Identifying Peer Influence in Online Social Networks Using Transfer Entropy

Saike He[1], Xiaolong Zheng[1], Daniel Zeng[1,2], Kainan Cui[3], Zhu Zhang[1], and Chuan Luo[1]

[1] The State Key Laboratory of Management and Control for Complex Systems, Institute of Automation, Chinese Academy of Sciences, Beijing, 100190, China

[2] Department of Management Information System, University of Arizona, Tucson 85719, Arizona, USA

[3] The School of Electronic and Information Engineering, Xi'an Jiaotong University, Xi'an, China

{saike.he, xiaolong.zheng, dajun.zeng, zhu.zhang, chuan.luo}@ia.ac.cn, kainan.cui@live.cn

Abstract. The past few years have witnessed the rapid growth of online social networks, which have become important hubs of social activity and conduits of information. Identifying social influence in these newly emerging platforms can provide us with significant insights to better understand the interaction behaviors among online users. However, it is difficult for us to measure the influence quantitatively among user peers, since many key factors such as homophily and heterogeneity, can not be observed in our real world conveniently. More recent work mainly focuses on developing theoretical models based on explicit causal knowledge. Nevertheless, such knowledge is usually not available and often needs to be discovered. In this paper, we introduce a model free approach to formulate causal inferences of behaviors among user peers. Experimental results show that influence measured by our approach could successfully reconstruct the underlying networks structure. Furthermore, two additional case studies based on this approach reveal that influentials wield power through specific venues, which constitute a comparatively small portion of the whole channels.

Keywords: Transfer entropy, causal inference, peer influence

1 Introduction

Online social networks have become the hub of information generation and contagion. The ever-increasing amount of information flowing through online social networks forces the participants of these networks to struggle for attention and influence through social messaging [1, 2], adoption of political opinions and technologies [3, 4]. Consequently, identifying influential users among them and quantify their influence becomes an important problem with applications in viral marketing [5], information dissemination [6, 7], search [8], security informatics discovery [9], and influence standings prediction [10].

However, identifying peer influence in online social networks is challenging because of several confounding factors, such as homophily [11], unobserved heterogeneity [12], simultaneity [13], time-varying factors [14], and other contextual effects [15]. Correlated outcome patterns in homophily may lead to upward bias estimations of influence by involving plausible causal influence. Population heterogeneity may also confound casual inference. Simultaneity emerges when the superficially related outcomes of peer users occur within the same time interval randomly. Time-related factors may vary peer influence measured at different timestamps. As for contextual effects, users' intension to behave varies with the exogenous traits of his neighbors to some extent.

Unfortunately, most existing work is pale to solve these problems well. Roughly speaking, recent work targeting at the aforementioned confounding factors mainly used structural measures and dynamic measures. For structural measures, one particularly salient characteristic is their heavy dependence on the assumed network structure. However, Cha et al. suggest that structural measures alone reveals very little about the influence of a user [16]. This point is further supported by the weak correlation between popularity and influence uncovered by Romero et al. [1]. In addition, structural measures are relatively unreliable, since the ranking of the most influential users differs depending on the measure [17-19]. As an advance over structural measures, researchers attempt to introduce dynamic information through epidemiology simulation [20, 21], information cascades [18], and Influence-Passivity conception [1]. One serious drawback of such work is the requirement of explicit causal knowledge, which is highly scarce in many scenarios.

To obtain causal knowledge, a new line of research has examined casual relationships [22]. According to this research essence, if we could discover causal relationships in user peers, we then identify influence based on their capacity to predict the behavior of other users. However, identifying causal relationships covers only part of dynamics in social media — a large part of its participants' activity is internally generated. Actually, for internally generated dynamics, it is hard to infer causal relationships, since a comprehensive observation of specific user's behavior is extremely difficult.

Fortunately, information theoretic techniques provide an ideal basis to accurately formulate causal inference in a model free manner. In this paper, we introduce a model free approach, named as transfer entropy (TE) or interchangeably information transfer, to identify peer influence in online social networks. Transfer entropy is originally formulated by Saito and Harashima [23], and further developed by Kamitake et al. [24]. Since then, transfer entropy has been widely used to study causal relationships in neuroscience, such as complex nonlinear behavior analysis [25], influence of intelligent agents over their environments [26], and inducing emergent neural structure [27]. Later, this approach was introduced to measure influence in social media by Ver and Galstyan [28], where no specific modeling for dyadic interaction are requested as before. From the perspective of information theory, transfer entropy can be viewed as a nonlinear generalization of Granger causality (GC) [29], and surpasses GC and other model based approaches because of its sensitivity to all order correlations. This is particularly useful for unknown non-linear interactions. As it is inherently asymmetric, transfer entropy incorporates directional and dynamical information based on transition probability. This is a key advantage over mutual information measure [30]. Also interestingly, transfer entropy can be reformulated as a conditional mutual information [31, 32]. This brings a convenient way for calculating, and will be shown in section 3 afterwards.

In brief, our work in this paper contributes in two folds.

(1) We propose a new estimator for transfer entropy. Choice of this estimator is critical to the final performance of casual relationship inference. In contrast to existing statistical based work, our estimator is data-driven and could give higher accuracy in entropy estimation.

(2) By applying our approach d in two case studies, we conduct quantitative analysis for the pattern of wielding venue for peer influence.

The rest of the paper is structured as follows. Section 2 reviews related work in influence identification. In section 3, we present the architectural design and detailed technical information of our influence identification framework. Section 4 reports the results of our evaluation study. Section 5 concludes this paper with a summary.

2 Related Work

To deal with the confounding factors in influence identification, there has been a long history of extensive study on information diffusion in general, and the attributes and roles of influencers specifically [18, 33]. Irrespective of various approaches used, previous work can be broadly categorized into two families: structural measures and dynamic measures.

For structural measures, researchers have tried numerous topological characteristics such as in/out-degree [34], number of followers [16, 19, 35, 36], and so on. As this line of approaches heavily depend on the underlying network structure assumed, this may cause drawbacks for three reasons. First, the underlying networks assumed is empirical to some extent, and could reveal little information about the actual social dynamics. This point is well exemplified by the classic link farming problem [37-39]. Secondly, it is challenging to select the most influential nodes based on network structure, which proves to be NP-hard [5]. What is worse, when a graph is updated, the measurement score need to be recomputed from scratch. Finally, the ranking of the most influential users differ depending on the measure used [16].

In this regard, more recent work tried to involve the dynamic information into influence model. One of the first and most influential work in this direction is proposed by Bass [40]. This work does not explicitly consider the underlying structure of social networks. Rather, it involves the concept of adoption rate among online users. Afterwards, Romero et al. devise an Influence-Passivity algorithm by defining user passivity in a social network and propose an algorithm accordingly to measure peer influence in the network. A limitation inherent in all these approaches, however, is that they require the a priori assumptions of a model to describe the interaction mechanism. Since the required model parameters are usually unknown, there calls for approaches to depict temporally varying causal interactions.

As such, researchers turned to another line of research to examine casual relationships, which traces back to Granger [22]. Various measures of causal relationships exist, and they can be divided into two large classes: those quantify causal relationships based on the information of random variables [41], and those based on specific models of data generation. Methods in the latter class are widely used to study casual relationships in neuroscience, with Granger causality [22] and dynamic causal modeling (DCM, [42]) two aptly paragons in this filed. However, GC requests a linear restriction on the two units under observation, which is not guaranteed in user behavioral dynamics. Whereas DCM assumes a bilinear state space model (BSSM), which could cover non-linear interactions. As a cost, DCM requires a priori knowledge about the network of connectivity under investigation. This scenario requires a potential new method to be as model-free as possible, and naturally leads to the application of information theoretic techniques.

3 Methodology

While Wiener defines causal dependencies resting on an increase of prediction power [43], a causality measure can be naturally expressed in terms of information theoretic concepts by associating prediction enhancement with uncertainty reduction.

Essentially, transfer entropy is a rigorous derivation of a Wiener causal measure within the information theoretic framework [41]. It generalizes the mutual entropy measure, with even fewer formal restrictions. In the following section, we will present detail descriptions about this approach.

3.1 Definition of Transfer Entropy

Theoretically, transfer entropy specifies the directed information flow between two signal sources using their multivariate nonparametric signal statistics. In light of this backdrop, two stochastic processes $X = \mathrm{x_t}$ and $Y = \mathrm{y_t}$ can be approximated by Markov processes according to the generalized Markov condition:

$$\mathrm{p}(\mathrm{y_{t+1}}|\mathbf{y_t^n}, \mathbf{x_t^m}) = \mathrm{p}(\mathrm{y_{t+1}}|\mathbf{y_t^n}) \quad (1)$$

where $\mathbf{x_t^m} = (x_t, \dots, x_{t-m+1})$, $\mathbf{y_t^n} = (y_t, \dots, y_{t-n+1})$, while m and n are the orders (memory) of the Markov process X and Y, respectively. Eq. (1) is fully satisfied when the transition probabilities or dynamics of Y is independent of the past of X, this is in the absence of causality from X to Y. To measure the departure from this condition (i.e. the presenceof causality), Schreiber [41] used the expected Kullback-Leibler divergence between the two probability distributionsat each side of Eq. (1) to define the transferentropy from X to Y as:

$$TE(X \to Y) = \sum_{\mathrm{y_{t+1}}, \mathbf{y_t^n}, \mathbf{x_t^m}} p(\mathrm{y_{t+1}}, \mathbf{y_t^n}, \mathbf{x_t^m}) \log\left(\frac{p(\mathrm{y_{t+1}}|\mathbf{y_t^n}, \mathbf{x_t^m})}{\mathrm{p}(\mathrm{y_{t+1}}|\mathbf{y_t^n})}\right) \quad (2)$$

Paluš has shown that transfer entropy can be rewritten as a conditional mutual information [31, 32]. Then we can rewrite Eq. (2) as:

$$TE(X \to Y) = -\sum_{\mathrm{y_{t+1}}, \mathbf{y_t^n}} p(\mathrm{y_{t+1}}, \mathbf{y_t^n}) \log\left(\frac{p(\mathrm{y_{t+1}}, \mathbf{y_t^n})}{\mathbf{y_t^n}}\right)$$

$$+\sum_{y_{t+1},\mathbf{y}_t^{\mathbf{n}},\mathbf{x}_t^{\mathbf{m}}} p(y_{t+1},\mathbf{y}_t^{\mathbf{n}},\mathbf{x}_t^{\mathbf{m}})\log\left(\frac{p(y_{t+1},\mathbf{y}_t^{\mathbf{n}},\mathbf{x}_t^{\mathbf{m}})}{\mathbf{y}_t^{\mathbf{n}},\mathbf{x}_t^{\mathbf{m}}}\right) \quad (3)$$

To estimate the entropy given by Eq. (3) based on limited amount of data samples, we tend to use binning method from numerical techniques. To this end, we need to reformulate Eq. (3) as follows. First, we deem the chronological user activities as a stochastic process. And then, for each user (denoted as X) in online social networks, we record the history of his activities into a stochastic process S_X. The activities can be arbitrarily in your interests, e.g. tweet, retweet, reply, or a hybrid of them. For convenience, we adopt compress storage by record the timestamps of specific activity chronologically.

$$S_X = \{t_j : 0 < t_1 < t_2 < \cdots\} \quad (4)$$

As we are limited by finite data to investigate the casual relationship between two stochastic point processes, a binary indicator variable need to be introduced to tell where an event occurred in given time intervals or bins [28]:

$$B_X(a,b) = \begin{cases} 1 & \text{if } \exists\ t_j \in S_X \cap (b,a], \\ 0 & \text{otherwise.} \end{cases} \quad (5)$$

Then, based on the observation of user actions within a long time span T, we define the probabilities with fixed $\delta \in \mathcal{R}$ as:

$$P(B_X(t,t-\delta) = X_t) \equiv \frac{1}{T-\delta}\int_{\delta}^{T} dt\,[B_i(t,t-\delta) = X_t] \quad (6)$$

And a joint probability distribution can be defined similarly over a sequence of adjacent bins:

$$P(B_X(t,t-\delta_0) = X_t, B_X(t-\delta_0,t-\delta_0-\delta_1) = X_{t-1},\cdots) \quad (7)$$

With more succinct representation:

$$X_t^{(t-k)} \equiv \{X_t,\cdots,X_{t-k}\} \quad (8)$$

For two distinct users X and Y, their joint probability can be defined over a common set of bins denote with widths $\delta_0,\ \delta_1,\ \ldots\delta_k$ as $P\left(X_t^{(t-k)}, Y_t^{(t-k)}\right)$.

With all the notations given above, we can now redefine transfer entropy from Eq. (3) to:

$$TE(X \to Y) = H\left(Y_t \middle| Y_{t-1}^{(t-k)}\right) - H\left(Y_t \middle| Y_{t-1}^{(t-k)}, X_{t-1}^{(t-l)}\right) \quad (9)$$

In particular, the behavior of user X is said to cause that of user Y when the future behavior of user Y is better predicted by adding knowledge from the past and present behavior of user X than by using the present and past behavior of Y alone. For the sake of simplicity, we take l=k henceforth. Besides, previous study has shown that the distribution of user response exhibit a long tail [44]. This indicates that more recent bin width should be narrower than older ones. According to our data statistics, we set the width of bins as: $\delta_0 = 1\ \mathrm{min}$, $\delta_1 = 1\ \mathrm{hour}$, and $\delta_k = 2\ \mathrm{hours}$ in the experiment section.

3.2 Computation of Transfer Entropy

Any estimator of the transfer entropy based on limited data will inevitably lead to biases and statistical errors [32, 45]. Sources of bias mainly come from two aspects: systematical deviation and statistical deviation. Systematical deviation can be tackled with randomized experiments, as will be involved in our experimental design in section 4. While for statistical deviation, a myriad of methods are available from the literature of computational neuroscience [46].

Generally speaking, statistical deviation could be eliminated through two ways: ex-ante limitation and ex-post elimination. For the former one, statistical deviation can be restrined at reasonable range with respect to the given data. While ex-post elimination works by first estimating the bias itself and then adjusting final result accordingly. Since the estimation of bias in the latter approach is derived based on some general priori knowledge (e.g. Panzeri-Treves bias estimate [46]), it just works like post hoc remedy. Therefore, we believe ex-ante limitation approach is more data efficient by depicting the subtle characteristics of samples more accurately.

In this work we choose to use Simpson's rule [47] to estimate the integration presented in Eq. (9), and finally for that in calculating transfer entropy. The process of estimating transfer entropy can be formulated as:

$$\int_a^b TE^{'}(t)dt = \int_a^b \left[\frac{(t-\mathrm{c})(t-\mathrm{b})}{(a-\mathrm{c})(a-\mathrm{b})}TE^{'}(a) + \frac{(t-\mathrm{a})(t-\mathrm{b})}{(c-\mathrm{a})(c-\mathrm{b})}TE^{'}(c) + \frac{(t-\mathrm{a})(t-\mathrm{c})}{(b-\mathrm{a})(b-\mathrm{c})}TE^{'}(b)\right]dt \tag{10}$$

$$= \cdots = \frac{b-a}{6}\left[TE^{'}(a) + 4TE^{'}\left(\frac{a+b}{2}\right) + TE^{'}(b)\right]$$

where $TE^{'}(t)$ is the derivation of function $TE(t)$.

In numerical analysis, Simpson's rule corresponds to the 3-point Newton-Cotes quadrature rule [47], and is derived by approximating the integrand by a quadratic interpolant function. Since Simpson's method approximates the function by a "piecewise" quadratic, thus, if a function is already quadratic, then the result is exact. This is particularly suitable for our situation, where Eq. (5) can be considered as a special quadratic function with second order coefficient equals to zero.

4 Experiments and Results

One of the primary goals in this paper is to infer transfer entropy in user peers by analyzing their patterns of activities. Though many activity patterns can be used to calculate transfer entropy [28], in the following experiments, we focus on the recommendation adoption behaviors of users. The propensity of an individual to adopt or reject the items recommended to him is influenced by the behaviors of his neighbors in some way, and his future actions will, in turn, port influence to his neighbors through the process of social contacts, either directly or indirectly. Though such complications are hard to model, we can detect whether there are causal relationships among these behaviors through transfer entropy, and ascribe the outcomes to the effects of influence accordingly.

4.1 Dataset

Tencent Weibo[1] is a Twitter-like microblogging system in China provided by Tencent, one of the largest Chinese Internet content providers. The dataset used in the following experiments (hereafter referred to as Tencent dataset) is from KDD Cup2012[2], a sampled snapshot of Tencent Weibo users' preferences for various items — items recommended to users and the history of their 'following' history. Millions of users in volume together with rich information in multiple domains (such as user behaviors, social graph, item categories), makes it an ideal resource to study social influence.

4.2 Experimental Design

As mentioned before, estimation of transfer entropy is confounded by several sources of factors. To obtain an unbiased estimation, we design randomized trial to minimize the potential effects caused by such factors, as described below.

Our method for randomization is effective yet easy to implement. We randomly select adoption behaviors with corresponding timestamps from the whole dataset. This procedure is imperative, and brings benefit in three ways. First, without any prior knowledge of how the data are collected, this manipulation can alleviate the affection caused by selection bias. Secondly, we guarantee that the sampled data used are representative enough for the whole volume. As users may be clustered in local network that differ in important ways from users in other counterpart of the whole network, randomization can guarantee that characteristics of those users under experiment present no statistical differences from others. Thirdly, the Tencent dataset suffers from information incompleteness, and this phenomenon does not occur rigorously at random. Randomized trial could mitigate the affections induced by this factor statistically. Consequently, we obtain a representative sample of 920,110 Tencent users, which constitutes 10% of the total population.

In the following, we first evaluate the performance of transfer entropy in recovering the underlying network structure. We then use transfer entropy in two case studies, namely how do online users wield influence through two different information venues: keywords and topics.

---

[1]http://t.qq.com/

[2]Official website: http://www.kddcup2012.org/

4.3 Performance of Transfer Entropy

Network structure recovering is a challenging task in the literature of social dynamics. The key in this task consists in the choice of criteria for establishing an edge within user peers. In our implementation, we consider there is an edge from user X to user Y if $TE(X \rightarrow Y) \geq T_0$, where T_0 is a predefined threshold. By taking true positive rate (TPR) and false positive rate (FPR) as a function of T_0, we can draw the ROC curve, as showing in Fig. 1.

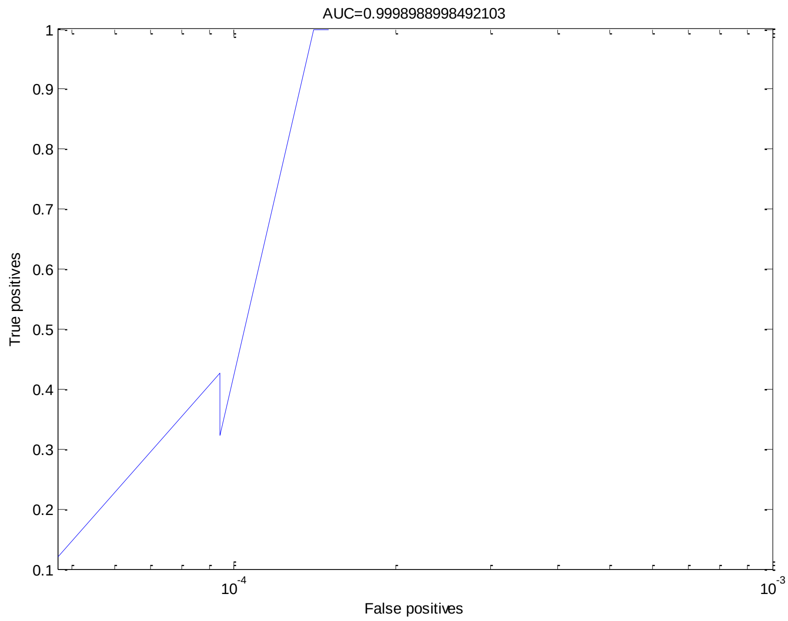

Fig. 1. ROC curve for transfer entropy.

To balance both the fitting ability and computational cost, we set the maximum estimator error in transfer entropy to be 0.01. Experiment result shown in Fig. 1 reveals that transfer entropy can reconstruct the network structure well. In addition, the FPR is restricted within as narrow interval [0, 0.00015] where TPR varies from 0 to 1.0. This characteristic allows us to adjust TPR without prohibitively drop in FPR, and is vital for situations wherein rigorous fitting over the underlying graph structure is not always optimal.

4.4 Case studies

Users in online social networks compete for attention and spread influence mainly through messaging. Here, we are highly intriguing in the patterns of how users wield influence through two kinds of information venues: keywords and topics.

4.4.1 Influence concentration on keywords

Messages are the medium for users to wield their influence, and we believe influence is mainly conveyed by the keywords used in the content. We are interested in the pattern of user's influence assigned over these keywords: do they have any preference in choosing keywords when wielding influence, and to which degree?

To answer these questions, we first calculate each user's influence by cumulating all the influence wielded on his neighbors. Without any further information, we make a rough yet reasonable assumption that the influence of an individual distributes evenly among all the keywords he used. By summarizing all the sub-portions of influence exerted by all users through the keyword, we can draw the influence distribution among all the keywords, as shown in Fig. 2.

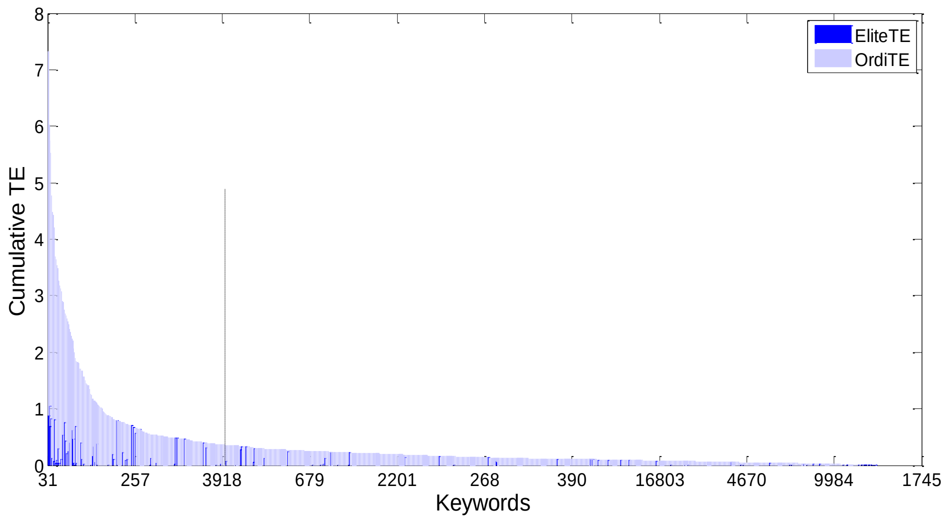

Fig. 2. Distribution of users' influence (measured by transfer entropy) over keywords. Left side of the black dash line corresponds to 80% of the total influence. 'Elite TE' and 'Ordi TE' represents influence of elite and ordinary users respectively.

Fig. 2 reveals that users tilted specific keywords to spread influence. Specifically, they wield 80 percent of their influence merely through 20 percent of the total keywords. In addition, we also differentiate elite users from ordinary ones. By elite, we means online users whose neighbors ranks top 1% of the target volume considered [48]. Therefore, though elite users are optimal vehicles for disseminating information [49], the most cost-effective performance can also be can be realized using "ordinary influencers" [18]. This provides beneficial advice for online campaigns competing aiming at catalyzing the diffusion of opinions, behaviors, innovations, and products in society [50, 51].

As we find that users generally exert influence through a small amount of keywords. Specifically, for influential users, we tend to further check the degree to which they concentrate influence on keywords. To this end, we first need to define a criterion to select influential users. So, we predefine a threshold 'InfluRatio', if a user's influence ranks among top InfluRatio of the whole population, we deem him as influential. We then define the degree of concentration as the partition of keywords used by these influential users to those used by all users. Finally, the concentration degree defined in this way can be formulated as a function of the parameter InfluRatio, and its covariance with InfluRatio are shown in Figure 3.

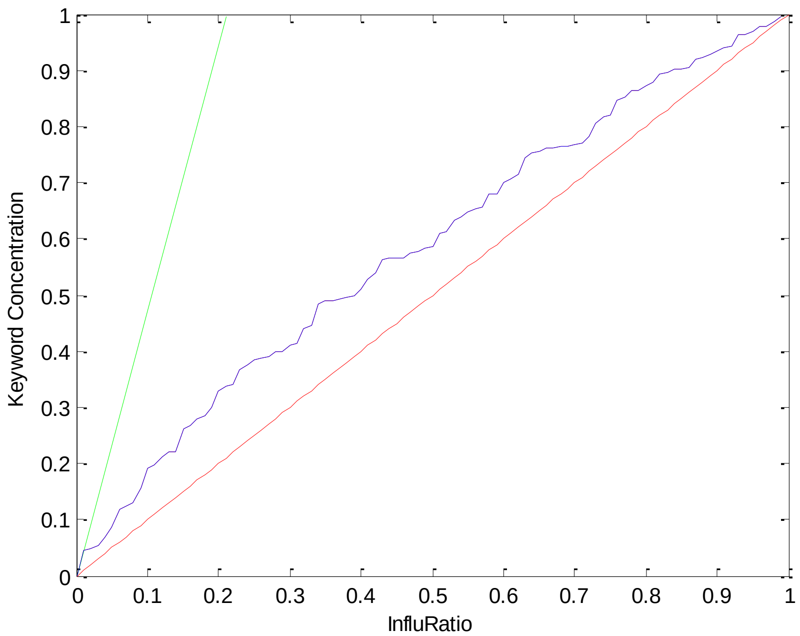

Fig. 3. Concentration of influential user's influence on keywords (the blue curve). Red line corresponds to function y=x, while green line corresponds to y=4.735x (x and y represent InfluRatio and Keyword Concentration respectively). 'Elite TE' and 'Ordi TE' represents influence of elite and ordinary users respectively.

If influential users wield influence evenly over keywords without any concentration, the concentration degree should increase linearly as InfluRatio becomes larger (red line in Fig. 3). If this is the case, considering the average keywords used by each user (about 4.735), the proportion of total keywords used by all influential users will be reformulated as the green line in Fig. 3. Actually, the actual concentration curve lies below the green line and close to the red line in Fig. 3.

This indicates a significant degree of influence concentration on keywords. As the variate InfluRatio becomes larger, keywords concentration begins to decentralize, yet still hold concentration to some extent.

4.4.2 Influence concentration on topics

Usually, keywords are not used solely. More often, they are used together to express specific topics. In this situation, the influence conveyed by each keyword is mixed up as a united output of each topic. Though related topics for each user are not available directly in Tencent Weibo, user tags can partly reflect topics one concerns.

Under this backdrop, we compute the influence distribution among various topics, and the result is shown in Fig. 4.

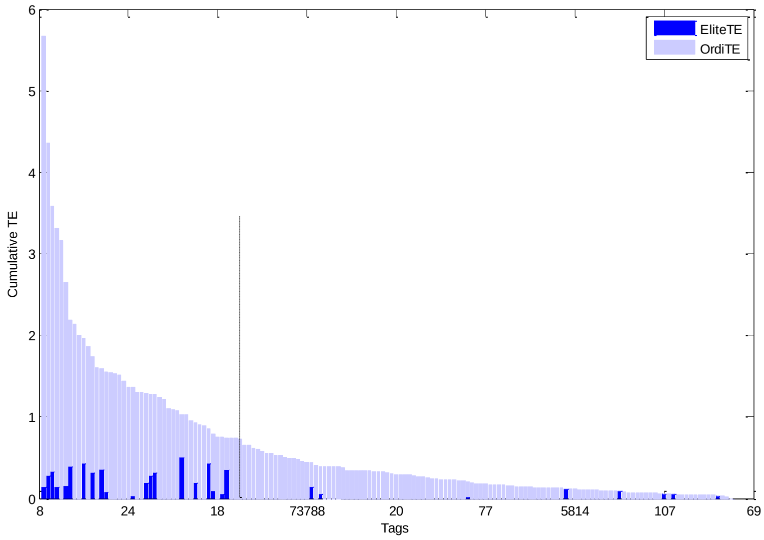

Fig. 4. Distribution of users' influence among topics. Left side of the black dash line corresponds to 70% of the total influence.

From Fig. 4, we find that users tend to wield influence through a small amount of topics, and our statistics suggests that 30% users wield over 70% of the total influence. In more detail, the most intensive concentration is achieved when the ratio of influential users is 0.01. Intriguingly, this coincides with the exact ratio of Twitter users who produce 50% of its content [48]. Further, we want to get a close scrutiny of the concentration degree of influential users on topics, and the result is presented in Fig. 5.

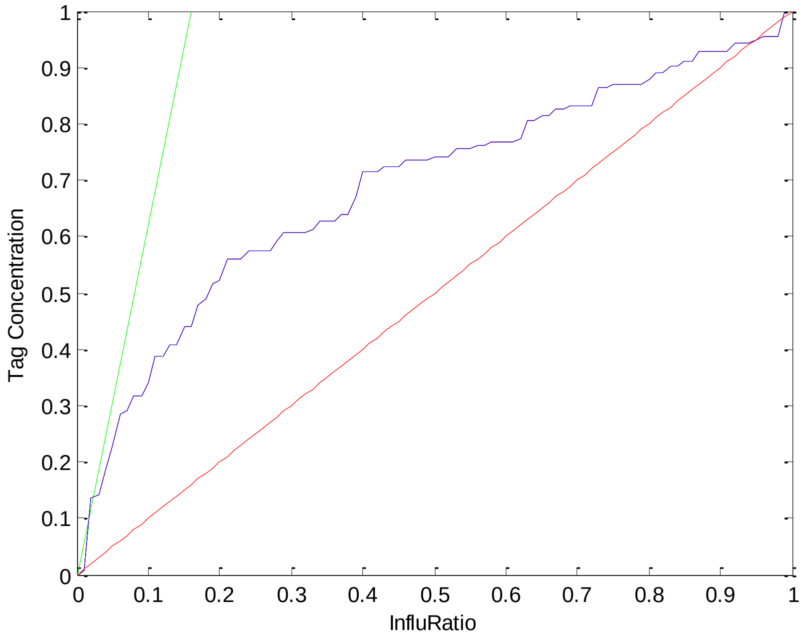

Fig. 5. Concentration of user's influence over tags (the blue curve). Red line corresponds to function y=x, while green line corresponds to y=6.2x (x and y represent InfluRatio and Tag Concentration respectively).

Influential users also wield influence intensively over topics as shown in Fig. 5. Comparing Fig. 5 with Fig. 3, the concentration degree on topics is less than that on the keywords. This may be explained by the fact that users averagely concern topics more (6.2 topics per user) while message less with other users.

5 Conclusions and future work

In this paper, we introduce a model free approach to identify influence in online social networks. With our refinement in entropy estimation, this approach can successfully uncover the underlying network structure. During the adjustment for fitting accuracy, we can improve true positive rate without prohibitively drop in false positive rate. When applying his approach to two case studies, we find that users have preference to concentrate their influence in specific information venues, and the concentration degree is higher in keywords than that in topics. Our findings can be utilized to leverage analytics in security informatics, viral marketing, public opinion, among others.

Our future work lies in four aspects. First, to justify our approach's effectiveness in other disciplines, we intend to make a lateral comparison of its performance with more datasets. Secondly, considering the high computation cost, we desire to prose more efficient algorithms for entropy estimation. Thirdly, we wonder how other types of user behavior could influence the final result, which is temporally uncovered in this paper. Finally, we want to further check whether involving contextual information could leverage influence identification.

Acknowledgments

This work was supported in part by the following grants: NNSFC grants #91124001, #71103180,#71025001, and #91024030, and MOH grants #2013ZX10004218 and #2012ZX10004801.

References

1. Romero, D., Galuba, W., Asur, S., Huberman, B.: Influence and passivity in social media. Machine Learning and Knowledge Discovery in Databases 18-33 (2011)
2. Zheng, X., Li, H., Sun, A.: Exploring social dynamics in online bookmarking systems. Intelligence and Security Informatics, pp. 390-391. Springer (2008)
3. Granovetter, M.: Threshold models of collective behavior. American journal of sociology 1420-1443 (1978)
4. Watts, D.J., Strogatz, S.H.: Collective dynamics of `small-world' networks. Nature 393, 440-442 (1998)
5. Kempe, D., Kleinberg, J., Tardos, É.: Maximizing the spread of influence through a social network. Proceedings of the ninth ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 137-146. ACM (2003)
6. Gruhl, D., Guha, R., Liben-Nowell, D., Tomkins, A.: Information diffusion through blogspace. Proceedings of the 13th international conference on World Wide Web, pp. 491-501. ACM (2004)
7. Leskovec, J., Adamic, L.A., Huberman, B.A.: The dynamics of viral marketing. ACM Transactions on the Web (TWEB) 1, 5 (2007)
8. Adamic, L., Adar, E.: How to search a social network. Social Networks 27, 187-203 (2005)
9. Glass, K., Colbaugh, R.: Web analytics for security informatics. Intelligence and Security Informatics Conference (EISIC), 2011 European, pp. 214-219. IEEE (2011)
10. Ghosh, R., Lerman, K.: Predicting influential users in online social networks. arXiv preprint arXiv:1005.4882 (2010)
11. Aral, S., Muchnik, L., Sundararajan, A.: Distinguishing influence-based contagion from homophily-driven diffusion in dynamic networks. Proceedings of the National Academy of Sciences 106, 21544-21549 (2009)
12. Van den Bulte, C., Lilien, G.L.: Medical Innovation Revisited: Social Contagion versus Marketing Effort1. American journal of sociology 106, 1409-1435 (2001)
13. Godes, D., Mayzlin, D.: Using online conversations to study word-of-mouth communication. Marketing Science 23, 545-560 (2004)
14. Bemmaor, A.C.: Modeling the diffusion of new durable goods: Word-of-mouth effect versus consumer heterogeneity. Research traditions in marketing, pp. 201-229. Springer (1994)
15. Manski, C.F.: Identification of endogenous social effects: The reflection problem. The review of economic studies 60, 531-542 (1993)

16. Cha, M., Haddadi, H., Benevenuto, F., Gummadi, K.P.: Measuring user influence in twitter: The million follower fallacy. 4th international aaai conference on weblogs and social media (icwsm), vol. 14, pp. 8 (2010)
17. Wu, S., Hofman, J.M., Mason, W.A., Watts, D.J.: Who says what to whom on twitter. Proceedings of the 20th international conference on World wide web, pp. 705-714. ACM (2011)
18. Bakshy, E., Hofman, J.M., Mason, W.A., Watts, D.J.: Everyone's an influencer: quantifying influence on twitter. Proceedings of the fourth ACM international conference on Web search and data mining, pp. 65-74. ACM (2011)
19. Kwak, H., Lee, C., Park, H., Moon, S.: What is Twitter, a social network or a news media? Proceedings of the 19th international conference on World wide web, pp. 591-600. ACM (2010)
20. Bailey, N.T.: The mathematical theory of infectious diseases and its applications. Charles Griffin & Company Ltd, 5a Crendon Street, High Wycombe, Bucks HP13 6LE. (1975)
21. Anderson, R.M., May, R.M., Anderson, B.: Infectious diseases of humans: dynamics and control. Wiley Online Library (1992)
22. Granger, C.W.: Investigating causal relations by econometric models and cross-spectral methods. Econometrica: Journal of the Econometric Society 424-438 (1969)
23. Saito, Y., Harashima, H.: Tracking of Information within Multichannel {EEG} record Causal analysis in {EEG}. Yamaguchi N, Fujisawa K (eds) Recent advances in {EEG} and {EMG} data processing. Elsevier 133-146 (1981)
24. Kamitake, T., Harashima, H., Miyakawa, H.: A time-series analysis method based on the directed transformation. Electronics and Communications in Japan 67, 1-9 (1984)
25. Varn, D.P., Crutchfield, J.P.: From finite to infinite range order via annealing: The causal architecture of deformation faulting in annealed close-packed crystals. Physics Letters A 324, 299-307 (2004)
26. Klyubin, A., Polani, D., Nehaniv, C.: All else being equal be empowered. Advances in Artificial Life 744-753 (2005)
27. Lungarella, M., Sporns, O.: Mapping information flow in sensorimotor networks. PLoS Computational Biology 2, e144 (2006)
28. Ver Steeg, G., Galstyan, A.: Information transfer in social media. Proceedings of the 21st international conference on World Wide Web, pp. 509-518. ACM (2012)
29. Barnett, L., Barrett, A.B., Seth, A.K.: Granger causality and transfer entropy are equivalent for Gaussian variables. Physical review letters 103, 238701 (2009)
30. Pillai, S.U.: Probability, Random Variables, and Stochastic Processes. Tata McGraw-Hill Education (2002)
31. Paluš, M., Komárek, V., Hrnčíř, Z., Štěrbová, K.: Synchronization as adjustment of information rates: detection from bivariate time series. Physical Review E 63, 046211 (2001)
32. Hlaváčková-Schindler, K., Paluš, M., Vejmelka, M., Bhattacharya, J.: Causality detection based on information-theoretic approaches in time series analysis. Physics Reports 441, 1-46 (2007)
33. Zheng, X., Zhong, Y., Zeng, D., Wang, F.-Y.: Social influence and spread dynamics in social networks. Frontiers of Computer Science 6(5), 611-620
34. Ohara, K., Saito, K., Kimura, M., Motoda, H.: Effect of in/out-degree correlation on influence degree of two contrasting information diffusion models. Social Computing, Behavioral-Cultural Modeling and Prediction 131-138 (2012)
35. Page, L., Brin, S., Motwani, R., Winograd, T.: The PageRank citation ranking: bringing order to the web. (1999)
36. Jeh, G., Widom, J.: Scaling personalized web search. Proceedings of the 12th international conference on World Wide Web, pp. 271-279. ACM (2003)
37. Wu, B., Davison, B.D.: Identifying link farm spam pages. Special interest tracks and posters of the 14th international conference on World Wide Web, pp. 820-829 (2005)
38. Gyöngyi, Z., Garcia-Molina, H.: Link spam alliances. Proceedings of the 31st international conference on Very large data bases, pp. 517-528. VLDB Endowment (2005)

39. Castillo, C., Donato, D., Gionis, A., Murdock, V., Silvestri, F.: Know your neighbors: Web spam detection using the web topology. Proceedings of the 30th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 423-430. ACM (2007)
40. Bass, F.M.: A New Product Growth for Model Consumer Durables. MANAGEMENT SCIENCE 15, 215-227 (1969)
41. Schreiber, T.: Measuring information transfer. Physical review letters 85, 461-464 (2000)
42. Friston, K.J., Harrison, L., Penny, W.: Dynamic causal modelling. Neuroimage 19, 1273-1302 (2003)
43. Vicente, R., Wibral, M., Lindner, M., Pipa, G.: Transfer entropy—a model-free measure of effective connectivity for the neurosciences. Journal of computational neuroscience 30, 45-67 (2011)
44. Barabasi, A.-L.: The origin of bursts and heavy tails in human dynamics. Nature 435, 207-211 (2005)
45. Kraskov, A., Stögbauer, H., Grassberger, P.: Estimating mutual information. Physical Review E 69, 066138 (2004)
46. Victor, J.D.: Approaches to information-theoretic analysis of neural activity. Biological theory 1, 302-316 (2006)
47. Hildebrand, F.B.: Introduction to numerical analysis. Dover Publications (1987)
48. Hopcroft, J., Lou, T., Tang, J.: Who will follow you back?: reciprocal relationship prediction. Proceedings of the 20th ACM international conference on Information and knowledge management, pp. 1137-1146. ACM (2011)
49. Leavitt, A., Burchard, E., Fisher, D., Gilbert, S.: The influentials: New approaches for analyzing influence on twitter. Web Ecology Project, http://tinyurl. com/lzjlzq 29, (2009)
50. Katz, E.: The two-step flow of communication: An up-to-date report on an hypothesis. Public Opinion Quarterly 21, 61-78 (1957)
51. Valente, T.W.: Network models of the diffusion of innovations. Computational & Mathematical Organization Theory 2, 163-164 (1996)