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

Daifeng Li and Jie Tang*
Department of Computer Science and Technology, 1-308, FIT Building, Tsinghua University, Beijing, 100084, China, PR. E-mail: daifli_3000@163.com, jery.tang@gmail.com

Ying Ding, Xin Shuai, and Tamy Chambers
School of Informatics and Computing, Indiana University, 1320 East 10th Street, LI 025, Bloomington, IN 47405, USA. E-mail: {dingying, tischt, xshuai}@indiana.edu

Guozheng Sun
Tencent Company, No. 66, North Sihuan West Road, Haidian District, Beijing, P.R. China. E-mail: gordon.gzsun@gmail.com

Zhipeng Luo
Beijing University of Aeronautics and Astronautics, No. 37, Xueyuan Road, Haidian District, Beijing, P.R. China. E-mail: patrick.luo2009@gmail.com

Jingwei Zhang
Department of Electronic Engineering, Rohm Building, Tsinghua University, Beijing, China, 100084. E-mail: iceboal@gmail.com

Text mining has been widely used in multiple types of user-generated data to infer user opinion, but its application to microblogging is difficult because text messages are short and noisy, providing limited information about user opinion. Given that microblogging users communicate with 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.

Introduction

Online social media, including microblogging, contain rich information about user opinion. This information is valuable to public services and initiatives such as stock prediction (Bollen, Mao, & Zheng, 2011), marketing (Domingos & Richardson, 2001), political campaigns and sports (Guerra, Veloso, Meira, & Almeida, 2011), and even traffic prediction (Yu, Song, Fu, & Song, 2013). For example, by understanding users’ real-time sentiment toward new events related to the stock market and making correct predictions (a time delay always exists for users’ information toward a certain event), investors are able to make more informed decisions (Bollen et al., 2011; Bordino et al., 2012; Mao et al., 2012; Zhang et al., 2010). 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 research shows that incorporating topic model into opinion mining can help identify users’ opinion distributions on different topics (Lin & He, 2009; Mei, Ling, Wondra, Su, & Zhai, 2007). Additionally, social influence can mitigate the data sparsity...
problem and help improve inference using users’ relationships (Guerra, Veloso, & Almeida, 2011; Tan et al., 2011). Thus, different from existing opinion mining and prediction algorithms, we mainly consider incorporating the topic level opinion influence into our model. According to recent research, incorporating a topic model into opinion mining can help identify users’ opinion distributions on different topics (Lin & He, 2009; Mei et al., 2007). Social influence can also help improve the inference of users’ relationships, especially for the data sparsity problem (e.g., insufficient data for analyzing users’ behaviors; Guerra et al., 2011; Tan et al., 2011).

There are 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. In this paper, we focus on these users and 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. Such opinion influence is topic specific because 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 her/his communication records with neighbors are combined by TOIM to infer user opinion toward a specific object related to a discussion topic. We tested TOIM on Tencent Weibo, one of the largest Chinese microblogging sites, and the results show that TOIM can infer user opinion more accurately at the topic level in the social network than some common baseline methods, if the user is active and frequently communicates with his/her neighbors. We also demonstrate some typical applications of TOIM in analyzing specific users’ behaviors, attitudes, and influence on Tencent Weibo.

Related Work

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 perspectives. Pang, Lee, and Vaithyanathan (2002) studied the problem of classifying documents using overall sentiment identified through machine-learning methods. Hu and Liu (2004) mined opinion features from customers’ online reviews. L. Liu, Tang, Han, Jiang, and Yang (2010) analyzed document-level sentiment by first extracting the comment target, then predicting the polarity of opinions about the target. In recent research, sentiment analysis has been widely applied in social media. Gerald et al. (2013) analyzed users’ behaviors in different social medias, such as Twitter, Facebook, Amazon, and so on, and provided frameworks and solutions to different social media to obtain optimized sentimental analysis results. Moghaddam and Ester (2013) applied different models, such as LDA, HMM, and CRF to realize opinion mining in online reviews, they used tweetfeel, Google Products, Stock Sonar, and so on, as testing data sets to evaluate different methods. Vailios, Daniel, and Trevor (2013) provided a structured multitask regularization method to infer users’ voting intention on Twitter. Leon, Diana, Niraj, and Kalina (2013) demonstrated the importance of making approaches specific to the microblog genre; they used Twitter, Facebook, and so on, as examples, to make a more accurate semantic annotation. Opinion detection and prediction algorithms have also been widely applied in market analysis; for example, Bollen et al. (2011) used the public mood mined from Twitter to predict the stock market. Gruhl, Guha, Kumar, Novak, and Tomkins (2005) predicted book sales by analyzing online chat. Mishne and Glance (2006) analyzed blogger sentiment to predict movie sales. Liu et al. (2007) also used sentiment information from blogs to predict sales performance. However, what distinguishes these works from ours is that their proposed approaches only extract the overall sentiment of a document, and 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 query, and their associated sentiments; authors used OpinionMind (OpinMind.com) to label positive and negative sentiment polarity for queries and topics to formulate a training data set. They generated formulas $P(z_{n,i}, S)$ to express the sentiment probability $S$ of a certain entity $w$ related with document $d$ under a certain topic $z_{j}$, $P(w | \theta)$ to represent entity-topic distribution and $P(w | \theta)$ for entity-sentiment distribution; EM updating formulas were also built for learning all required parameters. Lin and He (2009) proposed a joint sentiment/topic (JST) model based on LDA (Blei, Ng, & Jordan, 2003; Ramage, Hall, Nallapati, & Manning, 2009; Rosen-Zvi, Griffiths, Steyvers, & Smyth, 2004; Zhai, Liu, Xu, & Jia, 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 a 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 her/his textual information (Kim & Hovy 2004; Liu, Zhao, Qin, & Liu, 2010; Riloff & Wiebe, 2003). The main idea of this research was to first build grammar rules, the domain features for sentiment analysis, then apply machine-learning algorithms such as SVM, CRFs, or a propagation algorithm to learn those rules. The rules are mainly about identifying of a word’s sentiment polarity (Kamps, Mokken, Marx, & Rijke, 2004; Takamura, Inui, & Okumura, 2005), identification of subjective and objective sentiment (Mihalcea, Banea, & Wiebe, 2007; Su & Markert, 2009; Wiebe & Mihalcea, 2006), identification of sentiment
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,” 2013). One main purpose of social influence analysis is to detect and evaluate the existence and context of specific forms of social influence (Anagnostopoulos, Kumar, & Mahdian, 2008). Anagnostopoulos et al. (2008) focused on identifying and understanding social influence. They applied 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 proposed a model to identify customers' influence on each other in the customer network. They also built a probabilistic model to mine the spread of influence for viral marketing (Richardson & Domingos, 2002). Similar works maximize the spread of influence through a social network; for example, Kempe, Kleinberg, and Tardos (2003) considered that influence maximization is an NP-hard problem and applied an approximate method to solve it. Tang, Sun, Wang, and Yang (2009) identified social influence on different topics and proposed topical affinity propagation (TAP) to model topic-level social influence. Nallapati and Cohen (2008) applied link-plsa-LDA to model influence in blogs. L. Liu et al. (2010) designed a LDA-based social influence model to detect influential relationships among individuals. Crandall Cosley, Huttenlocher, Kleinberg, and Suri (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 Twitter, put sentiment analysis in the context of a social network. For example, Tan et al. (2011) used 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 a Twitter data set to identify who supported or was against a political celebrity. Guerra et al. (2011) applied transfer learning to use users' communication features to build opinion agreement graphs, thus inferring user opinion about each aspect of an event; for example, they used their model to analyze public 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, L. Liu et al. (2010) suggested conservation and nonconservation methods for mining indirect influence in different heterogeneous networks; however, they focused mainly on user behaviors such as citing, following, and replying.

This research provides a new approach to 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.

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, that is, 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, as it 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.

In Figure 1, David comments on both Allen and Mike’s messages and replies to both of them on the topic of inflation in 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 between David, Allen, and Mike are taken into consideration to calculate the opinion influential relationship about the topic of global economics among them. Specifically, with regard 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 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 in global economics, then David’s opinion on the same object related to 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 (Figure 2) to model that process, which is a multilayer graphical model consisting of four groups of random variables. In the user layer, we use $U = \{u_1, u_2, \ldots, u_V\}$ to
denote all users in Tencent Weibo, who can comment on each other’s messages. In some current research, forwarding behaviors are considered an important index to measure users’ opinion-influential relationships. For example, if user A forwarded user B’s weibos related to a certain topic many times, then A has a high probability to agree with B’s opinion on that topic. However, most of this research seldom considers inferring users’ influential relationships when analyzing 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, 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 \(N = \{n_1, n_2, \ldots, n_k\}\) to denote all entities or objects occurring in all messages represented by \(M = \{m_1, m_2, \ldots, m_l\}\). Some NLP and statistical techniques are used to detect opinions about \(n_i (i \in [1, X])\). In the topic layer, we use \(T = \{t_1, t_2, \ldots, t_k\}\) to denote all topics discussed by users. Those topics are latent variables, which cannot be observed or detected. In the opinion layer, \(o_i\) represents the opinion of user \(u_i (i \in [1, V])\) about topic \(t_i\). 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).

Four probabilistic matrices are defined to connect multiple variables from the four layers. Message layer and topic layer are connected by \(\Phi = \{\phi_i\}_{i \in X}\), where \(\phi_i\) denotes the probability that object \(n_i\) belongs to topic \(t_i\). User layer and topic layer are connected by \(\Theta = \{\theta_j\}_{j \in X}\), where \(\theta_j\) denotes the probability that topic \(t_j\) is selected by user \(u_i\). Opinion layer and user layer are connected by two matrices. One is \(\Psi = \{\psi_{i, m} \}_{i \in X, m \in T}\), where \(\psi_{i, m}\) denotes the probability that user \(u_i\) prefers opinion \(o\) given topic \(t\) and \(o \in \{-1, 0, +1\}\) (−1 represents negative, +1 represents positive, 0 represents no opinion). The other is \(\Omega = \{\omega_{i, j, \text{agree}}, \omega_{i, j, \text{disagree}}\}_{i \in X, j \in T, \theta \in \Theta}\), where \(\omega_{i, j, \text{agree}}\) denotes the probability that \(u_i\) agrees with \(u_j\)’s opinion and \(\omega_{i, j, \text{disagree}}\) denotes the probability that \(u_i\) disagrees with \(u_j\)’s opinion on topic \(t\). Thus, \(\Omega\) represents the agreement/disagreement opinion influence. Besides, \(S = \{s_{i, j, \text{agree}}\}_{i \in X, j \in T, \theta \in \Theta} \cup \{s_{i, j, \text{disagree}}\}_{i \in X, j \in T, \theta \in \Theta}\), where \(s_{i, j, \text{agree}}\) and \(s_{i, j, \text{disagree}}\) are defined as the confidence/strength of \(\omega_{i, j, \text{agree}}\) and \(\omega_{i, j, \text{disagree}}\), respectively. Finally, to reduce the high dimensions of \(X \times Y \times \Theta\), we only consider 2,000 highly active users and their followers are limited to the 1,000 most active and related users. The details are introduced in the next section. The opinion mining problem can be formulated as a mapping function as follows:

\[
f(u_i, n_j, \Phi, \Psi, \Omega, S) \rightarrow o_{n_j}^{u_i}
\]

where \(o_{n_j}^{u_i}\) 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, \(\Theta, \Psi, \Omega,\) and \(S\) are learned simultaneously utilizing statistical and NLP techniques. In the inference phase, given an entity \(n_i\), the opinion about \(n_i\) is inferred using \(\Phi, \Psi, \Omega,\) 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 \(\Theta\) 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.

| TABLE 1. Notations. |
|----------------------|
| **Notations** | **Definitions** |
| \(U\) | The set of users in Tencent Weibo, assume \(U = \{u_1, u_2, \ldots, u_l\}\) |
| \(M\) | The set of weibos in Tencent Weibo, assume \(M = \{m_1, m_2, \ldots, m_l\}\) |
| \(N\) | The set of noun entities, assume \(N = \{n_1, n_2, \ldots, n_k\}\) |
| \(T\) | The set of topics, assume \(T = \{t_1, t_2, \ldots, t_k\}\) |
| \(\alpha x\) | \(\alpha \in \{-1, 1, +1\}\) (−1 represents negative, +1 represents positive, 0 represents no opinion) |
| \(\theta_{ij}\) | User-Topic distribution, \(x \in U, z \in T\) |
| \(\phi_{ij}\) | Topic-Entity distribution, \(z \in T, n \in N\) |
| \(\psi_{ij}^{+}\) | The probability of user \(u_i\)’s positive opinion on topic \(t_j\) |
| \(\psi_{ij}^{-}\) | The probability of user \(u_i\)’s negative opinion on topic \(t_j\) |
| \(\omega_{ij}\) | The probability of user \(u_i\) agrees with \(u_j\)’s opinion on topic \(t_j\) |
| \(\bar{\omega}_{ij}\) | The probability of user \(u_i\) disagrees with \(u_j\)’s opinion on topic \(t_j\) |
| \(\Theta\) | The matrix of author-topic distributions, \(\Theta = \{\theta_j\}_{j \in T, \theta \in \Theta}\) |
| \(\Psi\) | The matrix of topic-entity distributions, \(\Psi = \{\phi_i\}_{i \in X, m \in T}\) |
| \(\Omega\) | A structure to record users’ opinion preference on different topics: \(\Omega = \{\omega_{ij}\}_{i \in X, j \in T, \theta \in \Theta}\) |
| \(S\) | A chain structure to record the confidence/strength of \(\Omega\): \(S = \{s_{ij}\}_{i \in X, j \in T, \theta \in \Theta}\) |

**Method**

In this section, we illustrate the details about how to learn the parameters of users’ opinion influence and make inference. The next section gives the computational equations of five parameters of TOIM (\(\Theta, \Phi, \Psi, \Omega,\) and \(S\)) separately and introduces the unified probability model to estimate the five parameters. The Inference section demonstrates how to use the learned parameters to infer a certain users’ opinion taking the opinion influence of his neighbors into consideration.

**Learning**

We formulize the calculation of five parameters of TOIM (\(\Theta, \Phi, \Psi, \Omega,\) 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.

**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 was proposed by several previous studies (Kamps et al., 2004; Kim & Hovy, 2004; H. Liu et al., 2010; Riloff & Wiebe, 2003; Takamura et al., 2005), and the process can be summarized as follows: First, a parse tree developed by the 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 the Tsinghua NLP group, which consists of 5,567 positive and 4,479 negative words. Finally, two grammar rules are applied to identify the sentimental relation: (a) whether there exists a negation word, such as not or don’t, and (b) whether there exists an adversative relation between $ni$ and $ow(ni)$, such as but or however. As an example, the parse tree for “This product is good and cheap” is shown in Figure 3. All of the 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 a sentence is complete and grammatically correct. However, our preliminary study shows only a small portion (5% to 10%) of messages can be successfully parsed, whereas most other messages are short and noisy because the corresponding part of speech tag. Here, the entity is “product,” and the opinion words are “good” and “cheap.” By searching the sentiment lexicon we conclude that the opinion is positive.

User-topic and topic-entity distribution ($\Theta$ and $\Psi$). Gibbs sampling is used to estimate $\Theta$ and $\Psi$, with two prior hyperparameters $\alpha$ and $\beta$, respectively (Asmussen & Glynn, 2007; Blei & Lafferty, 2006; Gilks, Richardson, & Spiegelhalter, 1995; Walsh, 2004). We construct a Markov chain that converges to a joint posterior distribution on random variables $z$, $x$, $\omega$, and noun $w$, which can then be used to infer $\Theta$ and $\Psi$ (Griffiths & Steyvers, 2004). The diameter between successive states of a 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 $ih$ noun found in $ui$’s message is $n_h$, we can sample a topic for $ui$ based on Equation 3.

$$P(z' = t_k | x = u_i, \omega = n_h, z^{-1}) \propto \frac{C_{z'}^{i+1} + \alpha}{\sum_{z \in U} C_{z}^{i+1} + K\alpha} \frac{C_{i}^{\omega} + \beta}{\sum_{w \in \omega} C_{i}^{w} + N\beta}$$

(3)

where $z' = t_k$ denotes the assignment of the $ih$ noun into topic $t_k$ and $z^{-1}$ denotes all topic assignments not including $n_h$. $C_{z'}^{i+1}$ and $C_{i}^{\omega}$ 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 $ih$ noun. For user $ui$, if $n_h$ also occurs in $uj$’s reply message, $n_h$ is also assigned to topic $t_k$ and $t_k$ is assigned to user $uj$. For all other nouns in $uj$’s replying message, the assignment of words and topics are calculated as shown in Equation 3. The final $\Theta$ and $\Psi$ can be estimated by:

$$\theta_{ik} = \frac{C_{i}^{k} + \alpha}{\sum_{z \in U} C_{z}^{i+1} + K\alpha}, \phi_{ik} = \frac{C_{i}^{\omega} + \beta}{\sum_{w \in \omega} C_{i}^{w} + N\beta}$$

(4)

**Topic-level user opinion distribution ($\Psi$).** Topic-level user opinion distribution characterizes the relative frequency of positive or negative opinions from one user about a certain topic. Similar to the studies by Mei et al. (2007) and Lin and He (2009), we define two counters $C_{i}^{k}$ and $C_{i}^{\omega}$, $i = 1, \ldots, \nu; k = 1, \ldots, K$ to record the number of times user $u_i$ express positive or negative opinions towards topic $t_k$ by scanning all $ui$’s messages. Then $\Psi$ can be estimated as:

$$\psi_{i}^{k} = \frac{C_{i}^{k}}{C_{i}^{\omega} + C_{i}^{k}}, \psi_{i}^{\omega} = \frac{C_{i}^{\omega}}{C_{i}^{\omega} + C_{i}^{k}}$$

(5)

**Topic-level opinion influence ($\Omega, S$).** 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 L. Liu et al.’s (2010) study. In many cases, users discuss 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 U.S. government, whereas user $u_j$ dislikes Gaddafi. Both “U.S. government” and “Gaddafi” belong to the same politics topic, but they 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 and $u_j$ 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 the LDA algorithm on the training corpus and set the number of topics at 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; n) to denote the sentimental relations between ni and n j with OSR(ni; n j) = 1 representing coherent sentiment and OSR(ni; n j) = −1 representing antagonistic sentiment.

We also define two counters, C i,j,agree and C 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 i = −1 · j; OSR(ni; n j) > 0; ni, n j ∈ ti, u i and u j agree on topic t i; otherwise, u i and u j disagree on topic t i. Therefore, Ω can be estimated as:

\[
\omega_{i,j}^{t,\text{agree}} = \frac{C_{i,j}^{t,\text{agree}}}{C_{i,j}^{t,\text{agree}} + C_{i,j}^{t,\text{disagree}}}, \quad \omega_{i,j}^{t,\text{disagree}} = \frac{C_{i,j}^{t,\text{disagree}}}{C_{i,j}^{t,\text{agree}} + C_{i,j}^{t,\text{disagree}}}
\]

(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 low. For user u i, we first assume that all neighbors who have discussed topic t i with u i constitute a set ON(ui, t i), and then we calculate the strength of influence from ON(ui, t i) on u j:

\[
\begin{align*}
S_{i,j}^{t,\text{agree}} &= \frac{C_{i,j}^{t,\text{agree}}}{\max \left( \rho, \sum_{u j \in \text{ON}(u i, t i)} C_{i,j}^{t,\text{agree}} \right)} \\
S_{i,j}^{t,\text{disagree}} &= \frac{C_{i,j}^{t,\text{disagree}}}{\max \left( \rho, \sum_{u j \in \text{ON}(u i, t i)} C_{i,j}^{t,\text{disagree}} \right)}
\end{align*}
\]

(7)

where \( \rho \) denotes a threshold of the minimum agreement/disagreement frequency.

The above detection of opinion influence works well when OSR(ni; n j) is known. However, the values of OSR(ni; n j) are unknown for most pairs of (ni; n j). Therefore, to quantify the opinion relation between u i and u j, other contextual information is used to generate the following opinion agreement index (OAI) that can be used to quantify the opinion influence of u i on u j:

\[
OAI(u_i, u_j) = a \times \text{Influence}(u_i) + b \times \text{Tightness}(u_i, u_j) + c \times \text{similarity}(u_i, u_j)
\]

(8)

where \( \text{Influence}(u_i) \) is the normalized function of u i’s followers, \( \text{Tightness}(u_i, u_j) \) is the normalized function of the interaction (i.e., comment) frequency between u i and u j, and \( \text{similarity}(u_i, u_j) \) is the cosine similarity between \( \Theta \) and \( \Theta \). \( a, b, \) 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 i:

\[
\text{NOAI}(u_i, u_j) = \frac{OAI(u_i, u_j)}{\sum_{u j \in \text{ON}(u_i, t_i)} OAI(u_i, u_j)}
\]

(9)

If u j comments on u i’s one message and opinions influence cannot be determined, then NOAI(u i, u j) can be used to approximate \( s_{i,j}^{t,\text{agree}} \).

**Parameter estimation.** Five parameters of TOIM (\( \Theta, \Phi, \Psi, \Omega, \) 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 i from her/his topic distribution depending on \( \Theta \), and then selects an entity n x associated with t i depending on \( \Phi \), and finally expresses his/her opinion \( o_{n x}^{u_i} \) towards n x. After that, \( \Psi \) 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 applied to u j and u j tends to select the same topic t i as u i. Finally, \( \Omega \) and \( S \) are updated based on both \( o_{n x}^{u_i} \) and \( o_{n x}^{u_j} \). The learning process is listed in Algorithm 1.

**ALGORITHM 1.** Estimation of \( \Theta, \Phi, \Psi, \Omega, \) and \( S \)

**Input:** \( U, M \)

**Output:** \( \Theta, \Phi, \Psi, \Omega, \) and \( S \)

**Initiation:** Iteration \( I \)

\( \text{Pre1:} \) Generate distinct word list \( N; \)

\( \text{Pre2:} \) Construct parse tree for \( m_i \in M; \)

\( \text{Pre3:} \) Calculate \( o_{\Theta}(n) \) for \( n_i \in N \), based on Equation 2;

\( \text{Pre4:} \) Calculate \( GSF(n_i, n_j), n_i, n_j \in N \)

\( \text{Pre5:} \) 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 j who posted \( m_i \); Find all comments CM i on \( m_i \);

For \( n_i \) in \( m_i \) do

Sample topic z j based on Equation 3;
Detect u j’s opinion \( o_{n x}^{u_j} \) of \( n_i \) based on parse tree or Equation 2;
If \( o_{n x}^{u_i} = +1, C_{i,j}^{+} = 1; \) else \( C_{i,j}^{-} = 1; \)
For \( m_i \) 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 \( o_{n x}^{u_j} \) of \( n_j \) based on parse tree or Equation 2;
If \( o_{n x}^{u_i} = +1, C_{i,j}^{+} = 1; \) else \( C_{i,j}^{-} = 1; \)
If \( z_j = z_i \) then

If \( o_{n x}^{u_i} \cdot o_{n x}^{u_j} \cdot \text{OSR}(n_i, n_j) > 0, C_{i,j}^{\text{disagree}} = 1; \) else

\( C_{i,j}^{\text{disagree}} = 1; \)
End

If \( o_i = \text{NULL} \) or \( o_j = \text{NULL} \) or \( \text{OSR}(n_i, n_j) = \text{NULL} \) then

Sample temp from \( NOAI(u_i, u_j); \)
If \( \text{temp} > 0.5, C_{i,j}^{\text{agree}} = 1; \) else

\( C_{i,j}^{\text{agree}} = 1; \)
End End End End End

Calculate \( \Theta, \Phi, \Psi, \Omega, \) and \( S \) based on Equation 4, 5, 6 and 7.
Inference

Once the four parameters are learned through Algorithm 1, we infer the opinion of user $u_i$ about a query object $n_q$, that is, $\phi_i^{n_q}$. First, we find the topic $t_i$ that is most practically related to $n_q$ under $\phi_i$. Second, we find all neighbors of $u_i$ under topic $t_i$, that is, $ON(u_i, t_i)$, and collect all messages between $u_i$ and $u_n \in ON(u_i, t_i)$. Third, for each $u_n \in ON(u_i, t_i)$ we directly detect the opinion of $u_n$ toward $n_q$ from the messages. Fourth, if opinion $\phi_i^{n_q}$ is not neutral, we sample the type of opinion influence (agree or disagree) of $um$ on $ui$ from $\psi$. Finally, we obtain the $\phi_i^{n_q}$ from a linear combination of the two types of opinions under a random sampling technique: $ui$’s historical opinion about topic $t_i$ sampled from $\psi_i$, and $ui$’s opinion influence purely deducted from $u_n \in ON(u_i, t_i)$ and the corresponding type of influence. The exact inference process is shown in Algorithm 2. It is worth noting that TOIM may not be able to infer a user’s opinion if the provided information is insufficient.

**Algorithm 2. Opinion Inference**

```
Input: $\theta$, $\phi$, $\psi$, $\omega$ and $\psi$, user $u_i$, object $n_q$, weight $\omega$
Output: opinion $\phi_i^{n_q}$

Initiation: Iterations Iter, $\omega$, SWO = 0
For e = 1:Iter do
    Find the most probabilistically related topic $t_i$ regarding $n_q$ from $\phi_i$;
    For user $u_n \in ON(u_i, t_i)$ do
        If $u_n$’s opinion $\phi_i^{n_q}$ is known then
            Set $temp = \omega \times \phi_i^{n_q} + (1-\omega)\times\phi_i^{n_q,agree}$; sample $\phi_i^{n_q}$ from $temp$;
            If $\phi_i^{n_q,agree} = \phi_i^{n_q}$ then
                SWO += $\phi_i^{n_q,agree}$; $\phi_i^{n_q,disagree}$;
            Else
                SWO += $\phi_i^{n_q,disagree}$;
            End
        End
    End
    If SWO $>$ = 0.5 then $\phi_i^{n_q}$; +1;
    Else if $\phi_i^{n_q} <= -0.5$ then $\phi_i^{n_q}$; -1;
    Else $\phi_i^{n_q}$ is unknown;
End
```

Experiment

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 $u_n$ are connected by an edge if $u_i$ (or $u_n$) comments on or replies to $u_i$ (or $u_n$)’s messages.

We collected messages from Tencent Weibo between October 1, 2011, and January 5, 2012, with an average of 30 to 60 million messages posted daily. We found that the most frequent user action was “forwarding,” followed by “commenting,” and few records were about “replying.” This indicates that a large number of Tencent Weibo users prefer to discuss with their friends and express their own ideas by commenting on messages around a similar topic, which provides historical records for us to learn the topic-level opinion influence needed by TOIM. To evaluate TOIM, we selected five hot entities that were discussed frequently in Tencent Weibo during this 3-month 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 searched all relevant users who once mentioned $Oi$ in their messages during the 3-month period. Among all retrieved users, the top 2,000 most active users (i.e., the users who created many “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 obtained plenty of textual information about their communication records, which is used by TOIM to better infer their topical opinion.

Given the 2,000 selected users, we further generated four types of data. First, all messages posted by the 2,000 users and related to $Oi$ were chosen and denoted $D1$. We designed a bag of keywords, which included alias, synonym, error-correction terms, and other important words to describe each $Oi$. For example, the movie The Flowers of War has many aliases in Chinese, and all aliases were collected to identify the right weibos; “School bus accident” was a controversial event that aroused the public’s attention toward public safety. We also used a combination of terms to detect weibos that were related to the entity “College talk from the president of Peking University,” such as “The name of the president” plus “College talk.” We manually labeled all 2,000 users’ opinions toward the five entities based on their weibos; for example, if user A held positive opinions toward “Chinese economics,” then we designed 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 to “Chinese economics” and our labeled opinion information “+1” of user A toward “Chinese economics.” $D1$ is considered testing data; in our experiment, we wanted to use the learned model to predict user A’s opinion (+1 or -1). If the model is reasonable and well-trained, we gain a high precision.

Second, all messages posted by the 2,000 users excluding $D1$ are denoted $D2$. 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 were crawled via one-hop commenting and replying actions,
where all related commenting and replying messages are extracted and denoted \( D_3 \) (to reduce the high dimension of the user-user matrix, we set the limitation as the 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 were also collected (did not include the 2,000 users themselves) as \( D_4 \). The learning of those users’ behaviors helps construct stronger influence relationships between current users and the selected 2,000 users.

Specifically, we used \( D_1 \) as the testing corpus to manually decide the related users’ opinions about \( O_i \). We used \( D_2, D_3 \) plus \( D_4 \) as a training corpus to learn the topic-level user opinion distribution and topic-level opinion influence, respectively. The data format of the training data are summarized as:

\[
\#IDX"The ID of current weibo"
\#"Content of current weibo"
\#@“Username of current weibo”
\#T"TimeStamp"
\#C"IDX"
\#END
\]

The data description is in Table 2.

**Results**

We used two classic measures to evaluate the inference performance of TOIM, precision and recall, but slightly modified them. 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, whereas 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), served 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, the three algorithms do not take opinion influence into consideration. We adopted SVM-light\(^4\) for SVM, adopted the code from Tang et al. (2009) for CRF, and implemented the work of Lin and He (2009) for JST. For those baseline models, the inference of user opinions are mainly based on their historical opinions of their posted messages. We extracted user name, key words, topics, entities, and users’ sentiment toward each entity and other information from each message as attributes and then applied SVM, CRF, and JST to make inferences. To generate the training and testing data set, four attributes of each user are defined as (a) username, (b) nouns with their weighted score, (c) qualifiers of the nouns, and (d) topics ID related with those nouns. For example, a user \( X_t \) 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\}. Noun \( i \) \((i = 1, 2, . . .) \) is a word, which is mentioned in \( mb \). KeyWord extraction technology was used to compute the weight of each noun and their qualifiers according to their grammar position in \( mb \): (a). The noun with highest score should be the most important core word. (b) User’s attitude toward this noun could be considered as the input of “Opinion.” (c) “Weight0” of “Username” is assigned as 1, while “WeightZ” of “Topic ID” is the score of “Topic ID” on \( mb \). All of the inferred results from each model were normalized from \(-1\) to 1. The normalization methods are summarized as:

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:

\[
f(S, X, T, W) = S \times P(SIX, T, W)
\]

s.t. \( S \in [-1, 0, +1] \) (10)

where \( S \) means sentimental polarities, \(+1\) means positive sentiment, \(-1\) means negative sentiment, and 0 means no sentiment preferences. The results are shown in Figures 4 and 5.

Figure 4 shows the precision of opinion inference for entities \( O_1, O_2, O_3 \) using four algorithms including TOIM and three baseline methods, against a different number of topics. We 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 it reaches a plateau. By contrast, the other three methods do not show such a 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, the topics “politics” and “military” may be mixed together when

\(^4\)http://svmlight.joachims.org/
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 opinion influence, are obviously not affected by \( K \) because we did not predefine 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 \( O_1 \), \( O_2 \), \( O_3 \). 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 of correct or incorrect) given the sparseness of the data. Therefore, when \( K \) is small, we have a high probability of tracking 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 small, thus TOIM fails to detect the opinion influence at the topic level, leading to low recall. While for the other three algorithms (SVM, CRF and JST as baseline) without considering the opinion influence, the experimental results show that precision and recall are not significantly affected by the number of topic assignment \( K \).

The overall inference performance of TOIM exhibits low recall but higher precision values when compared with the other three baseline methods. Because text messages on microblogging are generally noisy and short, many 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 users’ opinions towards a certain topic from a large number of opinion-detectable users, then TOIM is not a good choice because too many topic partitions 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 inference precision.

To better illustrate the performance of TOIM, we assigned 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 a receiver operating characteristic (ROC) for objects \(O_1\), \(O_2\), \(O_3\), as in Figure 6.

In Figure 6, we plot ROC curves to further evaluate the performance of our proposed model. The 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 see that the AUC of TOIM is significantly larger than other methods, and its ROC curves are all above the diagonal line for three selected entities, \(O_1\), \(O_2\), and \(O_3\), implying that TOIM is a more accurate method to judge a user’s sentiment toward a certain topic than other baseline methods.

According to L. Liu et al.’s (2010) study, incorporating indirect influence can further improve the performance of the influence model. Indirect influence represents the influence relations between two users who do not have direct connections. For example, users 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 important 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 L. Liu et al.’s (2010) and Sun and Han’s (2013) work, we mainly used conservative propagation (CP) to calculate indirect influence. The performance of combining indirect influence is shown in Figure 7.

Figure 7 shows that incorporating indirect influence in TOIM not only can improve average precision but also can improve average recall and the F1-measure significantly. The reason is that indirect influence helps to build influence connections between two users who are not directly connected and thus provide more information regarding the users’ opinions inferences.

**Case Study**

As mentioned previously, 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.

**Opinion Influence Mining**

Influence vs. noninfluence opinion mining. In the Results section, we show that TOIM outperforms the other three baseline methods in opinion inference precision because TOIM considers the opinion influence deduced from the social networks. Here we select five representative users and infer their opinions regarding 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.
As an example, we 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 a positive opinion toward “Muammar Gaddafi,” and the other three disagree with an average probability of 0.3253 to have a positive opinion. One infers 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 toward “Muammar Gaddafi.” The five examples indicate that, given sufficient information about a user’s neighbor’s opinion and his or her corresponding opinion influence types and strengths, TOIM can detect the topic-level opinions more accurately than other baseline methods, such as CRF.

**Opinion leader recognition.** Opinion leaders refer to those users whose opinions on some topics 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 \( o_\omega \) on topic \( tz \) with user \( u_x \), then we increase \( CUTOxzw \) by one. We select the top three users ranked by influential scores and show them in Figure 8.

**TABLE 3. The comparison between topic-level opinion influence model (TOIM) and conditional random field (CRF) in opinion inference for five selected users.**

| User (true opinion) | Method | Agree (average probability) | Disagree (average probability) | Inferred results |
|---------------------|--------|-----------------------------|-------------------------------|-----------------|
| 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 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 a positive opinion toward “Muammar Gaddafi,” and the other three disagree with an average probability of 0.3253 to have a positive opinion. One infers 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 toward “Muammar Gaddafi.” The five examples indicate that, given sufficient information about a user’s neighbor’s opinion and his or her corresponding opinion influence types and strengths, TOIM can detect the topic-level opinions more accurately than other baseline methods, such as CRF.

**Opinion leader recognition.** Opinion leaders refer to those users whose opinions on some topics 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 \( o_\omega \) on topic \( tz \) with user \( u_x \), then we increase \( CUTOxzw \) by one. We select the top three users ranked by influential scores and show them in Figure 8.

**We check the profiles of those opinion leaders and find their backgrounds are 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 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 3-month period and find that they are all active and that their messages are forwarded and commented on by thousands of users. For instance, YZ Qiu often went to conflict zones (i.e., Libya) to interview local residents and published up-to-date news reports. He is also popular among Tencent users and often participates 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 on by thousands of users.

**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 9 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 economics 32 times. Those values can be considered as a confidence for influential degree. If two users have agreed for times, the influence inference will be more accurate.

In particular, 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 economics, and have high probabilities of agreement with his opinions.

**Topic-Level Opinion Mining**

**Topic-opinion preference detection for celebrities.** We selected four user accounts already verified by Tencent Weibo and detected 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, she/he gains a high influence score. The third column lists the frequently discussed topics of each user, where the words in parentheses are the expression of that topic, and the value behind each topic is the weight of a users’ opinion preferences. For example, for Topic 1, user LYang (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 used often to describe the current user’s opinions on a certain topic. For example, “university” means that user LYang (622007070)
FIG. 9. The direct opinion influence map of XP Lang on Chinese economics. The central node is XP Lang and the other nodes are opinion neighbors who frequently communicate with XP Lang and show agreement with him. Some marked nodes with identification information are verified celebrities.

TABLE 4. User’s opinion preference on different topics represented by different words.

| User ID      | Profile                  | Topic preference | Opinion probability | Influential score (max) | Representative words                                      |
|--------------|--------------------------|------------------|----------------------|-------------------------|----------------------------------------------------------|
| L Yang       | Famous host              | Topic 1: College & Education | 0.7101+              | 0.0063+(0.034)          | POS: daughter+, exam+, university+                       |
| (622007070)  |                          | Topic 2: Daily Emotion          | 0.6845+              | 0.0154+(0.0235)         | NEG: primary school–, student–                          |
|              |                          | Topic 3: Country & History     | 0.4729+              | 0.0132+(0.0274)         | POS: happiness+, expectation+, women+                    |
|              |                          | # of followers: 10,116,317     |                      |                         | NEG: encounter–, home–, violence–                        |
|              |                          | # of messages: 28              |                      |                         | POS: women+, china+, army+                               |
|              |                          | # of comments: 30,700           |                      |                         | NEG: war–, power–, hurt–                                |
| Y Qin        | Famous financial officer | Topic 1: College & Education | 0.7907+              | 0.0154+(0.034)          | POS: research+, university+, student+                    |
| (394865678)  |                          | Topic 3: Social Affairs        | 0.5463+              | 0.0102+(0.0274)         | NEG: protection+, children+, driver+                     |
|              |                          | Topic 23: Country Development  | 0.4729+              | 0.0052+(0.0193)         | NEG: accident–, corruption–                             |
|              |                          | # of followers: 2,186,021       |                      |                         | POS: innovation+, investment+                            |
|              |                          | # of messages: 36              |                      |                         | NEG: enterprise–, lawsuit–, stock–                      |
|              |                          | # of comments: 6,242            |                      |                         | POS: university+, student+, china+                       |
| Christopher   | Famous researcher        | Topic 1: College & Education  | 0.5777+              | 0.0087+(0.034)          | NEG: education–, research–, lost–                        |
| (1484189007) |                          | Topic 13: Economics            | 0.3227+              | 0.0064+(0.0422)         | POS: innovation+, medical+, economics+                   |
|              |                          | Topic 27: International Politics | 0.3939+              | 0.0042+(0.0223)         | NEG: cost–, technology–, company–                        |
|              |                          | # of followers: 800,770         |                      |                         | POS: charm+, people+, solution+                          |
|              |                          | # of messages: 78              |                      |                         | NEG: relationship–, criticism–, protest–                |
|              |                          | # of comments: 31,642           |                      |                         | POS: reform+, tax+, finance+                             |
| XJ Yang      | Vice editor in finance   | Topic 13: Economics           | 0.2988+              | 0.0076+(0.0422)         | NEG: industry–, monopoly–, debt–                       |
| (790147482)  |                          | Topic 14: Social Study         | 0.3267+              | 0.0062+(0.0186)         | POS: democracy+, institution+, law+                     |
|              |                          | Topic 23: Country Development  | 0.5232+              | 0.0058+(0.0193)         | NEG: officers–, society–, market–                      |
|              |                          | # of followers: 146,718         |                      |                         | POS: charm+, people+, solution+                          |
|              |                          | # of messages: 47              |                      |                         | NEG: income–, ination–, welfare–                       |
|              |                          | # of comments: 14,140          |                      |                         |                                                           |

The superscript symbol “+” denotes positive opinion or agree relationship, whereas “−” denotes negative opinion or disagree relationship.
has ever held positive opinions toward that word on Topic 1. To detect the representative words for topics under different opinions, we define a chain counter $CUTOW_{xyzw}$ 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 $CUTOW_{xyzw}$ by one, during the learning process. Then we can use $CUTOW_{xyzw}$ to calculate the weight of each entity toward user $u_x$ under topic $t_y$.

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 a certain degree of positive influences on their neighbors in their interested topics; “influential score” evaluates the influence of each celebrity in each topic, “max” means the maximum influential score, which is from the most influential users, in a corresponding topic. We find that although 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, although they are less well-known. Although the celebrities may not be the most influential users in their interested topics, their influence ranks are all in the top 20, implying that their opinions are still generally supported by many Tencent followers.

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 a Chinese economic index: Hushen-300 Index 5. Figure 10a shows the volatility of Tencent moods about O3 (i.e., Chinese economics), calculated using TOIM during the 3-month period; Figure 10b shows the volatility of Hushen-300 trend for the same time period. 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 10a vs. area 1 in Figure 10b). The statistical analysis shows that the R square of such a 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 the real-world economy than other simple opinion mining method.

Conclusions and Future Work

This paper investigated the problem of inferring user opinion by identifying opinion influence in social networks. A topic-level opinion influence model (TOIM) is proposed and tested on Tencent Weibo, an established 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 regarding specific topics. To test and evaluate the proposed model, an experiment was conducted based on 3-month data from Tencent Weibo. The results show that the proposed TOIM effectively combines social influence and topic preference and outperforms 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 frequently. We demonstrate that the inferred opinion from TOIM can be applied to detect celebrities’ opinions toward various topics, identify the collective opinion correlations with real-world phenomena, visualize opinion influence structure, and identify opinion leaders in different domains.

There are several limitations of TOIM, which need further study. First, a user’s true opinions can be misunderstood. The reason is that human language expression can be complicated in terms of innuendo, irony (users use lots of “innuendo” and “irony” in weibos, TOIM can not identify them and make wrong judgment), 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 challenges including how to design and obtain dedicated topic labels, and how to effectively preprocess experimental data sets (e.g., delete noisy information, use training data to denote constraint rules for learning algorithms).

Acknowledgments

This study 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).

References

Anagnostopoulos, A., Kumar, R., & Mahdian, M. (2008). Influence and correlation in social networks. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 7–15). New York: ACM.
Asmussen, S., & Glynn, P.L. (2007). Stochastic simulation: Algorithms and analysis. Series: Stochastic Modeling and Applied Probability, Vol. 57, 2007. Berlin: Springer.
Blei, D., & Lafferty, J.D. (2006). Dynamic topic models. In Proceedings of the 23rd International Conference on Machine Learning, Pittsburgh, PA.
Blei, D.M., Ng, A.Y., & Jordan, M.I. (2003). Latent Dirichlet allocation. Journal of Machine Learning Research, 3 (4–5), 993–1022.
Bloom, K., Garg, N., & Argamon, S. (2007). Extracting appraisal expressions. In C.L. Sidner, T. Schultz, M. Stone, & C. Zhai (Eds.), HLT-NAACL (pp. 308–315). The Association for Computational Linguistics.
Bollen, J., Mao, H., Zheng, X. (2011). Twitter mood predicts the stock market. Journal of Computational Science, 2, 1–8.
Bordino, I., Battiston, S., Caldarelli, G., Cristelli, M., Utkonen, A., & Weber, I. (2012). Web search queries can predict stock market volumes. PLoS One, 7.
Crandall, D., Cosley, D., Huttenlocher, D., Kleinberg, J., Suri, S. (2008). Feedback effects between similarity and social influence in online communities. In Proceedings of the 14th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 160–168). New York: ACM.
Domingos, P., & Richardson, M. (2001). Mining the network value of customers. In Proceedings of the Seventh ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 57–66). New York: ACM.
Gerald, P., Michal, K., Harald, F., Andreas, A., Vlach, S., & Andreas, H. (2013). Opinion mining on the Web 2.0—Characteristics of user generated content and their impacts. HCI-KDD 2013, LNCS 7947, 35–46.
Gilkis, W.R., Richardson, S., & Spiegelhalter, D. (1995). Markov Chain Monte Carlo in practice. New York: Chapman & Hall/CRC Interdisciplinary Statistics.
Griffiths, T.L., & Steyvers, M. (2004). Finding scientific topics. In Proceedings of the National Academy of Sciences (PNAS), 101, 5228–5235.
Grulh, D., Guha, R., Kumar, R., Novak, J., & Tomkins, A. (2005). The predictive power of online chatter. In Proceedings of the 11th ACM SIGKDD International Conference on Knowledge Discovery in Data Mining (pp. 78–87). New York: ACM.
Guerra, P.H.C., Veloso, A., Meira, W., Jr., & Almeida, V. (2011). From bias to opinion: A transfer-learning approach to real-time sentiment analysis. In Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 150–158). New York: ACM.
Hu, M., & Liu, B. (2004). Mining opinion features in customer reviews. In Proceedings of the 19th National Conference on Artificial Intelligence (pp. 755–760). Menlo Park, CA: AAAI Press.
Kamps, J., Mokken, R.J., 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, vol. IV, pp. 1115–1118). Paris: European Language Resources Association.
Kempe, D., Kleinberg, J., & Tardos, E. (2003). Maximizing the spread of influence through a social network. In Proceedings of the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 137–146). New York: ACM.
Kinn, S.M., & Hovy, E. (2004). Determining the sentiment of opinions. Retrieved from http://dx.doi.org/10.3115/1220355.1220555
Leon, D., Diana, M., Niraj, A., & Kalina, B. (2013). Microblog-genre noise and impact on semantic annotation accuracy. In Proceedings of 24th ACM Conference on Hypertext and Social Media. Paris.
Li, D., Shuai, X., Sun, G., Tang, J., Ding, Y., & Luo, Z. (2012). Mining topic-level opinion influence in microblog. In Proceedings of the 21st ACM International Conference on Information Management and Knowledge Management (CIKM ’12, pp. 1562–1566). New York: ACM. Retrieved from http://doi.acm.org/10.1145/2196761.2198473
Lin, C., & He, Y. (2009). Joint sentiment/topic model for sentiment analysis. Retrieved from http://doi.org/10.1145/1645953
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 (SIGIR ’07, pp. 607–614). New York: ACM. Retrieved from http://doi.acm.org/10.1145/1277741.1277845
Liu, L., Tang, J., Han, J., Jiang, M., & Yang, S. (2010). Mining topic-level influence in heterogeneous networks. In Proceedings of the 19th ACM International Conference on Information and Knowledge Management (pp. 199–208). New York: ACM.
Liu, H., Zhao, Y., Qin, B., & Liu, T. (2010). Comment target extraction and sentiment classification. Journal of Chinese Information Processing, 24, 84–88.
Mao, Y., Wang, B., Wei, W., & Liu, B. (2012). Correlating S&P 500 stocks with Twitter data. HotSocial’12, Beijing, China.
Mee, 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 (WWW ’07, DOI: 10.1002/asi 2672 JOURNAL OF THE ASSOCIATION FOR INFORMATION SCIENCE AND TECHNOLOGY—December 2015
