An Efficient Critical Incident Propagation Model for Social Networks Based on Trust Factor

XiaoMing Li[1,2], Limengzi Yuan[1], ChaoChao Liu[1], Wei Yu[1], Xue Chen[1], and Guangquan Xu[1(✉)]

[1] School of Computer Science and Technology, Tianjin University, Tianjin, China
lxm696@163.com, {ylmz, chaochaoliu, weiyu, xuechen, losin}@tju.edu.cn
[2] School of Information Science and Engineering, Zaozhuang University, Shandong, China

Abstract. Studying patterns of social behavior among users based on micro blogs, QQ posts, and comments is essential to understanding the information propagation process during critical incidents. A common problem of information propagation models based on epidemic dynamics is that they regard the probability of information being propagated successfully across different nodes as a constant. But in real-world scenarios, infection probability varies depending on the trust relationship between people. In this paper, a novel information propagation model for critical incidents is proposed that takes into account the trust factor based on information propagation theory.

Keywords: Trust factor · Cps · Social networks · Incident propagation

1 Introduction

Trust models are classified according to different classification dimensions. For example, the authors in [1] classified existing trust models into six categories. The authors in [2] performed fine-grained classification of trust models, and the categories included BeliefmM [3, 4], a straightforward mathematical formula model [5], a fuzzy theory-based model [6], a Bayesian model [7], and recommendation algorithms based on the relationship and depth of relationship [7–9], where trust factor is computed based on the social relationship. In [10], the authors divided trust into two categories: trust among friends and trust based on similarity. A trust recommendation method based on a reliable path in the social network was presented in [11]. A recommendation-based trust chain model was developed in [12]. This model could inhibit the behavior of dishonest recommendation nodes effectively. However, their work did not rely on the information propagation process that is exhibited across social networks with different trust factors during critical and normal incidents.

The work in this paper intends to address the above limitations. We jointly consider information propagation data during critical and normal incidents, and use a trust-based

I. Romdhani et al. (Eds.): CollaborateCom 2017, LNICST 252, pp. 416–424, 2018.
https://doi.org/10.1007/978-3-030-00916-8_39

information propagation model to indicate the difference in information propagation with varying trust factor in critical and normal incidents.

2 Data Sets

Data from two typical real-world networks (Sina micro blog and QQ Zone) are selected as the data sets for simulations. Sina micro blog has become one of China's most used micro blogs. We collected information propagation records from 576 volunteers through QQ Zone and Sina micro blog during crisis and normal incidents that occurred between Jan. 1, 2008 and Mar. 31, 2013 [13]. Data time stamps are accurate to within a second. For the sake of experiment, posts from public accounts are removed to simplify the data. The Baidu local news service (news.baidu.com) is used to search for news across the country for the time of data sets in critical incidents. Incidents include emergencies like "explosion", "earthquake", and "group incidents", and ordinary incidents like "BRT is put in operation", "festival shows", and "concerts". These words are used as keywords to extract the data. Table 1 shows the information propagation for the two data sets.

Table 1. Data sources

Source	Nodes	Side	Diameter	Clustering coefficient	Density	Weighted degree
Micro blog Sina	958006	6156091	4	0.022	0.001	20.63
QQ zone	820484	16044572	5	0.107	0.01	34.629

3 Modeling

The trust-based SEIR model is used to study information propagation across different social networks during critical and normal incidents. Data from QQ Zone and Weibo constitutes an information propagation network, where each node represents a user in the network, the degree of a node refers to the number of users for the node, and the edge denotes communication between a pair of nodes. We define all nodes to have four states: susceptible (S), infected (I), immune to the propagated message (R), and unknown after infection (E). Depending on the trust factor g, state E is likely to change to R or S [14].

Consider a network with N(t) nodes at time t. Then, we have:

$$\mathrm{S(t)} + \mathrm{E(t)} + \mathrm{I(t)} + \mathrm{R(t)} = \mathrm{N(t)} \tag{1}$$

Assume that node j is at state E at time t. Let P_{es}^{j} denote the probability that node j turns from E to S at time $[t, t + \Delta t]$, and P_{er}^{j} denote the probability that node j turns from E to R at time $[t, t + \Delta t]$. Then, we have $P_{ss}^{j} + P_{se}^{j} = 1$.

According to the definition, a node turns from E to S at a probability P. Then, we have:

$$P_{se}^{j} = \Delta t \mathrm{p}(t) \tag{2}$$

where $\mathrm{P(t)} = g\beta\frac{I(t)}{N(t)}$, g denotes the trust factor that a node turns from E to S or R, β denotes the node degree, $\frac{I(t)}{N(t)}$ denotes the ratio of infected nodes to the total number of nodes in the network at time t, and P(t) is in the range [0, 1].

Similarly, we can obtain node expressions for state E at time $[t, t+\Delta t]$.

$$\begin{aligned} \mathrm{E}(t+\Delta t) &= \mathrm{E}(t) + \mathrm{Gs}(t)p_{se} - E(t)p_{ei} - \mathrm{E}(t)p_{es} \\ &= \mathrm{E(t)} + g\mathrm{S}(t)\Delta\beta\frac{I(t)}{N(t)} - gE(t)\Delta t\varepsilon - gE(t)\Delta t\gamma \end{aligned} \tag{3}$$

Since unknown node E(t) is likely to turn into I(t) or R(t) under the influence of propagation time and trust among friends, we define a function as follows.

The trust factor of node a in its social network L is computed as:

$$g(a,L) = \frac{1}{\sum_{\substack{b\in V(L)\\ b\neq a}} g(b,L)} \sum_{\substack{b\in V(L)\\ b\neq a}} (g(b,L)\times e_{ba}) \tag{4}$$

where g(a, L) denotes the trust factor of node a in social network L, V(L) denotes the set of all nodes in L, and e_{ba} denotes the direct trust factor of b for a. If there is no direct interaction between b and a, then $e_{ba} = 0$.

In a social network with n nodes, g(a, L) can be computed through iterations. The steps are as follows [15]:

Step 1: The trust factor g(a, L) for each node in the social network is set to 0.5.

Step 2: Update g(a, L) of node a in the social network via e_{ba} in $\overline{g(i,R)}$ using Eq. (4):

$$\overline{g(a,L)} = \frac{1}{\sum_{\substack{b\in V(L)\\ b\neq a}} g(b,L)} \sum_{\substack{b\in V(L)\\ b\neq a}} (g(b,L)\times e_{ba}) \tag{5}$$

Step 3: All nodes in the social network can be updated iteratively. The value of $g(a,L)$ converges to $\overline{g(a,L)}$. Hence, $g(a,L) = \overline{g(a,L)}$, which yields the trust factor g:

$$g = s(a,L)|a \in V(r) \tag{6}$$

Based on the above equations, when $\Delta t \to 0$, we have:

$$\mathrm{E}' = \{s(a, L)|a \in V(r)\}\left(\beta \frac{I(t)}{N(t)} S(t) - \varepsilon E(t) - \gamma E(t)\right) \tag{7}$$

where $g \in [0, 1]$. The higher the value of g, the higher the probability that a message is sent to other individuals. This means that the message is spread more quickly to more nodes.

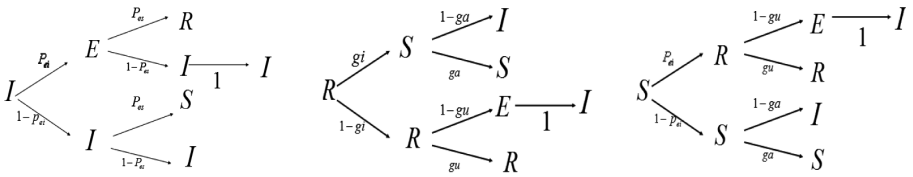

Fig. 1. The probability tree on the graph describes the states of dynamics in different trust networks, i.e. susceptible (S), infected (I), immune to the propagated message (R), and unknown after infection (E). gi is the probability of node infection in the case of high trust degree, ga is the probability of node infection in the case of low trust, and gu is the probability of infected node turning to R due to the influence of the trust degree.

Based on this selected scheme, a person can have four different states, i.e., susceptible (S), infected (I), immune to the propagated message (R), and unknown after infection (E). We suggest using the probability tree, as shown on Fig. 1, to reveal the possible states of the nodes and their changes based on trust. According to the total probabilities of different states given on Fig. 1, the MMCA equation of coupling dynamics in multiplexing can be derived.

$$\begin{aligned} ga(t) &= \prod_j (1 - a_{ji} p_j^I(t) p_{is}) \\ gu(t) &= \prod_j (1 - a_{ji} p_j^I(t) p_{ir}) \end{aligned} \tag{8}$$

Using Eq. 7 we can develop the Microscopic Markov Chains for the coupled processes for each node i as following:

$$\begin{aligned} p_i^R(t+1) &= p_i^I(t) p_{ei}\mu + p_i^R(t) gi(t) gu(t) + p_i^S(t) p_{ei} gu(t) \\ p_i^S(t+1) &= p_i^I(t)(1 - p_{ei}) p_{es} + p_i^R(1 - gi(t)) ga(t) + p_i^S(t)(1 - p_{ei}) ga(t) \\ p_i^I(t+1) &= p_i^I(t)(1 - p_{es}) + p_i^R[(1 - gi(t))(1 - ga(t)) + gi(t)(1 - gu(t))] \\ &\quad + p_i^S(t)[p_{ei}(1 - gu(t)) + (1 - \delta)(1 - ga(t))] \end{aligned} \tag{9}$$

The MMCA can be extended near the critical point, assuming that the probability of the node is the same in the Sina Weibo and QQ zones. Using stationarity we are now in

the position of computing the on set of the epidemic β_c Near the critical point the MMCA can be expanded assuming that the probability of nodes to be infected is $p_i^I = \varepsilon_I \leq 1$. Inserting this in Eq. 8 and we obtain:

$$\sum_j \left[(1-(1-\gamma)i) g_{ji} - \frac{p_{es}}{\beta^{p_{es}}} p_{ei} \right] \varepsilon_j = 0 \tag{10}$$

Even if there are only two different phases in the steady state, for those nodes who correspond to the specific value of trust, they have initial number of inflammatory nodes. Later, as the level of trust rises, the infection level falls back to the normal stage.

4 Simulations and Verification

Simulations are performed to verify the effectiveness of the proposed model. We assume for QQ Zone and Micro blog data that a node is chosen separately as the promulgator to initiate information propagation and that all remaining nodes are unknown nodes.

4.1 Analysis of Time on Information Propagation

The trust factor of the model, g, is set to 0.5 in order to observe the number of the four types of nodes as a function of time. As shown in Fig. 2, due to widespread publicity, the number of susceptible nodes S(t) dwindles quickly in the initial stage for Weibo and QQ Zone. This means that information is spread very quickly across the social network, and that the number of immune nodes R(t) increases quickly in the initial stage until it approaches 1, indicating that all users receive this information. As shown in Fig. 2, a peak occurs earlier in QQ Zone than in Weibo. But a message in Weibo is spread to more nodes than in QQ Zone because QQ Zone relies more so on a circle of acquaintances than Weibo. Most nodes in QQ Zone are friends, family members, or colleagues. Hence, QQ Zone users can spread information more quickly or be immune to the information.

4.2 Influence of Trust Factor on the Promulgating Node

Simulations are conducted on critical and normal incidents across different data sets to explore the influence of trust factor on the promulgating node. The trust parameter is determined by analyzing simulation results obtained with varying trust factors. In order to discuss the influence of trust factor on information propagation, a node is randomly chosen as the source of information propagation. We set the number of people as N = 10,000, $\varepsilon = 0.5$, $\gamma = 0.5$, and p = 0.5. We observe the variation in the number of promulgating nodes with time in each data set. Based on Eqs. (4–9), we analyze the simulation results during critical and normal incidents when the trust factor is 0.2, 0.4, 0.6, and 0.8.

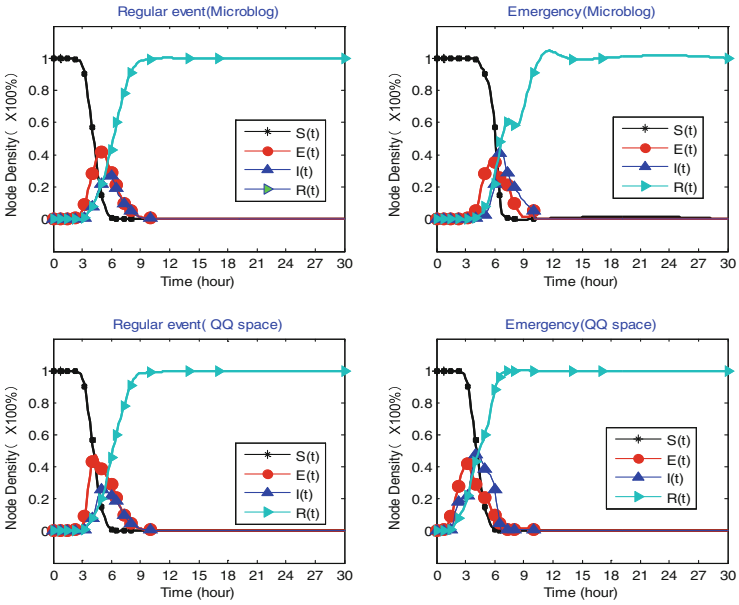

Fig. 2. Black, red, blue, and green lines represent the susceptible node, unknown node, infected node, and immune node, respectively. The number of nodes as a function of time within 30 h, given a trust factor of 0.5 for each node, are shown. (Color figure online)

From Fig. 3, we can see that the propagation range of activities is relatively small when the trust threshold g = 0.4, which indicates that the current activity has no reliable or unreliable nodes. However, when we increase the influence of reliable neighbors, the proportion of active nodes rises rapidly when the threshold of trust equals to 0.8, which indicates that the nodes easily cover the entire propagation area.

Simulations conducted with a trust factor of 0.8 show that, during critical incidents, information propagation via QQ Zone matches the simulation model well, but information propagation via Macro-blog only partially matches the simulation model. This agrees with reality, because during crisis incidents people tend to send a message to old acquaintances and accept a message from whom they trust. An exception is the case of group incidents, when the model matches the two data sets well. This means that after a group incident, if the trust factor reaches a threshold, people tend to follow the herd or put too much trust in who they think is reliable, rather than making their own decision based on information they have. As a result, individuals may be influenced or incited during group incidents. But at some time after the crisis incident, the two data sets show almost the same response and match the simulation well. This is consistent with reality because at some time after the crisis happens, the truth comes out and thus only the true message is accepted. An exception is with communication censorship, where the amount of information propagation in Macro-blog and QQ Zone slumps during crisis incidents.

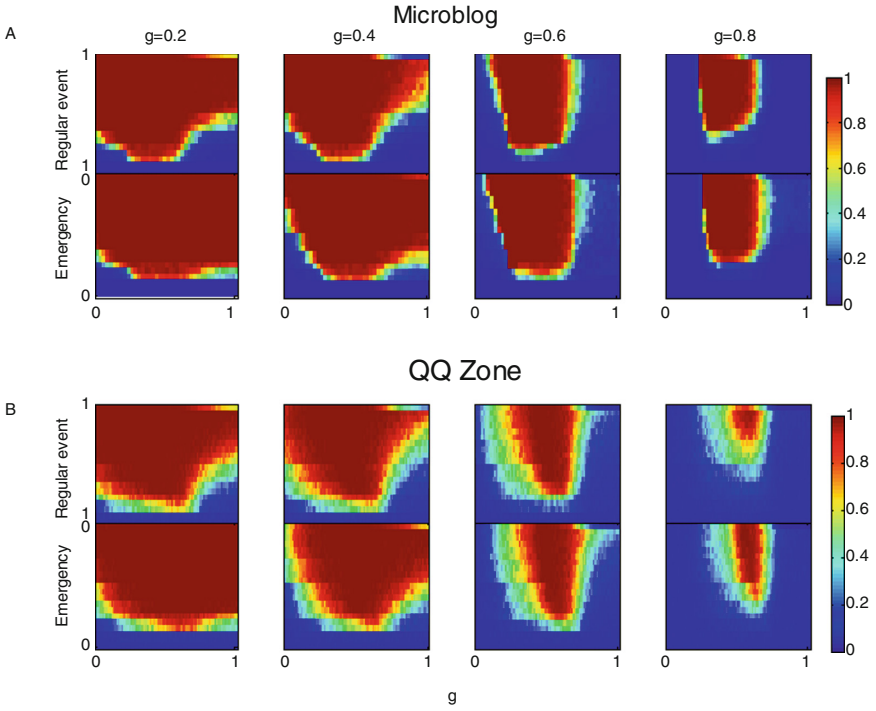

Fig. 3. The figure shows the results of the model SEIR based on the influence of trust factor g in two different social networks of QQ Zone and Micro-blog, when the trust factor is 0.2, 0.4, 0.6, and 0.8. The first line presents the results of simulating routine events in monte carlo, and the second one presents the results of emergency events by the monte carlo simulation.

The simulation model is more consistent with real-world data from the old acquaintance-based QQ Zone. This agrees with reality because during critical incidents, people tend to trust friends rather than strangers.

4.3 Comparison of Trust-Based SEIR and Traditional SEIR

In this subsection, we compare trust-based SEIR and traditional SEIR models. Figure 4 shows a comparison of the two models during critical incidents for a trust factor of 0.8 and also during normal incidents for a trust factor of 0.6. Due to the existence of trust factor in the trust-based SEIR model, information propagation varies in terms of propagation speed and scope with trust factor. Hence, the trust-based SEIR model takes more time to reach stability. In fact, information propagation behavior among people varies with trust factor during critical and normal incidents.

Figure 4 shows information propagation in QQ Zone with the trust-based SEIR model. The promulgating nodes reach the peak earlier than Weibo with a smaller trust factor. This agrees with reality because users of QQ Zone trust information from one another more than users of Weibo. Also, QQ Zone users are more prone to be

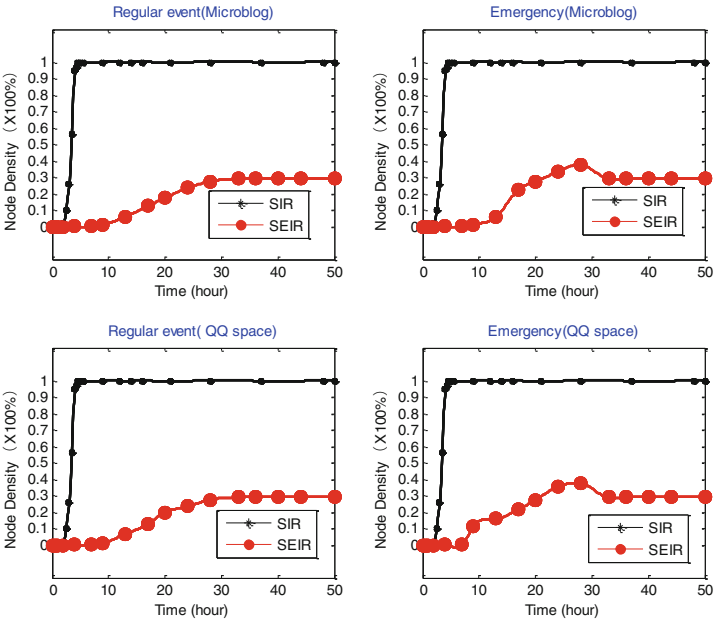

Fig. 4. Black lines represent the variation of promulgating nodes in the SEIR model as a function of time, and red lines represent the variation of promulgating nodes in the trust-based SEIR model as a function of time. (Color figure online)

influenced by other users in the social network. As a result, if one user posts inaccurate information, other friends may respond and comment. In other words, the false information fabricated by one user of the social network may be accepted by many other users in the same network. But, the SEIR model assumes that most nodes accept a certain message without taking trust factor into account. This disagrees with reality. As shown in Fig. 4, the curve of the SEIR model for QQ Zone is similar to that of Weibo for both critical and normal incidents; the promulgating nodes stabilize at the peak level soon. Hence, SEIR is an ideal model, while trust-based SEIR is more consistent with reality.

5 Conclusions and Future Work

A trust-based model of information propagation across a social network is proposed using data sets from QQ Zone and Sina Weibo. The proposed model is helpful for understanding information propagation mechanisms of social networks, exploring key factors that influence information propagation, and ascertaining propagation patterns of social network user relationships during different incidents given different trust factors. Moreover, our work offers guidance on prediction and direction development for diverse real-world incidents, and provides insight into crisis management and public opinion guidance for decision makers.

However, our work has some limitations since we only analyze the influence of trust factor on unknown nodes E during information propagation across a social network. Our ability to collect, process, and mine the data is limited. The amount and scope of data is insufficient to describe reality. Hence, we do not perform an in-depth study on the influence of content, user preferences, and social factors of a social network. These issues will be the focus of future work.

Acknowledgment. This research is partially supported by the Major Project of National Social Science Fund of China (Grant No. 14ZDB153), the National Science Foundation of China (61572355), and Tianjin Research Program of Application Foundation and Advanced Technology under Grant No. 15JCYBJC15700, and the fundamental research of Xinjiang Corps (Grant No. 2016AC015).

References

1. Roshani, F., Naimi, Y.: Effects of degree-biased transmission rate and nonlinear infectivity on rumor spreading in complex social networks. Phys. Rev. E **85**, 036109 (2012)
2. Jøsang, A., Ismail, R., Boyd, C.: A survey of trust and reputation systems for online service provision. Decis. Support Syst. **43**, 618–644 (2007)
3. Jøsang, A.: A logic for uncertain probabilities. Int. J. Uncertain. Fuzz. **9**, 279–311 (2001)
4. Yu, B., Singh, M.P.: An evidential model of distributed reputation management. In: Proceedings of the First International Joint Conference on Autonomous Agents and Multiagent Systems, Bologna, Italy, July 15–19, 2002, pp. 294–301 (2002)
5. Resnick, P., Kuwabara, K., Zeckhauser, R., Friedman, E.: Reputation systems. Commun. ACM **43**, 45–48 (2000)
6. Wang, Y., Vassileva, J.: Bayesian network-based trust model. In: Proceedings of IEEE/WIC International Conference on Web Intelligence, Halifax, NS, Canada, Canada, 27 October 2003, pp. 372–378 (2003)
7. Alves, L.G., Ribeiro, H.V., Lenzi, E.K., Mendes, R.S.: Distance to the scaling law: a useful approach for unveiling relationships between crime and urban metrics. PLoS ONE **8**, e69580 (2013)
8. Song, X., Yan, X.: Influencing factors of emergency information spreading in online social networks: a simulation approach. J. Homel. Secur. Emerg. Manag. **9**, 1–14 (2012)
9. Kenett, D.Y., Portugali, J.: Population movement under extreme events. Proc. Natl. Acad. Sci. USA **109**, 11472–11473 (2012)
10. Lu, X., Bengtsson, L., Holme, P.: Predictability of population displacement after the 2010 Haiti earthquake. Proc. Natl. Acad. Sci. USA **109**, 11576–11581 (2012)
11. Wesolowski, A., et al.: Quantifying the impact of human mobility on malaria. Science **338**, 267–270 (2012)
12. Clauset, A., Gleditsch, K.S.: The developmental dynamics of terrorist organizations. PLoS ONE **7**, e48633 (2012)
13. Fu, Z., Ren, K., Shu, J., Sun, X., Huang, F.: Enabling personalized search over encrypted outsourced data with efficiency improvement. IEEE Trans. Parallel Distr. **27**, 2546–2559 (2016)
14. Ma, T., et al.: Social network and tag sources based augmenting collaborative recommender system. IEICE Trans. Inf. Syst. **98**, 902–910 (2015)
15. Xie, S., Wang, Y.: Construction of tree network with limited delivery latency in homogeneous wireless sensor networks. Wirel. Pers. Commun. **78**, 231–246 (2014)