Modeling Opinion Influence with User Dual Identity

Chengyao Chen, Zhitao Wang, Wenjie Li
Department of Computing, The Hong Kong Polytechnic University, Hong Kong
{cscchen,csztwang,cswjli}@comp.polyu.edu.hk

ABSTRACT

Exploring the mechanism that explains how a user's opinion changes under the influence of his/her neighbors is of practical importance (e.g., for predicting the sentiment of his/her future opinion) and has attracted wide attention from both enterprises and academics. Though various opinion influence models have been proposed for opinion prediction, they only consider users' personal identities, but ignore their social identities with which people behave to fit the expectations of the others in the same group. In this work, we explore users' dual identities, including both personal identities and social identities to build a more comprehensive opinion influence model for a better understanding of opinion behaviors. A novel joint learning framework is proposed to simultaneously model opinion dynamics and detect social identity in a unified model. The effectiveness of the proposed approach is demonstrated through the experiments conducted on Twitter datasets.

CCS CONCEPTS

•Information systems → Data mining; •Applied computing → *Sociology;*

KEYWORDS

Opinion influence modeling; Dual identity; Joint learning

1 INTRODUCTION

Social media services provide effective platforms for people to make friends, share interests and exchange opinions on all facets of life. During the communication on a specific topic, people accumulate the required information and are likely to be influenced to change their opinions [6]. Naturally, social media offers a great chance for companies to monitor the opinions of their customers and to adjust their marketing strategies to get more positive responses. In order to achieve these goals, it is of great importance to understand the intrinsic mechanism behind the dynamics of opinion behaviors and to further predict the future opinions that have not yet been delivered.

Informational influence has been taken as a primary process of opinion formation about commercial products in social media [6]. It describes the following scenario: when people lack the necessary information, they will seek for the opinions from their neighbors (e.g., their friends in social media) and update their future opinions (or opinion behaviors). Note that opinion and opinion behavior are two interchangeable terms in this paper. Based on the theory of informational influence, researchers have proposed several opinion influence models to infer the interpersonal influence between user pairs, and predict a user's future opinion with the aggregated opinions of her/his neighbors [1, 3, 5, 6]. These studies all assume that the future opinions of users only depend on their relationships with the communicating neighbors while other users and the relationships to them within the whole social network have no effect. The existing opinion influence models are therefore targeted to model the opinion dynamics of each user individually, and learns the interpersonal influence a user receives from her/his neighbors according to their historical communication records.

Apart from being a unique person (i.e., *personal identity*) in communication, a user also possesses another identity, which is her/his *social identity* within the whole social network [8]. With the adopted social identity, a user's behavior tends to fit the expectations of the members in the same group. For example, a user who plays as an "expert" in the discussion on a specific product would like to post the professional information and advices to maintain her/his position. Usually, an expert is trusted by people and have positive influence. To capture the effects of social identities in the dynamics of user behaviors, Yang et al. propose to learn the influences of three structural roles and exploit them for reposting behavior prediction [8]. Though personal identity and social identity have been independently studied, to the best of our knowledge the effect of users' dual identities on their behaviors has not yet been explored. The focus of this paper is to study the dynamics of opinion behaviors for future opinion sentiment prediction by taking into account both personal identities and social identities.

To achieve this goal, we propose a Dual Identity based opinion Influence model (DI^2), which has the abilities to recognize users' social identities and to capture users' opinion influence based on their dual identities through opinion dynamics. Because the criteria used to define the social identities of the users who are interested in different topics varies, it is very difficult to devise a unified standard to categorize users' social identities. We cast social identity detection as a task of user clustering. By representing users with the textual features corresponding to their opinion contents and structural features corresponding to their network properties, we divide users into different groups with different social identities. Considering the dual identities of users, the individual-based influence and group-based influence are integrated in DI^2. The group-based influence enhances the individual-based influence, especially when the latter is difficult to learn from the insufficient communication between two users. Due to the interplay of social identity detection and opinion influence modeling [8], an efficient algorithm is developed to jointly infer users' social identities and learns the dual identity based opinion influence model. Using the learned DI^2, the sentiment of a user' future opinion can then be predicted using

CIKM'17, November 6–10, 2017, Singapore.
 ISBN 978-1-4503-4918-5/17/11…$15.00
DOI: https://doi.org/10.1145/3132847.3133125

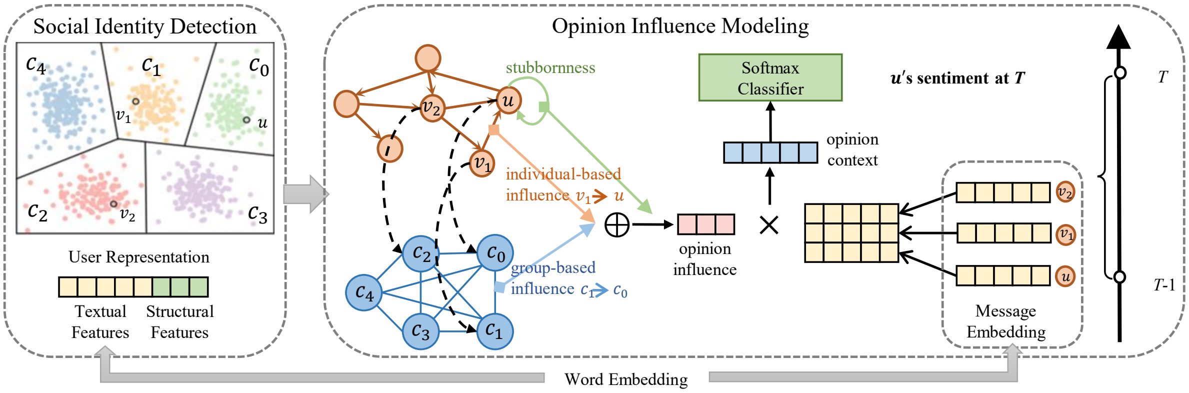

Figure 1: General view of DI^2

her/his historical opinions and the influence from her/his neighbors. The experiments conducted on three Twitter datasets demonstrate the effectiveness of DI^2 with its creative idea of exploring dual identities and the joint learning framework.

2 METHOD

The proposed DI^2 contains two major components, as briefly illustrated in Figure 1. The backbone of social identity detection is a soft K-means clustering algorithm. The opinion influence model is developed based on the neural network architecture to take advantage of its powerful semantic representation ability. Social identity detection and opinion influence modeling are learned jointly by the shared word embeddings.

2.1 Social Identity Detection

User Representation

We characterize users in two types of features. One is the textual features, which describe a user's personal interests and opinions. The other is the structural features, which reflect a user's position within the whole social network.

Given a set of N users in the set V, for each $u \in V$, her/his posting sequence is $S_u = \{S_u(1), \cdots, S_u(i), \cdots, S_u(n(u))\}$. Each element $S_u(i) =< M_u(i), o_u(i), t_u(i) >$ represents a message $M_u(i)$ posted by u at time $t_u(i)$ and with the sentiment of opinion $o_u(i)$. $o_u(i)$ can take the values of -1, 0 and 1 indicating the negative, neutral and positive sentiment respectively. The message $M_u(i)$ is represented by the opinion words included, i.e., $M_u(i) = \{w_{u,1}(i), w_{u,2}(i), \cdots\}$. Inspired by the semantic representation ability of word embeddings [2], we represent each opinion word w as a dense and continuous vector $\Phi(w)$ with the dimension d_w. We capture a user's textual features by all of her/his posted messages as follows.

$$\boldsymbol{h_u} = \frac{\sum_{i=1}^{n(u)} \Phi(M_u(i))}{n(u)} \tag{1}$$

where $\Phi(M_u(i)) = \sum_j \Phi(w_{u,j}(i))$.

In addition to the textual features, three structural features are selected to constitute the vector $\boldsymbol{r_u}$. They are the number of followers, the number of followees and a binary value indicating whether a user account has been verified to be authentic or not.

Given the textual features and structural features, we construct the user representation denoted as $\boldsymbol{x_u}$ with the dimension d_u.

$$\boldsymbol{x_u} = [\boldsymbol{h_u}; \boldsymbol{r_u}] \tag{2}$$

User Clustering

The popular K-means clustering algorithm [4] is utilized to aggregate users with similar characteristics into the same group. The center of cluster c_j is denoted by the vector $\boldsymbol{\theta_j}$, and its dimension is d_u, which is same as the dimension of user representations. The total number of social identities is K. Intuitively, a user may play multiple social identities with respect to different communities or groups. Here, we use the soft version of K-means, which allows each user to be assigned to multiple clusters with a probabilistic distribution $\boldsymbol{z_u}$ over all clusters. The probabilities are calculated as the distances between a user and K cluster centers, i.e.,

$$\boldsymbol{z_u} = softmax(-\boldsymbol{l_u}) \tag{3}$$

where $l_{uj} = \|\boldsymbol{x_u} - \boldsymbol{\theta_j}\|_2$

Note that to better illustrate the proposed model in Figure 1, we only show one assigned cluster for each user instead of presenting her/his probability distribution.

2.2 Dual-identity Opinion Influence Model

Neighboring influence has been demonstrated as the most important factor for a user to change her/his future opinion [6]. DI^2 learns two types of opinion influence between a user and her/his neighbors through a neural network framework. One is the individual-based influence corresponding to the personal identity and the other is the group-based influence corresponding to the social identity.

We denote the set of network connections by $E \subseteq V \times V$, which represents the following relationships between two users. We use a matrix $\boldsymbol{A} \in \mathbb{R}^{N \times N}$ to represent the influence between a user pair. $\boldsymbol{A}$ is naturally constrained by the network connections, i.e., α_{uv} is the individual-based influence v exerts on u if $(u, v) \in E$, and 0 otherwise.

The influence of group c_i on group c_j is denoted as β_{ij}, where the matrix $\boldsymbol{B} = [\beta_{ij}] \in \mathbb{R}^{K \times K}$. As a user plays multiple social identities with a probabilistic distribution, the group-based influence between v and u is:

$$e_{uv} = \boldsymbol{z_u} \boldsymbol{B} \boldsymbol{z_v}^T \tag{4}$$

We integrate the individual-based influence and group-based influence through a balance weight to describe the whole opinion influence between a user pair.

$$inf_{uv} = sigmoid(\lambda_{uv}) * \alpha_{uv} + (1 - sigmoid(\lambda_{uv})) * e_{uv} \quad (5)$$

λ_{uv} is the weight to balance the contribution of individual-based influence and group-based influence. We assume that the contributions of two types of influence vary for each user pair, and it is relevant to their interaction frequencies. We set $\lambda_{uv} = \lambda_0 + \omega * f_{uv}$, where λ_0 is a parameter shared by all users, f_{uv} is the interaction frequency between each user pair and ω denotes its weight. f_{uv} is measured by the percentage of the messages that u changes opinions after receiving the opinions from v. In addition to the neighboring opinion influence, a user also tends to insist on her/his prior opinion during the communication [6]. We use α_{uu} to indicate the degree of u's stubbornness on the prior opinion.

In DI^2, a neural network framework is applied to capture the opinion influence based on dual identities. It takes the prior messages from a user's neighbors and her/his own as the input, and after integrating the messages with the neighboring opinion influence and personal stubbornness, a hidden vector is constructed, which can be further utilized to infer a user's future sentiment in the output layer. Supervised by the gold-standard labels of a user's future sentiments, DI^2 learns both individual-based influence and group-based influence.

According to the linear property of influence [6], the hidden opinion context vector is constructed as:

$$\boldsymbol{c_u(i)} = \alpha_{uu}\Phi(M_u(i-1)) + \sum_{v}^{Nei(u)} inf_{uv}\Phi(M_v(t_v)) \quad (6)$$

where $Nei(u)$ represents the set of u's neighbors, and $t_u(i-1) < t_v < t_u(i)$. Given the opinion context vector, the output layer is a softmax function to output the probabilities over all types of sentiment.

$$P(o_u(i)|\boldsymbol{c_u(i)}) = softmax(\boldsymbol{V}\boldsymbol{c_u(i)} + \boldsymbol{b}) \quad (7)$$

where $\boldsymbol{V} \in \mathbb{R}^{m \times d_w}$, and $\boldsymbol{b} \in \mathbb{R}^m$. m is the number of sentiment polarities, which is 3.

2.3 Joint Learning Framework

There are two terms in the loss function of DI^2, including the loss function of the neural network for sentiment classification, and the loss function for soft K-means clustering.

$$\epsilon = -\sum_{u=1}^{N}\sum_{i=1}^{m(u)} \log P(o_u^*(i)) + \sum_{u=1}^{N}\sum_{j=1}^{K} z_{uj}\|\boldsymbol{x_u} - \boldsymbol{\theta_j}\|_2 \quad (8)$$

where $o_u^*(i)$ represents the sentiment label of u's message at $t_u(i)$.

The learning algorithm is summarized in Algorithm 1. Mini-batch is used during model training. In the forward pass, given the word embedding Φ updated in the previous batch, we compute the user representations and their probabilistic distributions over all the clusters for each user in the current batch. Based on the new cluster assignments of users, we estimate the cluster centers $\hat{\Theta}$. Because $\hat{\Theta}$ is computed according to the users within the current batch, it cannot represent the partition over all users. We use the online updating algorithm to update the cluster centers, i.e., the updated cluster center Θ is the combination of the former cluster center Θ' and the cluster center $\hat{\Theta}$ learned from the current batch. In the backward pass, given the estimation of cluster centers Θ, we update the word embedding and the other parameters of the opinion influence model with the stochastic gradient decent (SGD) algorithm. The training phase stops when the training error has a decrease less than 1 or reaches the maximum iteration $R = 100$.

Algorithm 1 Joint learning algorithm

```
Initialization: Maximum training iteration R, word representation Φ, cluster center Θ, former cluster center Θ', individual-based influence A, group-based influence B, balance weight λ_0, weight of interaction frequency ω, output parameters V, b and smooth weight δ
for iteration i =1,2,...,R do
    for each batch bs do
        Given the updated Φ in previous batch, compute user representation x_u for u ∈ BU, where BU is the set of users included in the batch bs
        Calculate the soft assignment z_u for each user according to Equation 3.
        Compute the new cluster center θ̂_j = 1/|BU| Σ_{u∈BU} z_uj x_u
        Smooth cluster centers Θ = δΘ̂ + (1 − δ)Θ'
        set Θ as Θ'
        Update the parameters Φ, A, B, V and b, λ_0, ω with the SGD algorithm given the cluster center Θ
    end for
end for
```

Table 1: Statistics of dataset

Method	Samsung Galaxy	Xbox	PlayStation
# of users	8921	4358	5158
# of avg messages	14.42	9.58	11.83
# of opinion words	878	1146	505
% of negative sentiment	11.07%	16.34%	12.00%
% of neutral sentiment	48.43%	41.57%	62.92%
% of positive sentiment	40.50%	42.09%	25.08%

3 EXPERIMENTS

3.1 Experimental Setup

Dataset. We evaluate the proposed model on three Twitter datasets [5]. The datasets contain the communication records of Twitter users and the structure of the user network on three topics, i.e., "**Samsung Galaxy**", "**Xbox**" and "**PlayStation**". Additionally, given users' ids, we collect their user profiles via the Twitter API to obtain the corresponding structural features. The statistics of the datasets are summarized in Table 1.

Sentiment Prediction. We evaluate the performance of DI^2 on the sentiment prediction task. Given the learned opinion influence and personal stubbornness, a user's future sentiment is predicted using her/his prior opinion and the opinions from her/his neighbors according to Equation 7.

Table 2: Experimental results

Method	Samsung Galaxy				Xbox				PlayStation			
	Acc	F_Neg	F_Neu	F_Pos	Acc	F_Neg	F_Neu	F_Pos	Acc	F_Neg	F_Neu	F_Pos
Degroot	0.5772	0.1688	0.6857	0.4408	0.4893	0.1801	0.5897	0.4265	0.6115	0.0701	0.7510	0.2023
Flocking	0.5819	0.2414	0.6939	0.3950	0.4407	0.0861	0.5830	0.2298	0.2023	0.1037	0.7332	0.1789
AsLM	0.5481	0.1555	0.6826	0.4597	0.4935	0.1963	0.7233	0.3056	0.5707	0.1903	0.7216	0.3571
PI^2	0.6590	0.2075	**0.7306**	0.6357	0.5694	0.2346	0.6272	0.6002	0.6653	0.1301	0.7813	**0.5136**
SI^2	0.6196	0.0560	0.7138	0.5280	0.5497	0.0670	0.6135	0.4973	0.6380	0.0752	0.7659	0.3474
PIPE_DI^2	0.6542	0.2452	0.7214	0.6252	0.5447	0.2698	0.5954	0.5812	0.6477	0.1586	0.7686	0.4737
DI^2	**0.6660**	**0.2874**	**0.7304**	**0.6450**	**0.5823**	**0.3017**	0.5962	**0.6137**	**0.6751**	**0.1957**	**0.7904**	0.5028

Compared methods. Existing opinion influence models often formulate opinions as the summarized opinion scores, but ignore the content information. For comparison purpose, three state-of-the-art value-based **Degroot** model [1], **Flocking** model [7] and **AsLM** model [6] are selected as baselines. We also exam the following three simplified versions of DI^2. **PI^2** is a variation of DI^2 involving the personal identity alone, whereas **SI^2** is another variation of DI^2 considering the social identity alone. Like DI^2, **PIPE_DI^2** investigates the dual identity. However, it employs a pipeline framework, which first clusters users into groups and then learns opinion influence based on the given clustering results.

Metrics and Parameter Setting. For each topic, we organize the tweets of each user in a temporal sequence, and use the first 90% tweets as training data and the remaining 10% as test data. We evaluate performances on four metrics. Accuracy (Acc) measures the percentage of correctly predicted sentiments among all testing instances. For a more detailed analysis, we compare models in terms of F-measure on negative, neutral and positive sentiment separately. The parameters of Degroot, Voter and AsLM are experimentally set for their best performances. For DI^2, we set the dimension of the word embedding d_w as 30, the initial values for the balance weights λ_0 and ω as 0 and 1, and the smooth weight δ as 0.5. We experimentally set the cluster number K as 5, which is same as the number of influence roles proposed in [4]. To make fair comparison, we use the same parameter settings for PI^2, SI^2, PIPE_DI^2 and DI^2.

3.2 Results and Analysis

As reported in Table 2, DI^2 performs the best in almost all evaluation metrics and on three topics. Based on the results, some important findings are concluded as follow:

Content information is important. On average, all content-based models outperform the value-based models. Compared with the summarized opinion scores, the texts included in the messages have a better capability to accurately depict the opinion information for influence modeling.

Dual identity better represents users. The better performance of DI^2 compared against PI^2 and SI^2 demonstrates that dual identities provide a more completed user representation, and the two types of opinion influence derived from dual identities better capture the relationships between user pairs. In the meanwhile, PI^2 outperforms SI^2, which indicates that the personalities of users play a dominant role during the process of opinion formation.

Joint learning framework brings benefits. DI^2 with the joint learning framework is superior to PIPE_DI^2 with the pipeline framework. With the novel joint learning framework, the tasks of opinion influence modeling and social identity detection benefit each other by learning interactively.

DI^2 copes better with the negative sentiment. Compared with the positive and neutral sentiment prediction, the negative sentiment prediction achieves the lowest results. Compared to the second best results, the F-measures of DI^2 on negative sentiment prediction are greatly improved by 17.2%, 11.82%, 23.39% on the topics "Samsung Galaxy", "Xbox" and "PlayStation", respectively. As shown in Table 1, less than 20% tweets express the negative sentiment. The dual identities allow DI^2 to better "understand" negative opinion formation when the communication is insufficient.

4 CONCLUSION

In this paper, we propose DI^2 to introduce user dual identity into opinion influence modeling. Our proposed model has the ability to learn opinion influence and detect social identities under a joint learning framework. By integrating both personal-based and group-based influence, DI^2 outperforms other compared opinion influence models when predicting the sentiment of users' future opinions.

ACKNOWLEDGMENTS

The work in this paper was supported by Research Grants Council of Hong Kong (PolyU 5202/12E, PolyU 152094/14E), National Natural Science Foundation of China (61272291) and The Hong Kong Polytechnic University (4-BCB5, G-YBJP).

REFERENCES

[1] Daron Acemoglu and Asuman Ozdaglar. 2011. Opinion dynamics and learning in social networks. *Dynamic Games and Applications* 1, 1 (2011), 3–49.

[2] Yoshua Bengio, Réjean Ducharme, Pascal Vincent, and Christian Jauvin. 2003. A neural probabilistic language model. *Journal of machine learning research* 3, Feb (2003), 1137–1155.

[3] David Bindel, Jon Kleinberg, and Sigal Oren. 2015. How bad is forming your own opinion? *Games and Economic Behavior* 92 (2015), 248–265.

[4] Chengyao Chen, Dehong Gao, Wenjie Li, and Yuexian Hou. 2014. Inferring topic-dependent influence roles of Twitter users. In *Proceedings of the 37th international ACM SIGIR conference on Research & development in information retrieval*. ACM, 1203–1206.

[5] Chengyao Chen, Zhitao Wang, Yu Lei, and Wenjie Li. 2016. Content-based influence modeling for opinion behavior prediction. In *Proceedings of the 26th International Conference on Computational Linguistics*. 2207–2216.

[6] Abir De, Sourangshu Bhattacharya, Parantapa Bhattacharya, Niloy Ganguly, and Soumen Chakrabarti. 2014. Learning a Linear Influence Model from Transient Opinion Dynamics. In *Proceedings of the 23rd ACM International Conference on Conference on Information and Knowledge Management*. ACM, 401–410.

[7] Rainer Hegselmann and Ulrich Krause. 2002. Opinion dynamics and bounded confidence models, analysis, and simulation. *Journal of Artificial Societies and Social Simulation* 5, 3 (2002).

[8] Yang Yang, Jie Tang, Cane Wing-ki Leung, Yizhou Sun, Qicong Chen, Juanzi Li, and Qiang Yang. 2015. RAIN: Social Role-Aware Information Diffusion.. In *AAAI*. 367–373.