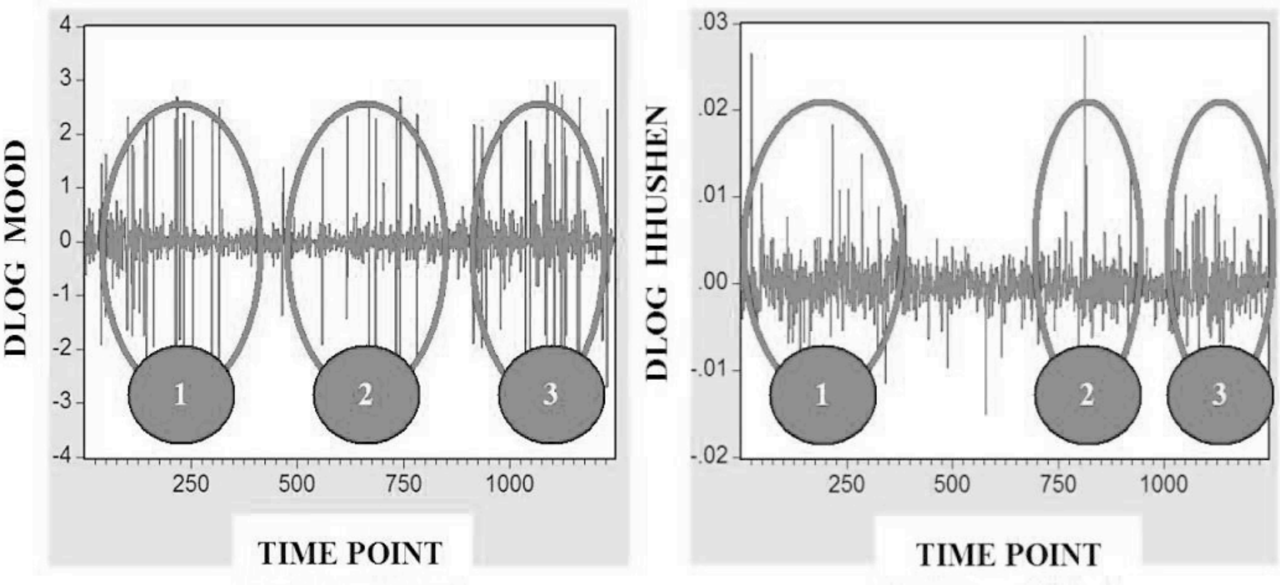
User ID	Profile	Topic Preference	Opinion Probability	Influential Score (max)	Representative words
L Yang (622007070)	Famous Host **# of followers:** 10,116,317 **# of messages:** 28 **# of comments:** 30,700	**Topic 1:** College & Education **Topic 2:** Daily Emotion **Topic 3:** Country & History	0.7101+ 0.6845+ 0.4729+	0.0063+(0.034) 0.0154+(0.0235) 0.0132+(0.0274)	**POS:** daughter+, exam+, university+ **NEG:** primary school-, student- **POS:** happiness+, expectation+, women+ **NEG:** encounter-, home-, violence- **POS:** women+, china+, army+ **NEG:** war-, power-, hurt-
Y Qin (394865678)	Famous Financial Officer **# of followers:** 2,186,021 **# of messages:** 36 **# of comments:** 6,242	**Topic 1:** College & Education **Topic 3:** Social Affairs **Topic 23:** Country Development	0.7907+ 0.5463+ 0.4729+	0.0154+(0.034) 0.0102+(0.0274) 0.0052+(0.0193)	**POS:** research+, university+, student+ **NEG:** college- **POS:** protection+, children+, driver+ **NEG:** accident-, corruption- **POS:** innovation+, investment+ **NEG:** enterprise-, lawsuit-, stock-
Christopherjing (1484189007)	Famous Researcher **# of followers:** 800,770 **# of messages:** 78 **# of comments:**	**Topic 1:** College & Education **Topic 13:** Economics **Topic 27:**	0.5777+ 0.3227+ 0.3939+	0.0087+(0.034) 0.0064+(0.0422) 0.0042+(0.0223)	**POS:** university+, student+, china+ **NEG:** education-, research-, lost- **POS:** innovation+, medical+, economics+ **NEG:** cost-, technology-, company- **POS:** charm+, people+, solution+

	31,644	International Politics			**NEG:** relationship-, criticism-, protest-
XJ Yang (790147482)	Vice Editor in Finance **# of followers:**	**Topic 13:** Economics	0.2988+	0.0076+(0.0422)	**POS:** reform+, tax+, finance+ **NEG:** industry-, monopoly-, debit-
	146,718 **# of messages:** 47	**Topic 14:** Social Study	0.3267+	0.0062+(0.0186)	**POS:** democracy+, institution+, law+ **NEG:** officers-, society-, market-
	# of comments: 14,140	**Topic 23:** Country Development	0.5232+	0.0058+(0.0193)	**POS:** charm+, people+, solution+ **NEG:** income-, ination-, walfare-

The opinions of celebrities towards their preferential topics can also be identified by TOIM. User *L Yang* (622007070) and *Q Ye* (394865678) hold relatively moderate attitudes towards Topic 1 and near-neutral opinion about other topics. By contrast, user *Christopherjing* (1484189007) and *XJ Yang*(790147482) seem more aggressive and often express negative opinions towards Topics 13, 14 and 27. As seen in the fifth column in Table 4, all celebrities exert certain degrees of positive influences on their neighbors in their interested topics, “influential score” is to evaluate the influence of each celebrity in each topic, “max” means the maximum influential score, which is from the most influential users, in corresponding topic. We find that though the selected celebrities are well-known in public, they may not be the most influential users on Tencent Weibo, because some other users' messages within the same topic seem to be more welcomed and discussed, though they are less well-known person in reality. Although the celebrities may not be the most influential users in their interested topics, their influence ranks are all in top 20, implying that their opinions are still generally supported by many Tencent followers.

6.2.2 Opinion and Real-World Correlation Identification

The correlation between public opinion (mood) and real world events has been confirmed by several previous studies. Here, we examine the correlation between Tencent opinion and Chinese economic index: Hushen-300 Index 5. Figure 10(a) shows the volatility of Tencent moods about *O3* (i.e. Chinese Economy), calculated using TOIM during the three-month period; Figure 10(b) shows the volatility of Hushen-300 trend for the same time periods. Three areas are marked in both figures, and the corresponding two areas marked with the same number exhibit negative correlations but with a time delay (for instance, area 1 in Figure 10(a) vs. area 1 in Figure 10(b)). The statistical analysis shows that the R square of such correlation is around 0.3. If we use a simple sentiment analysis strategy (only counting sentiment words without considering the topic and influence factor) to calculate the correlation, then the R square drops to 0.15. This indicates that TOIM can better capture the correlation between Tencent opinion and the trend of real world economy, than other simple opinion mining method.



(a) Tencent Moods (b) Hushen 300 Index

Figure 10: Intuitive Correlation between Hushen 300 Index and Tencent Mood. Hushen 300 is an important financial index to reflect real economics in China. Y axes represents opinion polarities in (a) and Hushen index values in (b), and the X axis represents the time. (a) Tencent Mood and (b) Hushen 300 share similar trend clustering in the response three areas (1,2,3) along the time (there exists a time delay between two time series), which means that TOIM model can also be used to detect the investors' sentiment in the future.

7. CONCLUSIONS AND FUTURE WORK

This paper investigates the problem of inferring user opinion by identifying opinion influence in social networks over different topics on microblogging. A Topic-level Opinion Influence Model (TOIM) is proposed and tested on Tencent Weibo, a famous microblogging website in China. Users' historical messages and social interaction records are leveraged by TOIM to construct their historical opinions and neighbors' opinion influences through a statistical learning process, which can be further utilized to infer users' future opinions towards specific topics. To test and evaluate the proposed model, an experiment is conducted based on three-month data from Tencent Weibo. The results show that the proposed TOIM effectively combines social influence and topic preference simultaneously, and outperform baseline methods in opinion inference accuracy, but has a relatively low recall, mainly due to the sparse data problem.

Therefore, the suggested use of TOIM is to detect behavior patterns within small sets of active users, who communicate with each other very frequently. Especially, we demonstrate that the inferred opinion from TOIM can be applied to detect celebrities' opinions towards various topics, identify the collective opinion correlations with real-world phenomena, visualize opinion influence structure and recognize opinion leaders in different domains.

There are several limitations of TOIM, which need further study. First, the user's true opinions can be misunderstood. The reason is that human language expression can be complicated in terms of mocks, analogies, and implications. Such ambiguity can be further aggravated in the microblogging environment, where users tend to create short, informal, and vague messages. Additionally, the opinion detection component of our model is still primitive, with potential challenges including how to design and obtain dedicated topic labels, and how to effectively pre-process experiment datasets (e.g. delete noisy information, use training data to denote constraint rules for learning algorithms).

8. ACKNOWLEDGEMENTS

This paper is supported by China Post Doc Funding(2012M510027, 023250015). National Basic Research Program of China(No.2011CB302302). He Gaoji Project, Tencent Company(No.2011ZX-01042-001-002). The National Natural Science Foundation of China (NSFC Program No.71072037).

9. REFERENCES

Anagnostopoulos A, Kumar R, Mahdian M (2008) Influence and correlation in social networks. KDD '08, pp. 7-15. doi:10.1145/1401890.1401897. URL http://doi.acm. org/10.1145/1401890.1401897.

Asmussen, S and Glynn, P. W. Stochastic Simulation: Algorithms and Analysis. Springer. Series: Stochastic Modeling and Applied Probability, Vol. 57, 2007.Blei DM, Ng AY, Jordan MI (2003) Latent dirichlet allocation. Journal of Machine Learning Research 3: 993-1022.

Blei, D., Lafferty, J, D. Dynamic Topic Models. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA, 2006.

Bloom K, Garg N, Argamon S (2007) Extracting appraisal expressions. In: Sidner CL, Schultz T, Stone M, Zhai C, editors, HLT-NAACL. The Association for Computational Linguistics, pp. 308-315. 17.

Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Ukkonen, A., & Weber, I.Web search queries can predict stock market volumes. *PLOS ONE*, 7. 2012.

Bollen J, Mao H, Zheng X (2011) Twitter mood predicts the stock market. Journal of Computational Science 2: 1-8.

Crandall D, Cosley D, Huttenlocher D, Kleinberg J, Suri S (2008) Feedback effects between similarity and social influence in online communities. KDD '08, pp. 160-168.

Domingos, P and Richardson, M. Mining the network value of customers. In KDD'01, pages 57–66, 2001.

Gerald P, Michal K, Harald F, Andreas, A, Vaclav S and Andreas H. Opinion Mining on the Web 2.0-Characteristics of User Generated Content and Their Impacts. HCI-KDD 2013, LNCS 7947, pp.35-46, 2013.

Gilks, W.R., Richardson, S., Spiegelhalter, D. Markov Chain Monte Carlo in Practice. Chapman & Hall/CRC Interdisciplinary Statistics. 1995.

Griffths TL, Steyvers M (2004) Finding scientific topics. Proceedings of the National Academy of Sciences 101: 5228-5235.

Gruhl D, Guha R, Kumar R, Novak J, Tomkins A (2005) The predictive power of online chatter. In: Proceedings of the eleventh ACM SIGKDD international conference on Knowledge discovery in data mining. New York, NY, USA: ACM, KDD '05, pp. 78-87. doi:10.1145/1081870.1081883. URL http://doi.acm.org/10.1145/1081870.1081883.

Guerra PHC, Veloso A, Jr WM, Almeida V (2011) From bias to opinion: a transfer-learning approach to real-time sentiment analysis. KDD '11, pp. 150-158.

Hu M, Liu B (2004) Mining opinion features in customer reviews. AAAI'04, pp. 755-760.

Kamps J, Mokken RJ, Marx M, de Rijke M (2004) Using WordNet to measure semantic orientation of adjectives. In: Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004). Paris, France: European Language Resources Association, volume IV, pp. 1115-1118.

Kim SM, Hovy E (2004) Determining the sentiment of opinions. COLING '04. URL http://dx.doi.org/ 10.3115/1220355.1220555.

Kempe D, Kleinberg J, Tardos E (2003) Maximizing the spread of influence through a social network. KDD '03, pp. 137-146.

Leon D, Diana M, Niraj A and Kalina B. Microblog-Genre Noise and Impact on Semantic Annotation Accuracy. In proceedings of 24th ACM Conference on Hypertest and Social Media. 1-3 May 2013, Pairs, France.

Li D, Shuai X, Sun G, Tang J, Ding Y, et al. (2012) Mining topic-level opinion influence in microblog. In: Proceedings of the 21st ACM international conference on Information 15 and knowledge management. New York, NY, USA: ACM, CIKM '12, pp. 1562-1566. doi:10.1145/2396761.2398473. URL http://doi.acm.org/10.1145/2396761.2398473.

Lin C, He Y (2009) Joint sentiment/topic model for sentiment analysis. CIKM '09, pp.375-384. doi:10.1145/1645953.1646003. URL http://doi.acm.org/10.1145/1645953. 1646003.

Liu H, Zhao Y, Qin B, Liu T (2010) Comment target extraction and sentiment classification. Journal of Chinese Information Processing 24: 84-88.

Liu L, Tang J, Han J, Jiang M, Yang S (2010) Mining topic-level influence in heterogeneous networks. CIKM 2010, pp. 199-208.

Liu Y, Huang X, An A, Yu X (2007) Arsa: a sentiment-aware model for predicting sales performance using blogs. In: Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval. New York, NY, USA: ACM, SIGIR '07, pp. 607-614. doi:10.1145/1277741.1277845. URL http://doi.acm. org/10.1145/1277741.1277845.

Mao, Y., Wang, B., Wei, W., & Liu, B. Correlating S&P 500 stocks with Twitter data. HotSocial'12, August, 12, Beijing, China. 2012.

Mei Q, Ling X, Wondra M, Su H, Zhai C (2007) Topic sentiment mixture: modeling facets and opinions in weblogs. In: Proceedings of the 16th international conference on World Wide Web. New York, NY, USA: ACM, WWW '07, pp. 171-180. doi:10.1145/1242572. 1242596. URL http://doi.acm.org/10.1145/1242572.1242596.

Mihalcea R, Banea C, Wiebe J (2007) Learning multilingual subjective language via cross-lingual projections. In: Proceedings of the 45th Annual Meeting of the Association of Computational Linguistics.

Mishne G, Glance N (2006) Predicting movie sales from blogger sentiment. In: AAAI Symposium on Computational Approaches to Analysing Weblogs (AAAI-CAAW). pp.155-158.

Moghaddam S and Ester M. Opinion Mining in Online Reviews:Recent Trends. Tutorial at WWW2013.

Nallapati R, Cohen W (2008) Link-plsa-lda : A new unsupervised model for topics and influence of blogs. Artificial Intelligence : 84-92.16.

Pang B, Lee L, Vaithyanathan S (2002) Thumbs up?: sentiment classification using machine learning techniques. EMNLP '02, pp. 79-86. doi:10.3115/1118693.1118704. URL: http://dx.doi.org/10.3115/1118693.1118704.

Qualities of a Leader - Online Leadership Guide - Personal MBTI Type Analysis. qualities-of-a-leader.com. December 26, 2011. Retrieved 8 April 2013.

Ramage D, Hall D, Nallapati R, Manning CD (2009) Labeled lda: a supervised topic model for credit attribution in multi-labeled corpora. EMNLP '09, pp. 248-256. URL http://dl.acm.org/citation.cfm?id=1699510.1699543.

Richardson, M and Domingos, P. Mining knowledge-sharing sites for viral marketing. In KDD'02, pages 61–70, 2002.

Riloff E, Wiebe J (2003) Learning extraction patterns for subjective expressions. EMNLP'03, pp. 105-112. doi:10.3115/1119355.1119369. URL http://dx.doi.org/10.3115/ 1119355.1119369.

Rosen-Zvi M, Gri_ths T, Steyvers M, Smyth P (2004) The author-topic model for authors and documents. UAI '04, pp. 487-494. URL http://dl.acm.org/citation.cfm?id=1036843.1036902.

Song X, Lin CY, Tseng BL, Sun MT (2005) Modeling and predicting personal information dissemination behavior. In: Grossman R, Bayardo R, Bennett KP, editors, KDD. ACM, pp. 479-488.

Su F, Markert K (2009) Subjectivity recognition on word senses via semi-supervised min- cuts. In: Proceedings of Human Language Technologies: The 2009 Annual Conference of the North American Chapter of the Association for Computational Linguistics. Stroudsburg, PA, USA: Association for Computational Linguistics,

NAACL '09, pp. 1-9.
Takamura H, Inui T, Okumura M (2005) Extracting semantic orientation of words using spin model. In: Proceedings of the Association for Computational Linguistics (ACL). pp. 133-140.
Tan C, Lee L, Tang J, Jiang L, Zhou M, et al. (2011) User-level sentiment analysis incorporating social networks. KDD '11, pp. 1397-1405. doi:10.1145/2020408.2020614. URL http://doi.acm.org/10.1145/2020408.2020614.
Tang J, Sun J, Wang C, Yang Z (2009) Social influence analysis in large-scale networks. KDD '09, pp. 807-816. URL http://doi.acm.org/10.1145/1557019.1557108.
Walsh, B. Markov Chain Monte Carlo and Gibbs Sampling. Lecture Notes for EEB 581, version 26. 2004.
Wiebe J, Mihalcea R (2006) Word sense and subjectivity. In: In: Proc. ACL-06. pp.1065-1072.
Vaileios L, Daniel P and Trevor C. A user-centric model of voting intention from Social Media. In Proceedings of the 51th Annual Meeting of the Association for Computational Linguistics, pages 993-1003, Sofia, Bulgaria, August 4-9. 2013.
Zhai Z, Liu B, Xu H, Jia P (2011) Constrained lda for grouping product features in opinion mining. PAKDD'11, pp. 448-459. URL http://dl.acm.org/citation.cfm?id=2017863.2017907.
Zhang, X., Fuehres, H., & Gloor, P. Predicting stock market indicators through Twitter: “I hope it is not as bad as I fear.” Procedia-Social and Behavioral Sciences. 2010.