Learning Support Method of Information Ethic by a Virtual Network Isolating Risky Messages to SNS

Hajime Iwata , Kento Inoue
Kanagawa Institute of Technology
Department of Information Network and Communication
Atsugi, Japan
e-mail: hajimei@nw.kanagawa-it.ac.jp, inoue@nwlab.org

***Abstract*—Social networking services (SNSs), which are global phenomena, allow users to share messages with others, resulting in a rapid widespread distribution. Because an inappropriate post may cause trouble for a user, it is important for a user to be aware of potential issues. In this study, we propose an information ethics system to determine the degree of risk of a post. When the system detects a high-risk post, it isolates the post on a virtual network.**

Keywords- SNS; Social Networking Service; Information Ethics; Virtual Network; OpenFlow .

I. INTRODUCTION

The Internet is ubiquitous, and many users enjoy social networking services (SNSs), which promote communications with others around the globe. On a SNS network, a user posts a message willingly and other users provide feedback in a short timeframe, allowing a user to receive approval from others. Occasionally a user receives negative feedback when sharing an inappropriate post, image, or video. Often inappropriate content spreads quickly before the user can retract the post. Thus, a user must always be aware of whether a post contains problematic content. However, users, especially those unfamiliar with SNSs, have a low consciousness of a post's risk.

In this study, we suggest a system that can learn information ethics by checking the degree of risk of a user's posts. The user posts on a SNS via a virtual network. The virtual network evaluates the risk of the post. Typically, the virtual network posts directly to the SNS. However, if the system detects a high-risk post, it isolates the post on a virtual network and sends a warning to the user, allowing the user to naturally learn about information ethics.

Section 2 describes the feature of our study. Section 3 describes the architecture of the proposed virtual network system. Finally, section 4 concludes this paper.

II. FEATURE OF OUR STUDY

Twitter is one example of a SNS with many users. In Twitter, the privacy settings are public and protected. When other users see a post, they can easily share a public post via the retweet function. Furthermore, many users post messages that unconsciously include personal information, which others can easily extract. Through this process, posts, including inappropriate ones or those with sensitive information, spread in a short time. In addition, the likelihood that emotional contents are included is high because Twitter users post in short sentences. Herein we suggest an information ethics learning support system for Twitter users.

We established a risk analysis system server on a virtual network. This system transmits a user's post once to a virtual network. If the content is problematic, this system stops the post from going public, reducing fears of inappropriate posts from being scattered on the Internet. This system also issues a warning to the user that a post is inappropriate by transmitting a message to the user from the virtual network. In addition, posts with privacy breaches of the users are stored in the virtual network. Because such posts are handled in a virtual network, the contents are not watched by a third party nor are they rapidly spread on the Internet.

III. SYSTEM SUMMARY

Fig. 1 overviews the constitution of the network. The virtual network increases the flexibility of the overall constitution and the placement of the system by controlling it using OpenFlow [1]. For OpenFlow controller, we set a packet of SNS client software installed in a client PC. The OpenFlow switch can divide a posting packet to SNS in two from a packet transmitted from a client PC.

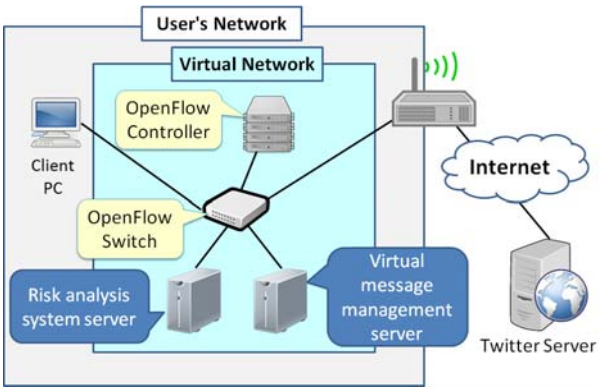

Figure 1. Constitution of the network

 ISBN: 978-1-61208-504-3

This system is easy to install in existing networks. Because this system is implemented using JAVA, it is independent of the platform. In addition, we intended for Twitter by this system as SNS. The access to Twitter uses Twitter4j [2] as a library and free (no cost) open use is possible.

Fig. 2 shows the structure of the risk analysis system. The system initially reads the new post of the user and evaluates it based on the user's previous Twitter posts. Then the system conducts a morphological analysis for a message. Currently, the system only supports posts written in English or Japanese. The English morphological analysis uses Stanford Log-linear Part-Of-Speech Tagger [3], while the Japanese morphological analysis uses lucene-gosen [4].

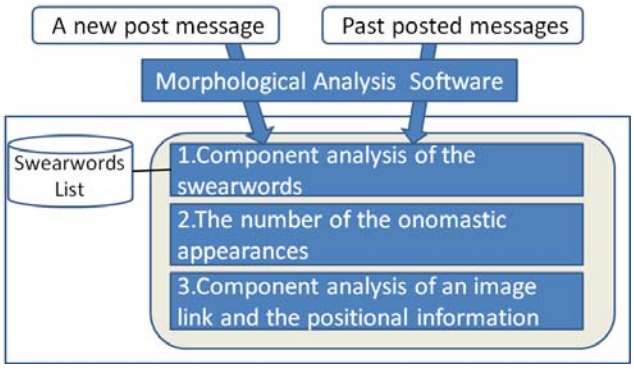

Figure 2. Structure of the risk analysis system

A morphological analysis library allows the system to evaluate the words in a post. This system is based on three factors:

1. Component analysis of the swearwords

The risk analysis system reads a list of swearwords stored on the server. This system compares each word in the message to the list of swearwords and analyzes the risk based on the number of swearwords. Currently, the list of swearwords is compiled manually. The swearwords include profanity, inflammatory, discriminative, and crime-related terms. In addition, the swearwords are converted into a basic form after the morphological analysis and a conjugation before converting to a list.

2. Number of the onomastic appearances

The risk analysis system stores proper nouns, including names and locations. Proper nouns in a message are not immediately considered risky. However, a proper noun becomes an identifiable factor when it appears consecutively in a short timeframe. When the number of the onomastic appearances continues, this system warns the user. Thus, this system can highlight risks to the user.

3. Component analysis of an image link and the positional information

When a link to an image is included in a post, the risk analysis system warns the user. Because automatically analyzing the content of image data is difficult, the system confirms that inappropriate contents are not included. In addition, this system determines whether a post contains GPS information because such information increases the risk that personal information such as the user's home address can be identified. When positional information is included in a post, this system provides a warning to the user.

If a message is problem-free according the risk analysis system, the virtual server posts the message to Twitter. When an item is risky, the virtual server sends a warning to the user and does not post the message to the actual Twitter server. This system encourages users to correct the content. Because the contents of the messages accumulate on the server in a virtual network, the user can ignore the warning. However, the message is not transmitted outside the virtual network, preventing public comments and retweets on a public server. However, the server on the virtual network allows replies and retweets from virtual users, and the system displays issues with the message. Thus, the user can experience the consequence of an inappropriate post virtually, allowing the user to learn information ethics on a non-public server.

IV. CONCLUSION

In this paper, we propose a system that learns and then teaches users about information ethics of SNSs. Our system teaches users about risky situations and the consequences of an inappropriate message on a virtual system, which is more effective than a user learning on a public SNS.
Future works include:

- Implementation of the server on a large-scale virtual whole network
- Automation of the system populating the swearwords list
- Improvement of the examples of trouble from a risky message

REFERENCES

[1] N. McKeown, et al., "OpenFlow: enabling innovation in campus networks", ACM SIGCOMM Computer Communication Review,38, 2 (March 2008), pp.69-74, 2008.

[2] "Twitter4J - A Java library for the Twitter API", http://twitter4j.org/en/ (July, 2016)

[3] The Stanford Natural Language Processing Group, "Stanford Log-linear Part-Of-Speech Tagger", http://nlp.stanford.edu/software/tagger.shtml, (July, 2016)

[4] "lucene-gosen", https://github.com/lucene-gosen/lucene-gosen, (July, 2016)

 ISBN: 978-1-61208-504-3