A Fake News Dissemination Model Based on Updating Reliability and Doubt among Individuals

Kento Yoshikawa*, Takumi Awa†, Risa Kusano‡, Hiroyuki Sato§, Masatsugu Ichino¶, and Hiroshi Yoshiura‖
Graduate School of Informatics and Engineering
University of Electro-Communications
Tokyo, Japan
Email: *k-yoshikawa@uec.ac.jp, †taku.awa@uec.ac.jp, ‡r1.Kusano@uec.ac.jp,
§h.sato@uec.ac.jp, ¶ichino@inf.uec.ac.jp, ‖yoshiura@uec.ac.jp

2020 11th International Conference on Awareness Science and Technology (iCAST) | 978-1-7281-9119-5/20/$31.00 ©2020 IEEE | DOI: 10.1109/ICAST51195.2020.9319485

***Abstract*—As social media has become more widely used, fake news has become an increasingly serious problem. The representative countermeasures against fake news are fake news detection and automated fact-checking. However, these countermeasures are not sufficient because people using social media tend to ignore facts that contradict their current beliefs. Therefore, developing effective countermeasures requires understanding the nature of fake news dissemination. Previous models related to this aim have been proposed for describing and analyzing opinion dissemination among people. However, these models are not adequate because they are based on the assumptions that ignore the presence of fake. That is, they assume that people believe their friends equally without doubting and that reliability among people does not change. In this paper, we propose a model that can better describe the opinion dissemination in the presence of fake news. In our model, each person updates the reliability of and doubt about his or her friends and exchanges opinions among each other. Applying the proposed model to artificial and real-world social networks, we found three clues to analyze the nature of fake news dissemination: 1) people can less accurately perceive that fake news is fake than they can perceive that real news is real. 2) it takes much more time for people to perceive fake news to be fake than to perceive real news to be real. 3) the results of findings 1 and 2 concerning fake news are because people become skeptical about friends in the presence of fake news and therefore people do not update opinions much.**

***Index Terms*—fake news, social network, opinion dissemination, truth discovery**

I. Introduction

Social media (e.g., Facebook, Twitter, and Instagram) have become an increasingly important means of sharing information. They provide easy access to and rapid dissemination of information, enabling people to express their opinions easily. However, the characteristics of social media enable people not only to find beneficial news quickly but also to spread misleading news on a large scale [1]. This misleading news, i.e., ”fake news” can seriously affect both individuals and society [2].

Fake news has led to many incidents. For instance, more than 500 people died of methanol poisoning in Iran because of fake news on social media about such alcohol preventing the spread of COVID-19 [3].

The representative countermeasures against fake news are fake news detection [4]–[7] and automated fact-checking [8]–[10], i.e., verification of the authenticity of the news. However, social media enhance the ”echo chamber effect,” and the ”backfire effect.” The former describes the inherent human tendency of people strengthening their beliefs through interactions with like-minded individuals [11]. The latter describes the inherent human tendency of people denying facts that contradict their beliefs [12]. These two effects degrade the effectiveness of fake news detection and automated fact-checking approaches because facts may be ignored. Therefore, developing effective countermeasures against fake news requires understanding the nature of fake news dissemination.

Previous work on opinion dissemination modeled the problem of people trying to share ”a correct opinion” that is the same as ground truth when friends' opinions might be incorrect [13]. The model was improved by the addition of the Autonomous Adaptive Tuning (AAT) model in which opinion is updated on the basis of the reliability of friends [14]. However, these models assume that people believe their friends equally without doubting and the reliability of friends does not change. Thus, these previous models for opinion dissemination do not adequately describe the behaviors of people in the presence of fake news.

On the other hand, previous work to discover the truth among from conflicting information led to models that enable reliability and doubt to be updated on the basis of friends' opinions [15]–[17]. However, these models only consider the situations that several persons send opinions independently to only one person and that they do not exchange their opinions with each other. Thus, these previous models for discovering the truth do not take people's interactions in social networks into account.

In this paper, we present an opinion dissemination model that describes the behavior of people in a social network under the presence of fake news. Our contributions are summarized as follows.

1) Our proposed model better reflects the opinion dissemination in the presence of fake news than previous models [13]–[17]. It can represent the communication of opinions as well as the update of the reliability of and doubt about friends.
2) Using our proposed model with artificial and real-world social networks as evaluation data, we demonstrate that people can less accurately perceive that fake news is

actually fake than they can perceive that real news is actually real. We also demonstrate that it takes much more time for people to perceive fake news to be fake than to perceive real news to be real.

3) We found that two results about fake news in contribution 2 are due to the following reason: people become skeptical about friends in the presence of fake news and hence people do not update their opinions much.

II. Related Work

A. Analyzing Dissemination of Fake News

Vosoughi, Roy, and Aral analyzed Twitter posts containing fake news [1]. They showed that fake news disseminates significantly faster and more broadly than real news. They found that most posts mentioning fake news include emotional words (e.g., surprise, fear, and disgust) and that the novelty of fake news content causes people to spread fake news significantly faster and more broadly than real news.

Kwon, Cha, and Jung analyzed the spread of fake news through Twitter posts and friendship among users [18]. They found that fake news is spread by a few fake news spreaders who spread fake news repeatedly. They also found that most posts mentioning fake news include hearsay (e.g., "I heard from a friend that ...") or a guess (e.g., "maybe").

It has also been shown that the program "bots," which post messages and share feeds automatically, are widely used to spread fake news [19].

B. Modeling Opinion Dissemination

Glinton, Scerri, and Sycara modeled an opinion sharing problem that people try to share "a correct opinion" that is the same as ground truth when friends' opinions might be incorrect (i.e., friends' opinions might be different from ground truth) [13]. To improve this model, Pryymak, Rogers, and Jennings presented the Autonomous Adaptive Tuning (AAT) model, which describes the update of an opinion on the basis of the reliability of friends [14]. They showed that each person should believe his or her friends' opinions to be correct about 60% to share a correct opinion.

Tsang and Larson modeled the situation in which people with extreme opinions affect people who do not have extreme opinions [20]. They showed that if 20% of people share extreme opinions, the majority of opinions can be swayed to extreme opinions.

In addition, Sasahara et al. modeled the echo chamber effect on social media based on the inherent human tendencies, in which people become more similar to acquaintances through interactions and connect a new person who has similar opinions [21]. They demonstrated that the echo chamber effect is facilitated when people tend to disconnect themselves from people with different opinions.

C. Identifying Truth among from Conflicting Information

A research field called "truth discovery" aims to identify the truth among conflicting information through estimating the reliability of information sources. An early approach for truth discovery is TruthFinder [22], which identifies the truth by iteratively updating the trustworthiness of the information sources and the confidence of facts that the information sources report. This approach uses the correlation among information sources to calculate the trustworthiness and confidence.

Wang, Kaplan, and Abdelzaher outperformed TruthFinder in terms of the accuracy to identify the truth [15]. They used the expectation maximization algorithm, which iteratively updates the reliability of and doubt about friends on the basis of their opinions about whether topics are true.

Furthermore, one work improved the accuracy of the model presented by [15] by incorporating the speed of sharing opinions in the expectation maximization algorithm (e.g., a friend shares an opinion immediately or late) [16]. Another work improved the accuracy of the model presented by [15] by incorporating an emotional-aware situation in the expectation maximization algorithm (e.g., a friend shares an emotional opinion that lacks grounds) [17].

III. Opinion Dissemination Model for Fake News

In this section, we propose an opinion dissemination model considering the presence of fake news, which overcomes the limitations of previous models [13]–[17] while preserving their strengths.

A. Requirements for Opinion Dissemination Model

Most of the previous work for modeling opinion dissemination does not take the truth of opinions into account [20] [21] They only investigate the influence of opinions through modeling opinion dissemination. The opinion dissemination model proposed by Glinton, Scerri, and Sycara [13] is an exception, which considers opinions as the comments about the truth of a topic as well as assuming its ground truth. Pryymak, Rogers, and Jennings improved the precision and scalability of the model of Glinton et al. by developing opinion dissemination model called AAT [14]. Because communicating the truth of opinions in the presence of ground truth is important for analyzing the nature of fake news, we develop our model on the basis of the previous work [13] [14]. However, as mentioned in Section I, the models proposed by [13] [14] have problems: these models assume that people believe their friends equally with doubting and that the reliability of friends does not change.

On the other hand, previous work on "truth discovery" developed methods, in which a person updates the reliability of and doubt about his or her friends on the basis of opinions received from them [15]–[17]. However, these methods do not consider social interaction (i.e., these models only consider the situations that several persons send opinions independently to only one person, and that they do not exchange their opinions with each other).

To address the limitations with previous models [13]–[17], we develop an opinion dissemination model in the presence of fake news that satisfies the following requirements.

1) Each person's opinion in our model is about the fact of the news, where the ground truth of the news is set to the model.
2) Each person updates his or her opinion as well as reliability and doubt of friends on the basis of the friends' opinions.
3) On updating his or her opinion, each person tells the opinion to his or her friends.
4) Every person satisfies requirements 2 and 3. This requirement means that people mutually exchange their opinions and update the reliability of and doubt about each other.

B. Problem Formulation

In our model, the users in set $U = \{u_1, \dots, u_N\}$ are connected by a social network $G(U, E)$ in which E is the set of edges (i.e., friendships among users). Each user u_i ($i = 1, \dots, N$, hereinafter called u_i) has a set of friends (i.e., followers in Twitter) $F_i = \{j : (i, j) \in E\}$, where $|F_i| = M_i$ $(1 \leq M_i \leq N-1)$. This indicates that the number of friends varies from user to user. We assume that each u_i communicates directly with only the other users in F_i.

To model the communication of users' opinions about the fact of the news, we first set $z \in \{\mathrm{True}, \mathrm{False}\}$ as ground truth of the news. If $z = \mathrm{True}$, the news is real news (e.g., The Tokyo 2020 Olympic Games was postponed due to COVID-19 [23]). If $z = \mathrm{False}$, the news is fake news (e.g., alcohol can prevent COVID-19 [3]). We support the assumption that each u_i's opinion o_i is ternary, where $o_i \in \{\mathrm{True}, \mathrm{False}, \mathrm{Undetermined}\}$. If $o_i = \mathrm{True}$, each u_i forms the opinion that the news is real (e.g., each u_i is sure that the news [23] is real news). If $o_i = \mathrm{False}$, each u_i forms the opinion that the news is fake (e.g., each u_i is sure that the news [23] is fake news, but it is real news actually). If $o_i = \mathrm{Undetermined}$, each u_i does not form the opinion about the news (e.g., each u_i is not sure about whether the news [23] is real). The settings in this paragraph satisfy requirement 1.

From the settings, we can denote that if $o_i = z$, each u_i correctly makes sure the fact of the news and that if $o_i \neq z$, each u_i misidentifies the fact of the news.

To decide which three opinions to adopt, each u_i has belief $P_i(z = \mathrm{True}) \in [0, 1]$, which is the probability that each u_i believes that the news is real. We also denote $P_i(z = \mathrm{False}) = 1 - P_i(z = \mathrm{True})$, which is the probability that each u_i believes that the news is fake. Each u_i updates his or her belief on the basis of initial belief P_i^0 and each u_j's ($\{j : j \in F_i\}$) opinion o_j. If $o_j = \mathrm{True}$, each u_i updates current belief $P_i^k(z = \mathrm{True})$ at belief update step k following (1), which is based on Bayes' theorem.

$$
\begin{aligned}
&P_i^{k+1}(z = \mathrm{T} | o_j = \mathrm{T}) \\
&= \frac{P_i^k(o_j = \mathrm{T} | z = \mathrm{T}) P_i^k(z = \mathrm{T})}{P_i^k(o_j = \mathrm{T} | z = \mathrm{T}) P_i^k(z = \mathrm{T}) + P_i^k(o_j = \mathrm{T} | z = \mathrm{F}) P_i^k(z = \mathrm{F})} \\
&= \frac{t_{ij}^k P_i^k(z = \mathrm{T})}{t_{ij}^k P_i^k(z = \mathrm{T}) + f_{ij}^k P_i^k(z = \mathrm{F})} \qquad (1)
\end{aligned}
$$

In this equation, $t_{ij}^k \in [0.0, 1.0]$ is the degree of u_i's updated reliability of each u_j at step k and $f_{ij}^k \in [0.0, 1.0]$ is the degree of u_i's updated doubt about each u_j at step k. These parameters correspond to the degree that each u_i believes or doubts each u_j from 0% to 100%. Every time each user receives an opinion from a friend, he or she updates the reliability and doubt regarding the opinion sharer simultaneously, as described in Section IV. After updating his or her reliability of and doubt about friends, each u_i updates his or her belief. The settings in this paragraph satisfy requirement 2.

Similarly, if $o_j = \mathrm{False}$, each u_i updates current belief $P_i^k(z = \mathrm{True})$ following (2).

$$
\begin{aligned}
&P_i^{k+1}(z = \mathrm{T} | o_j = \mathrm{F}) \\
&= \frac{(1 - t_{ij}^k) P_i^k(z = \mathrm{T})}{(1 - t_{ij}^k) P_i^k(z = \mathrm{T}) + (1 - f_{ij}^k) P_i^k(z = \mathrm{F})} \qquad (2)
\end{aligned}
$$

We also denote (3) and (4) as follows.

$$
P_i^{k+1}(z = \mathrm{F} | o_j = \mathrm{T}) = 1 - (1) \qquad (3)
$$

$$
P_i^{k+1}(z = \mathrm{F} | o_j = \mathrm{F}) = 1 - (2) \qquad (4)
$$

These equations reflect the situation in which users try to update current beliefs (i.e., P_i^k) using the reliability (i.e., t_{ij}^k) of and doubt (i.e., f_{ij}^k) about friends, who share opinions with them and then update their beliefs (i.e., P_i^{k+1}).

On updating his or her belief, each u_i forms opinion o_i^{k+1} at belief update step $(k+1)$ following (5):

$$
o_i^{k+1} = \begin{cases} \mathrm{Undetermined} & \text{if} \quad 1 - \sigma < P_i^{k+1} < \sigma \\ \mathrm{True} & \text{if} \quad P_i^{k+1} \geq \sigma \\ \mathrm{False} & \text{if} \quad P_i^{k+1} \leq 1 - \sigma \\ o_i^k & \quad \text{otherwise} \end{cases} \qquad (5)
$$

where σ is the threshold for users to form the opinion "True" $(0.5 < \sigma < 1.0)$. On the other hand, the threshold $1 - \sigma$ is the criterion for users to form the opinion "False." This threshold indicates that the higher it is, the more careful users form their opinions. If each u_i has formed an opinion $o_i^{k+1} = \{\mathrm{True}, \mathrm{False}\}$, he or she communicates it to all friends in F_i. If $o_i^{k+1} = \mathrm{Undetermined}$, the opinion is not communicated. This setting in this paragraph satisfies requirement 3.

All users in a social network $G(U, E)$ perform reliability and doubt update, opinion update, and opinion communication. This description corresponds to satisfying requirement 4.

IV. Updating Reliability and Doubt

In this section, we develop the model for updating the reliability and doubt formulated in Section III. We achieve this by using the Expectation Maximization (EM) algorithm, a general optimization algorithm for finding maximum likelihood estimates of parameters when the data are incomplete (e.g., part of the data are unobservable because of noises). Briefly, we input an opinion vector, and the algorithm outputs updated reliability and doubt.

A. Definition of Opinion Vector

For the sake of simplicity, we suppose a scenario in which each u_j $(j = 1, \ldots, M_i)$, who is the friends of u_i, tells u_i whether his or her opinion about the news is real or fake. We define one-dimentional opinion vector O_i in which the elements o_j $(j = 1, \ldots, M_i)$ are denoted as shown in (6).

$$o_j = \begin{cases} 1 & \text{if} \quad \text{a } u_j \text{ tells his or her opinion is true.} \\ 0 & \text{if} \quad \text{a } u_j \text{ tells his or her opinion is false.} \end{cases} \tag{6}$$

This vector is used as input to update reliability and doubt using the EM algorithm. In our model, all users in a social network have such a vector.

B. Formulation of Expectation Maximization Algorithm

Our model updates reliability and doubt by using the EM algorithm, which is an optimization scheme to find maximum likelihood estimates of unknown parameters $\boldsymbol{\theta}$ that depend on observed data $\boldsymbol{X}$ and latent (i.e., unobservable) variables $\boldsymbol{Z}$. It iterates following key steps, the ”E-step” and the ”M-step.”

- E-step: Compute the expectation of latent variables $\boldsymbol{Z}$ on the basis of observed data $\boldsymbol{X}$ and parameter $\boldsymbol{\theta}^s$ where s is current parameter update step and $L(\boldsymbol{X}, \boldsymbol{Z}; \boldsymbol{\theta})$ is a likelihood function. In general, (7) is called the Q function.

$$\begin{aligned} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(s)}) =& E_{\boldsymbol{Z}|\boldsymbol{X};\boldsymbol{\theta}^{(s)}}[\log L(\boldsymbol{X}, \boldsymbol{Z}; \boldsymbol{\theta})] \\ =& P(\boldsymbol{Z}|\boldsymbol{X}; \boldsymbol{\theta}^{(s)}) \log L(\boldsymbol{X}, \boldsymbol{Z}; \boldsymbol{\theta}) \end{aligned} \tag{7}$$

- M-step: Find parameters $\boldsymbol{\theta}^{(s+1)}$ that maximize the Q function. These parameters are used in the next E-step.

$$\boldsymbol{\theta}^{(s+1)} = \arg\max_{\boldsymbol{\theta}} Q(\boldsymbol{\theta}, \boldsymbol{\theta}^{(s)}) \tag{8}$$

Now, let us formulate the EM algorithm. We describe the situation in which each u_i updates the reliability of and doubt about all friends in F_i. We define observed data $\boldsymbol{X}$ as opinion vector O_i (i.e., (6)), latent variable $\boldsymbol{Z}$ as the ground truth of the news (i.e., $z \in \{\text{True}, \text{False}\}$), unknown parameters $\boldsymbol{\theta}$ as the reliability of and doubt about each friend (i.e., $\boldsymbol{\theta}$=$(t_{i1}, t_{i2}, \ldots, t_{iM_i}; f_{i1}, f_{i2}, \ldots, f_{iM_i})$). This algorithm can be considered as describing the situation in which each user optimizes reliability and doubt regarding his or her friends (i.e., $\boldsymbol{\theta}$) on the basis of their opinions (i.e., $\boldsymbol{X}$) when the ground truth of the news (i.e., $\boldsymbol{Z}$) cannot be observable. We denote the likelihood function $L(\boldsymbol{X}, \boldsymbol{Z}; \boldsymbol{\theta})$ for updating reliability and doubt as (9).

$$\begin{aligned} &L(\boldsymbol{X}, \boldsymbol{Z}; \boldsymbol{\theta}) \\ &= \prod_{i=1}^{N} \left\{ \prod_{j=1}^{M_i} t_{ij}^{o_j} (1 - t_{ij})^{(1-o_j)} z + \prod_{j=1}^{M_i} f_{ij}^{o_j} (1 - f_{ij})^{(1-o_j)} (1 - z) \right\} \end{aligned} \tag{9}$$

We also define the Q function as (10):

$$\begin{aligned} &Q\left(\boldsymbol{\theta}, \boldsymbol{\theta}^{(s)}\right) = P(\boldsymbol{Z}|\boldsymbol{X}; \boldsymbol{\theta}^{(s)}) \log L(\boldsymbol{X}, \boldsymbol{Z}; \boldsymbol{\theta}) \\ &= \\ &\sum_{i=1}^{N} \left\{ P_i^s(z = \mathrm{T}|o_j; \boldsymbol{\theta}^{(s)}) \sum_{j=1}^{M_i} (o_j \log t_{ij} + (1 - o_j) \log (1 - t_{ij})) \right. \\ &\left. + P_i^s(z = \mathrm{F}|o_j; \boldsymbol{\theta}^{(s)}) \sum_{j=1}^{M_i} (o_j \log f_{ij} + (1 - o_j) \log (1 - f_{ij})) \right\} \end{aligned} \tag{10}$$

where s is the current update step of reliability and doubt, $P_i^s(z = \mathrm{T}|o_j; \boldsymbol{\theta}^{(s)})$ is following (11) or (12)

$$\begin{aligned} &P_i^s(z = \mathrm{T}|o_j = \mathrm{T}) \\ &\quad = \frac{t_{ij}^{(s)} P_i^{s-1}(z = \mathrm{T})}{t_{ij}^{(s)} P_i^{s-1}(z = \mathrm{T}) + f_{ij}^{(s)} P_i^{s-1}(z = \mathrm{F})} \end{aligned} \tag{11}$$

$$\begin{aligned} &P_i^s(z = \mathrm{T}|o_j = \mathrm{F}) \\ &\quad = \frac{(1 - t_{ij}^{(s)}) P_i^{s-1}(z = \mathrm{T})}{(1 - t_{ij}^{(s)}) P_i^{s-1}(z = \mathrm{T}) + (1 - f_{ij}^{(s)}) P_i^{s-1}(z = \mathrm{F})}, \end{aligned} \tag{12}$$

and $P_i^s(z = \mathrm{F}|o_j; \boldsymbol{\theta}^{(s)})$ is following (13) or (14):

$$P_i^s(z = \mathrm{F}|o_j = \mathrm{T}) = 1 - (11) \tag{13}$$

$$P_i^s(z = \mathrm{F}|o_j = \mathrm{F}) = 1 - (12) \tag{14}$$

where $t_{ij}^{(s)}$ and $f_{ij}^{(s)}$ are following (15) and (16).

$$t_{ij}^{(s)} = \prod_{j=1}^{M_i} t_{ij}^{(s)o_j} (1 - t_{ij}^{(s)(1-o_j)}) \tag{15}$$

$$f_{ij}^{(s)} = \prod_{j=1}^{M_i} f_{ij}^{(s)o_j} (1 - f_{ij}^{(s)(1-o_j)}) \tag{16}$$

To optimize reliability and doubt, we apply the M-step to obtain parameters $\hat{\boldsymbol{\theta}} = (\hat{t}_{i1}, \hat{t}_{i2}, \ldots, \hat{t}_{iM_i}; \hat{f}_{i1}, \hat{f}_{i2}, \ldots, \hat{f}_{iM_i})$ that maximize the Q function. To achieve this goal, we solve $\dfrac{\partial Q}{\partial t_{ij}} = 0$, $\dfrac{\partial Q}{\partial f_{ij}} = 0$. However, we cannot define the elements of opinion vector if the opinions of friends are undetermined. To address this problem, we define A_i and $\overline{A_i}$, where A_i is the set of friends who have already formed an opinion, and $\overline{A_i}$ is the set of friends who have not yet formed an opinion. As a result, we can get maximum likelihood estimates of each u_i's reliability $t_{ij}^{(s+1)}$ and doubt $f_{ij}^{(s+1)}$ at step $(s + 1)$ using (17) and (18):

$$t_{ij}^{(s+1)} = \hat{t}_{ij} = \frac{\sum_{j \in A_i} P_i^s(z = \mathrm{T}|o_j; \boldsymbol{\theta}^s)}{\sum_{j=1}^{M_i} P_i^s(z = \mathrm{T}|o_j; \boldsymbol{\theta}^s)} \tag{17}$$

$$f_{ij}^{(s+1)} = \hat{f}_{ij} = \frac{K_i - \sum_{j \in A_i} P_i^s(z = \mathrm{T}|o_j; \boldsymbol{\theta}^s)}{M_i - \sum_{j=1}^{M_i} P_i^s(z = \mathrm{T}|o_j; \boldsymbol{\theta}^s)} \tag{18}$$

where K_i is the number of users in F_i who have already formed an opinion at step s.

The update of reliability and doubt continues until the log-likelihood function $\log L(\boldsymbol{X}, \boldsymbol{Z}; \boldsymbol{\theta})$ converges (i.e., it does not increase by more than a small threshold ϵ).

V. Overall Model

To put together the ideas described in Sections III and IV, we first consider a scenario in which each u_i in a social network performs the following steps 1 through 3 iteratively.

1) Send opinion to all friends in F_i, the set of each u_i's friends.
2) Update u_j's $(j = 1, \ldots, M_i)$ reliability t_{ij} and doubt f_{ij} using friends' opinions and the EM algorithm.
3) Update current belief P_i^k on the basis of updated reliability and doubt.

Fig. 1 illustrates an example of these steps. Firstly, users 1 and 2 send their opinions $o_1^k = \text{True}$ and $o_2^k = \text{True}$ to user 4, respectively, because their beliefs $P_1^k = 0.90$ and $P_2^k = 0.95$ exceeded the threshold $\sigma = 0.80$. Then, user 4 updates the current reliability of and doubt about users 1 and 2 (i.e., t_{41}^{k-1}, f_{41}^{k-1}, t_{42}^{k-1}, and f_{42}^{k-1}) based on their opinions. Finally, user 4 updates current belief $P_4^k = 0.70$ and then sends his or her opinion $o_4^{k+1} = \text{True}$ to users 5 and 7 because the updated belief $P_4^{k+1} = 0.85$ exceeded the threshold $\sigma = 0.80$.

Algorithms 1-3 describe these steps, in which K is the maximum update step of belief, S is the maximum update step of reliability and doubt, and δ is the threshold of accuracy R, which is the ratio that each u_i forms the same opinion o_i as the ground truth of the news $z \in \{\text{True}, \text{False}\}$ (i.e., $R = \frac{\sum_{i=1}^{N} |o_i = z|}{N}$). Algorithm 3 corresponds to the overall steps of our proposed model illustrated in Fig. 1.

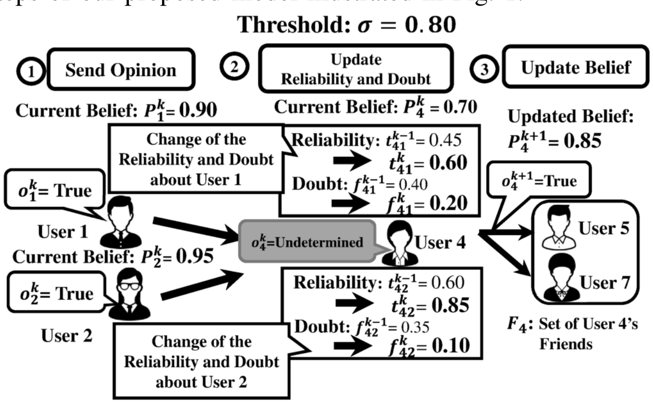

Fig. 1. Overall illustration of proposed opinion dissemination model

Algorithm 1 Reliability and Doubt Update

Input: Empty opinion vector O_i, initial reliability t_{ij}^0, and initial doubt f_{ij}^0

```
Calculate elements of opinion vector O_i using (6)
while log L(X, Z; θ) does not converge or s ≤ S do
    for j = 1 to M_i do
        Update t_ij^s and f_ij^s using O_i, (17), and (18)
    end for
    Set s = s + 1
end while
return Updated reliability t_ij^k and doubt f_ij^k at belief update step k
```

Algorithm 2 Belief Update at Update Step k

Input: Updated reliability t_{ij}^k, and updated doubt f_{ij}^k

```
for j = 1 to M_i do
    if o_j = True then
        if z = True then
            Update current belief P_i^k using (1)
        else if z = False then
            Update current belief P_i^k using (3)
        end if
    end if
    if o_j = False then
        if z = True then
            Update current belief P_i^k using (2)
        else if z = False then
            Update current belief P_i^k using (4)
        end if
    end if
end for
return Updated belief P_i^{k+1}
```

Algorithm 3 Proposed Opinion Dissemination Model

```
1:  Initialize social network G(U, E), ground truth z, thresholds (σ, ε, δ), and log-likelihood function
2:  Initialize belief, reliability, and doubt for all users in G(U, E)
3:  while Accuracy R < δ or k ≤ K do
4:      for i = 1 to N do
5:          /* The process in lines 6-10 reflects step 1 */
6:          if P_i^k ≥ σ then
7:              Communicate o_i^k = True to all u_j in F_i
8:          else if P_i^k ≤ 1 − σ then
9:              Communicate o_i^k = False to all u_j in F_i
10:         end if
11:         /* The process in line 12 reflects step 2 */
12:         Calculate updated reliability t_ij^k and doubt f_ij^k for all u_j in F_i using Algorithm 1
13:         /* The process in line 14 reflects step 3 */
14:         Calculate updated belief P_i^{k+1} using Algorithm 2
15:     end for
16:     Set k = k + 1 and calculate accuracy R
17: end while
18: Calculate convergence time k^c
```

VI. Evaluation

We conducted experiments to evaluate the performance of the proposed model and AAT using three artificial social networks and two real-world social networks. The metrics are accuracy R and convergence time k^c. Accuracy means the ratio that users form the same opinions as ground truth of a piece of news z, and convergence time means the first belief update step that satisfies $R \geq \delta$. The models and experiments were implemented using Python 3.7.3.

A. Artificial and Real-world Social Network Datasets

We first used three commonly used types of artificial networks as social network $G(U, E)$: a scale-free network, a small-world network, and a random network. A scale-free network is a network with a few nodes (i.e., users) called "hubs" connecting a large number of friends. A small-world network is a network that most nodes reach every other node in a small number of hops. A random network is a network in which any two nodes connect with probability p. Following the network settings of the AAT [14], we set the number of users $N \in \{150, 325, 500, 750, 1000, 1500, 2000\}$ and set the average number of each u_i's friends $|\bar{F}_i|$ from 4 to 12 in these networks.

We also used the publicly available Facebook and Twitter social network datasets as real-world social network $G(U, E)$ [24]. The Facebook dataset consists of $|U| = 4039$ users, $|E| = 88,234$ edges, and $|\bar{F}_i| = 44$ users, which is the average number of each u_i's friends. The Twitter dataset consists of $|U| = 81,306$ users, $|E| = 1,768,149$ edges, and $|\bar{F}_i| = 33$ users, which is the average number of each u_i's friends.

B. Experimental Settings

For the proposed model, we initialized reliability $t_{ij}^{(0)}$ and doubt $f_{ij}^{(0)}$ so that they followed a uniform distribution. In accordance with this initialization, we simulated the model using Algorithms 1-3.

For the AAT model, the reliability of each u_i is by definition $t_i = P(o_j = \mathrm{T}|z = \mathrm{T}) \in [0.5, 1.0]$, $1 - t_i = P(o_j = \mathrm{T}|z = \mathrm{F})$. The range $t_i \in [0.5, 1.0]$ indicates that each u_i equally believes his or her friends from 50% to 100%. That is, each u_i does not doubt any of his or her friends. Moreover, reliability does not change. Following the setting of reliability in [14], we randomly assigned a fixed value of reliability t_i, which is quantized by 0.05 in the range of $[0.5, 1.0]$, to each u_i in a social network. We simulated the AAT without executing Algorithm 1 (i.e., reliability was not updated).

We set the parameters as follows: the maximum belief update step $K = 3000$, the maximum reliability and doubt update step $S = 15$, the threshold to converge the EM algorithm $\epsilon = 0.00001$, the threshold to form opinion $\sigma = 0.8$, and the threshold of accuracy $\delta = 0.8$. These parameter settings follow those in previous work [14] [15]. We also initialized the belief of each u_i P_i^0 so that it followed a normal distribution $\mathcal{N}(\text{mean} = 0.5, \text{standard deviation} = 0.15)$.

When accuracy R was less than 80% within $K = 3000$ (i.e., the proposed and AAT models did not converge), we considered convergence time k^c as 3000.

C. Results with Artificial Networks

The results, shown in Figs. 2-4, are the average over the 50 times simulations in each ground truth $z \in \{\text{True}, \text{False}\}$.

First, we focus on the results for accuracy (Fig. 2). When $z = \text{True}$, the average accuracy of the proposed model for all networks and users was 88.5%, while that for AAT was 42.5%. In contrast, when $z = \text{False}$, the average accuracy of the proposed model for all networks and users was 60.3%, while that for AAT was 45.2%. These results indicate that people can less accurately perceive that fake news is actually fake (i.e., $z = \text{False}$) than they can perceive that real news is actually real (i.e., $z = \text{True}$).

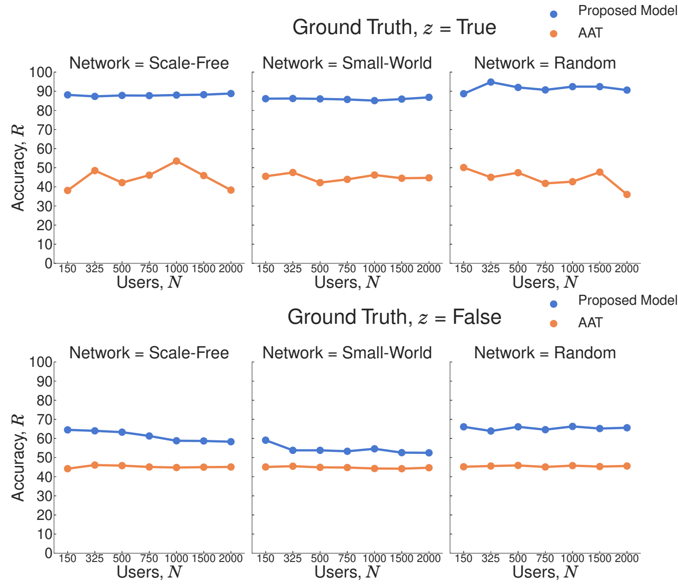

Fig. 2. Accuracy of the proposed model and AAT for artificial social networks.

Next, we focus on the results for convergence time (Fig. 3). When $z = \text{True}$, the proposed model converged much faster than the AAT. The average convergence time for all networks and users was $k^c = 3.18$, while that for AAT was $k^c = 1900$. When $z = \text{False}$, the average convergence time for all networks and users was $k^c = 2977$, while the AAT did not converge. These results indicate that people immediately perceive that real news is real (i.e., $k^c = 3.18$) and that it takes much time for people to perceive that fake news is fake (i.e., $k^c = 2977$).

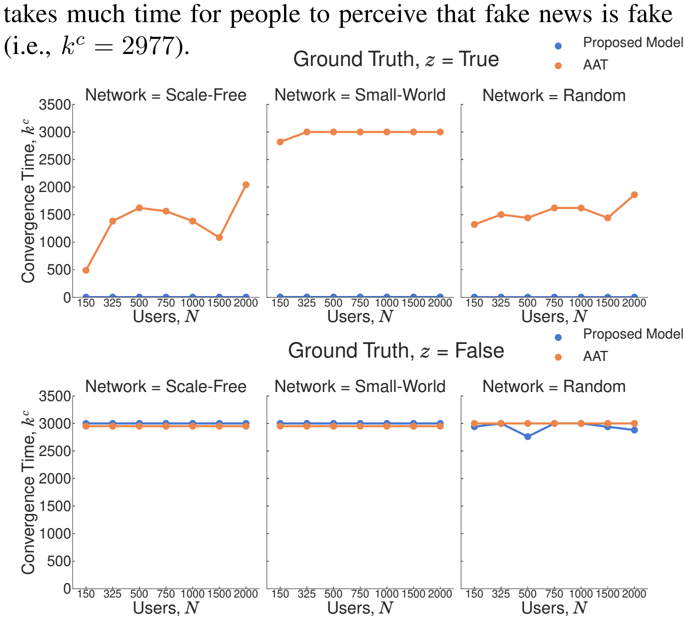

Fig. 3. Convergence time of the proposed model and AAT for artificial social networks.
Note: Convergence time for the scale-free and small-world network was 3000 both the proposed model and the AAT.

To consider the results for accuracy and convergence time in the proposed model, we visualized the change of distributions in reliability and doubt when the number of users $N = 2000$ as an example (Fig. 4), where these distributions initially followed a uniform distribution. In Fig. 4, the x-axis gives the value of reliability or doubt. The y-axis gives the number of users who have the value of reliability or doubt given in the x-axis. When $z =$ False, reliability was lower, and doubt was higher than when $z =$ True. These values were distributed around 0.5. The result indicates that people become skeptical about friends (i.e., they believe friends about 50% and doubt friends about 50%) in the presence of fake news (i.e., when $z =$ False). Consequently, most of the people did not update their beliefs and opinions much because reliability and doubt did not affect belief update. In concluding the consideration, the findings of accuracy and convergence time in $z =$ False are because people become skeptical about friends and thus people do not update their opinions much. On the other hand, the findings of accuracy and convergence time in $z =$ True are because people are relatively reliant on friends and thus people update their opinions.

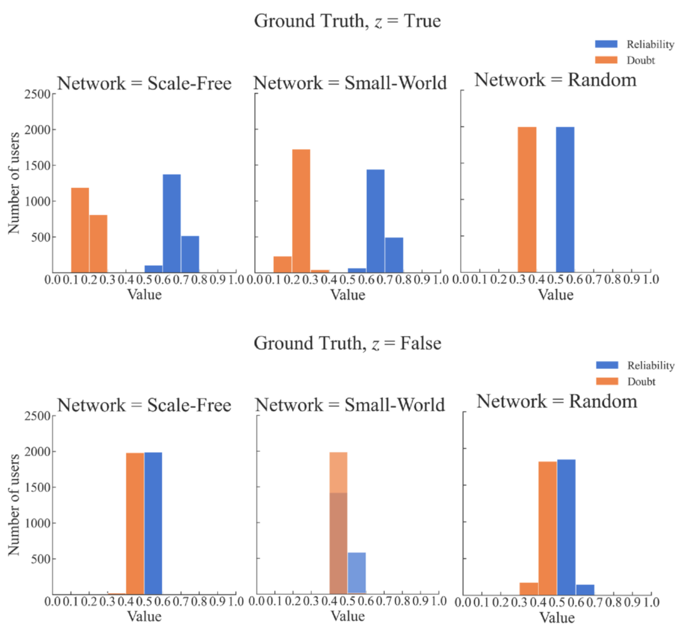

Fig. 4. Distribution of the reliability and doubt for artificial social networks when the number of users $N = 2000$.
Note: The overlapped bar shown in the small-world network when $z =$ False means that users have both reliability and doubt between 0.4 and 0.5.

D. Results with Real-world Social Networks

The results, shown in Figs. 5-6 and TABLE I, are the average over the 50 times simulations in each ground truth z. The accuracy and convergence time showed that the proposed model exhibited similar tendencies regarding accuracy and convergence time in Section VI-C (Fig. 5, TABLE I). Moreover, the result of the distributions of reliability and doubt also exhibited similar tendencies in Section VI-C (Fig. 6). These results demonstrate the effectiveness of using our proposed model to analyze opinion dissemination in the presence of fake news even when the proposed model is used in real-world social networks.

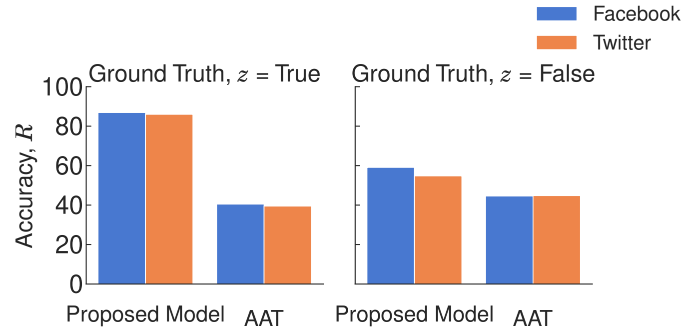

Fig. 5. Accuracy of the proposed model and AAT for Facebook and Twitter social networks.

TABLE I
CONVERGENCE TIME OF THE PROPOSED MODEL AND AAT FOR FACEBOOK AND TWITTER SOCIAL NETWORKS.

	Network	Ground Truth, z = True	Ground Truth, z = False
Proposed Model	Facebook	8.8	3000
Proposed Model	Twitter	13.1	3000
AAT	Facebook	3000	3000
AAT	Twitter	3000	3000

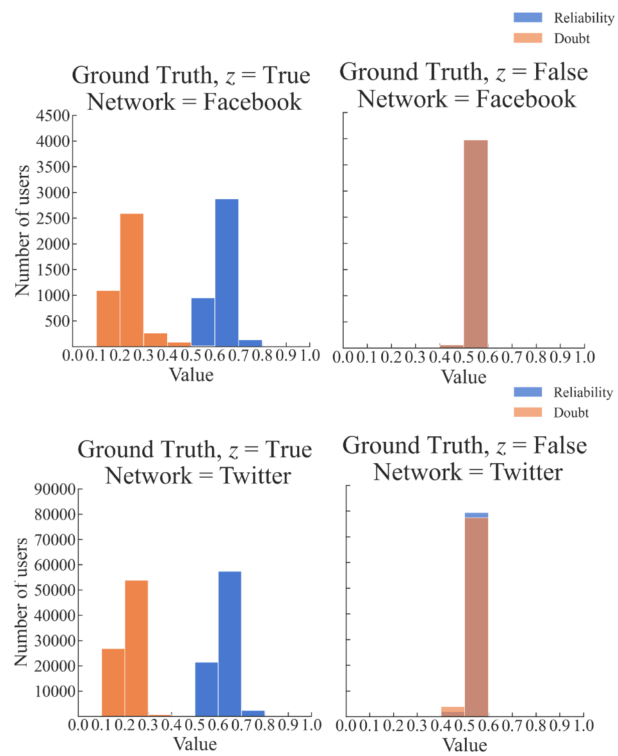

Fig. 6. Distribution of the reliability and doubt for Facebook and Twitter social networks.
Note: The overlapped bar when $z =$ False means that users have both reliability and doubt between 0.4 and 0.6.

VII. Conclusion

In this paper, we constructed an opinion dissemination model in which people in a social network communicate their opinions about the fact of news. In our proposed model, people update their opinions on the basis of their friends' opinions in accordance with Bayes' theorem and update the reliability of and doubt about each friend by using the expectation maximization algorithm. The proposed model can better represent opinion dissemination in the presence of fake news than previously proposed models.

Using our proposed model on three artificial social networks and two real-world social networks, we found three clues for understanding how fake news disseminates. Firstly, the number of people who shared the correct opinion for fake news decreased by about 28% compared to the number for real news, meaning that people can less accurately perceive that fake news is fake than they can perceive that real news is real. Secondly, people needed about 502 times more opinion updates to share correct opinions for fake news than for real news, meaning that it takes much more time for people to perceive fake news to be fake than to perceive real news to be real. The number "502" was calculated from the average ratio of convergence time k^c for real news and fake news in three artificial social networks $\left(\frac{\text{convergence time } z\text{=False}}{\text{convergence time } z\text{=True}} = \frac{2977}{3.18}\right)$ and two real-world social networks $(\frac{3000}{8.8}, \frac{3000}{13.1})$ (Section VI-C, TABLE I). Finally, these results regarding fake news are because people become half in doubt about friends in the presence of fake news, and hence people do not update their opinions much. These clues should be helpful in developing countermeasures against fake news.

Future work includes extending our model by incorporating the inherent human tendencies, such as the backfire effect and the echo chamber effect.

References

[1] S. Vosoughi, D. Roy, and S. Aral, "The spread of true and false news online," *Science*, vol. 359, no. 6380, pp. 1146–1151, 2018.

[2] H. Allcott and M. Gentzkow, "Social media and fake news in the 2016 election," *Journal of economic perspectives*, vol. 31, no. 2, pp. 211–36, 2017.

[3] M. Delirrad and A. B. Mohammadi, "New methanol poisoning outbreaks in iran following covid-19 pandemic," *Alcohol and Alcoholism (Oxford, Oxfordshire)*, 2020.

[4] V. Rubin, N. Conroy, Y. Chen, and S. Cornwell, "Fake news or truth? using satirical cues to detect potentially misleading news," in *Proceedings of the second workshop on computational approaches to deception detection*, 2016, pp. 7–17.

[5] A. Gupta, H. Lamba, P. Kumaraguru, and A. Joshi, "Faking sandy: characterizing and identifying fake images on twitter during hurricane sandy," in *Proceedings of the 22nd international conference on World Wide Web*, 2013, pp. 729–736.

[6] E. Tacchini, G. Ballarin, M. L. Della Vedova, S. Moret, and L. de Alfaro, "Some like it hoax: Automated fake news detection in social networks," *arXiv preprint arXiv:1704.07506*, 2017.

[7] Z. Jin, J. Cao, Y. Zhang, and J. Luo, "News verification by exploiting conflicting social viewpoints in microblogs," in *Thirtieth AAAI conference on artificial intelligence*, 2016.

[8] D. Graves, "Understanding the promise and limits of automated fact-checking," 2018.

[9] N. Hassan, G. Zhang, F. Arslan, J. Caraballo, D. Jimenez, S. Gawsane, S. Hasan, M. Joseph, A. Kulkarni, A. K. Nayak *et al.*, "Claimbuster: the first-ever end-to-end fact-checking system," *Proceedings of the VLDB Endowment*, vol. 10, no. 12, pp. 1945–1948, 2017.

[10] N. Vo and K. Lee, "The rise of guardians: Fact-checking url recommendation to combat fake news," in *The 41st International ACM SIGIR Conference on Research & Development in Information Retrieval*, 2018, pp. 275–284.

[11] W. Quattrociocchi, A. Scala, and C. R. Sunstein, "Echo chambers on facebook," *Available at SSRN 2795110*, 2016.

[12] B. Nyhan and J. Reifler, "When corrections fail: The persistence of political misperceptions," *Political Behavior*, vol. 32, no. 2, pp. 303–330, 2010.

[13] R. T. Glinton, P. Scerri, and K. Sycara, "Towards the understanding of information dynamics in large scale networked systems," in *Information Fusion, 2009. FUSION'09. 12th International Conference on*. IEEE, 2009, pp. 794–801.

[14] O. Pryymak, A. Rogers, and N. R. Jennings, "Efficient opinion sharing in large decentralised teams," in *Proceedings of the 11th International Conference on Autonomous Agents and Multiagent Systems-Volume 1*. International Foundation for Autonomous Agents and Multiagent Systems, 2012, pp. 543–550.

[15] D. Wang, L. Kaplan, H. Le, and T. Abdelzaher, "On truth discovery in social sensing: A maximum likelihood estimation approach," in *Proceedings of the 11th international conference on Information Processing in Sensor Networks*. ACM, 2012, pp. 233–244.

[16] C. Huang, D. Wang, and N. Chawla, "Towards time-sensitive truth discovery in social sensing applications," in *2015 IEEE 12th International Conference on Mobile Ad Hoc and Sensor Systems*. IEEE, 2015, pp. 154–162.

[17] J. Marshall and D. Wang, "Towards emotional-aware truth discovery in social sensing applications," in *2016 IEEE International Conference on Smart Computing (SMARTCOMP)*. IEEE, 2016, pp. 1–8.

[18] S. Kwon, M. Cha, and K. Jung, "Rumor detection over varying time windows," *PloS one*, vol. 12, no. 1, p. e0168344, 2017.

[19] E. Ferrara, O. Varol, C. Davis, F. Menczer, and A. Flammini, "The rise of social bots," *Communications of the ACM*, vol. 59, no. 7, pp. 96–104, 2016.

[20] A. Tsang and K. Larson, "Opinion dynamics of skeptical agents," in *Proceedings of the 2014 international conference on Autonomous agents and multi-agent systems*, 2014, pp. 277–284.

[21] K. Sasahara, W. Chen, H. Peng, G. L. Ciampaglia, A. Flammini, and F. Menczer, "On the inevitability of online echo chambers," *arXiv preprint arXiv:1905.03919*, 2019.

[22] X. Yin, J. Han, and S. Y. Philip, "Truth discovery with multiple conflicting information providers on the web," *IEEE Transactions on Knowledge and Data Engineering*, vol. 20, no. 6, pp. 796–808, 2008.

[23] M. Enya, M. Nagime, and D. Gusia, "Postponing the tokyo 2020 olympic games: stress-testing governance in the olympic system and limitations to agenda 2020," *Olimpianos-Journal of Olympic Studies*, vol. 4, pp. 49–63, 2020.

[24] J. Leskovec and A. Krevl, "SNAP Datasets: Stanford large network dataset collection," http://snap.stanford.edu/data, jun 2014.