Displaying Title Is Effective to Reduce Aggressive Comments in Pseudonym Type Computer Mediated Communications

Tetsuya MAESHIRO * and Hiroki YAMAGUCHI **

Abstract: We analyzed a mechanism to alleviate aggressions on computer mediated communications. A title that supposedly reflects the user’s behavior or status, such as calm, expert, optimist, or aggressive is displayed on the screen. Two types of title generation, title reflecting and not reflecting exact aggression level, are compared to the case without title using laboratory experiments, where participants select comments to be posted after reading the topics of discussion and other participants’ comments. The results indicate that displaying a title that is unrelated to the aggressiveness of comments posted by the user is equivalent to displaying a title based on accurately calculated aggression level of comments. Therefore, simply displaying their titles to users is effective to reduce aggressiveness of their comments in pseudonym type computer mediated communications.

Key Words: title assignment, social network service, computer mediated communication, aggressiveness.

1. Introduction

Computer mediated communication (CMC) is becoming one of the main communication tools, resulting in the growth of users and consequently in the increased number of posted comments. CMC is defined as a communication between two people intermediated by two or more computer devices. Besides the benefits of CMCs, a noticeable drawback is aggressions among users through hostile comments posted by users. No definitive solution, however, exists.

Suppose that a title reflecting comments posted by a user in computer mediated communication (CMC) services, such as calm, expert, optimist, or aggressive, which is constantly displayed to the user, is effective to reduce aggressive comments. Suppose further that randomly assigned title, without any association with her own posted comments or other users’ evaluations, is sufficient. We have analyzed the effectiveness of this title assignment and have found that it is a viable solution to reduce aggressive comments in CMC, with effectiveness equivalent to the algorithm-based title assignments.

This paper treats aggressions in CMC, particularly those services where users send their messages mainly in text form but are not limited to it. A widely accepted definition of face-to-face aggression is “an intentional act, with the potential to harm, committed by an individual in an aroused physical state and perceived as aversive by the victim” [1]. The aggression treated here is not face-to-face, but similar definitions exist for cyber aggression, which is “intentional behavior aimed at harming other persons through computers, cell phones, and other electronic devices, and perceived as aversive by the victim” [2]–[5]. Aggressive comments by the users in CMC services sometimes result in flaminis, deteriorating the quality of these services.

Our previous study [6] reported the results of an experiment using calculated, algorithm-based title assignments, where the title displayed to the user is based on the aggressiveness level measured from the comments selected by the user. In contrast to the algorithm-based title calculation, we reported the results of an experiment using random title assignment [7]. This paper analyzes and compares the effectiveness of the title assignment by introducing a new experiment parameter, which is the easiness for users to predict assigned titles, and presents details of the effects of the title assignment. Since a more detailed analysis is also presented, it enables a clearer understanding of the benefits of the title assignment. This paper also presents a new parameter not analyzed in the previous study [6], the time duration for users to select their comments, which is related to the time that users spend to post comments, which should also be related to the users’ attitude related to the aggressiveness of the comments to select.

This paper proposes a mechanism to reduce aggression in anonymous CMC services.

2. Methods

2.1 Experiment Design

Effects of displaying user’s title in pseudonym CMC services, such as calm, expert, optimist, or aggressive, are not well understood. When the title assigned to the user is displayed to her while using a CMC service, the user might be aware or not of her title.

To analyze the effects of displaying title to the user and the implications of the title generation algorithms, we designed an experiment with the following parameters.

(1) The existence of the displayed title. The title supposedly reflects the aggression level of the messages posted by the user to the CMC service.

(2) Easiness for a user to estimate the calculation method of the assigned title. In other words, it is the degree of intuitiveness that the user feels about the displayed title. It is also related

---

* Faculty of Library, Information and Media Studies, University of Tsukuba, Tsukuba 305–8550, Japan
** Graduate School of Library and Information Science, University of Tsukuba, Tsukuba 305–8550, Japan
E-mail: maeshiro@slis.tsukuba.ac.jp
(Received March 31, 2017)
(Revised September 2, 2017)
to the controllability of the assigned title by the user through managing the wordings and expressions employed in messages posted to CMC service.

Three experiments were conducted using the above parameters.

Control experiment: No displayed title.

Experiment 1 (Intuitive): Title displayed. The title reflects the actual aggressive levels of the comments posted by the user. This algorithm is denoted as Algorithm A.

Experiment 2 (Counterintuitive): Title displayed. The title does not reflect the actual aggressive levels of the comments posted by the user. This algorithm is denoted as Algorithm B.

The details of the title generation are explained in the following sections.

2.2 Experiment System

An experimental CMC system was built to verify the validity of the proposed mechanisms.

The experiment was conducted as follows.

1. The participant inputs the account name.

2. Repeat 20 times the following steps, one topic for each cycle.

   Fig. 1 Screenshot of experimental system at Step 1. The title of the user is shown at the top.
   
   Fig. 2 Screenshot of experimental system at Step 2.
   
   Fig. 3 Screenshot of experimental system at Step 3. Although shown on the actual screen, the title of the user is omitted in this figure.

Step 1 The participant reads the description of the discussed topic and comments from other four users in the system about the topic. Two agreeing and two opposing comments are displayed to the user as comments posted by four distinct users, i.e., using four different pseudonym account names (Fig. 1). One of the two agreeing comments has low aggression level, whereas the other one has a high aggression level. Similarly, the two opposing comments have high and low aggression levels. Then the participant chooses (i) agree or (ii) disagree (Fig. 1).

Step 2 The participant reads the displayed comments and then selects one comment among other users’ comments whom the participant wants to reply. If there are no suitable comments to reply, the participant does not have to select any user’s comment (Fig. 2).

Step 3 The participant reads the eight candidate comments that will be used as the replying comment to the user and comment selected in Step 2 and then selects one comment that the participant wants to use as the replying comment to the forum. If there are no suitable comments, the participant does not have to select any comments (Fig. 3). The eight comments shown in Step 3 were generated from two different arguments to improve the simulation accuracy of the real CMC user actions, each with four candidate comments having different aggression levels.

3. The participant answers a questionnaire.

Participants selected one comment for each topic from candidate comments which are prewritten sentences whose aggression levels were predetermined in a pre-analysis. The aggression levels of comments were hidden to the participants.

The number of topics was 20 so that participants repeat 20 times Steps 1 to 3. Topics used in the experiment were 20 topics...
that actually resulted in active discussions, and extreme opinions, including some kind of rivalry in existing CMC services, were extracted and used in the experiment, such as juvenile’s act, nuclear power plant, and education of the second language from childhood. Such topics were selected because they might originate aggressive opinions and comments and were suitable to verify the attenuation effect of aggressiveness.

The experiment consisted of four sets, where each set consisted of five topics (Fig. 4). The duration of one experiment was about 60 minutes. The experimental room was laid silent. There were a desk and a laptop computer to access the experimental forum system. Only one person participated in one experimental session.

At Step 1, four comments posted by four “virtual” users were presented to the participant, in order to simulate an on-going discussion. The four comments consist of two agreeing opinions to the presented topic and two disagreeing opinions. Twenty topics were provided in an experimental session; thus the total number of comments was 400. These comments consist of comments extracted from existing CMC services and created comments.

Only half (eight) of provided comments were presented to the participants in Step 3, based on the agreeing or disagreeing opinion selected by the subject in Step 2. Furthermore, the displayed comments consist of two arguments to provide a wide range of opinions so that participants are easier to find their opinion. Four comments based on each argument, totaling eight comments, were displayed to participants. For example, suppose a participant selects “Agree” for the topic “Do you agree or disagree with the juvenile act?” at Step 2. In eight agreeing candidate comments presented at Step 3, where four candidate comments are based on the first argument (“We must execute severe punishments on juvenile delinquency for regeneration”), and another four are based on the second argument (“Victims or bereaved family cannot be convinced of the reason for immunity of sentence for youth criminals”). In addition, the four comments based on each argument have different aggression levels A, B, C and D (the weakest is A, the strongest is D), each with the following preassigned aggression levels: A: 1.0–1.9; B: 2.0–2.9; C: 3.0–3.9; D: 4.0–5.0. In the experiment, the comments were displayed in random order to avoid habituation by the participants to the sequential order of aggression levels of comments presented to them.

2.3 Title Assignment Algorithm

The titles assigned and displayed to the participants were (1) rabid, (2) normal, and (3) clement. The title rabid denotes the most aggressive user title, clement denotes the least aggressive, and normal the average. The initial title was normal in the Set-1. The titles in the subsequent sessions were calculated differently for Algorithms A and B. Note that the title did not change during the same set. Basically, the title selection algorithm employs the average aggression levels of selected comments in the set \( S_n \) to determine the title in the next set \( S_{n+1} \), which is an intuitive method and relatively easy to be guessed by the participants.

All experiments consist of four sets, where four sessions compose a set (Fig. 4). The title remains the same during each set and changes between sets. Therefore, the title changes three times for a participant, between the sessions 1 and 2, 2 and 3, and 3 and 4.

The assigned title is calculated based on the aggression levels
of comments selected by the participant in the four sessions of the previous set, given by

$$ S_T = \sum_{i=1}^{4} v_{i,T} $$

(1)

where $T$ denotes the set number, $T = 1, \ldots, 4$, $S_T$ is the aggression score of the set $T$, and $v_{i,T}$ is the aggression level of the comment selected by the participant in session $i$ of the set $T$. Then the title assigned to the participant in the next set $T+1$ is calculated as follows. First, the aggression scores of all comments displayed in the four sessions of the set are collected. Since eight comments are shown in each session, the total number of comments in a set is $4 \times 8 = 32$. Then the quartile deviation $S_q$ is calculated. The assigned title will be

**Rabid:** if $S_T$ is above the upper $S_q$

**Clement:** if $S_T$ is below the lower $S_q$

**Normal:** otherwise, i.e., if $S_T$ is between the lower $S_q$ and the upper $S_q$.

The following two algorithms of title assignment were tested:

**Algorithm A:** Intuitive title. The title reflecting the aggressiveness of user’s comments is shown.

The assigned title is based on the average aggression level of the selected comments in the previous session. This title is the predictable title because the title reflects the aggression level of the selected comments. Figure 4 illustrates the title changing mechanism of Algorithm A.

**Algorithm B:** Counterintuitive title. Independent of the aggressiveness of user’s comments, randomly chosen title is shown.

The easily predictable title (title calculated by Algorithm A) is excluded from the assignment set. If the calculated title is normal, for instance, the other two titles, rabid and clement, are used as the possible titles, and one of them is randomly selected. Then the title is randomly assigned in the following three sets, changing three times at an interval (one minute) between the sets.

Figure 5 illustrates the title changing mechanism of Algorithm B.

The participants are easier to speculate the title shown by the CMC system using Algorithm A but are difficult when using Algorithm B because the displayed titles are counterintuitive.

### 2.4 Preparation of Comments

A total of 400 comments were provided for the experiment of one participant, since 20 comments were necessary for each session (four for other participants’ comments in Step 2 and eight for agreeing and eight for disagreeing in Step 3), five sessions constituted a set, and an experiment consisted of four sets $(4 + 8 + 8) \times 5 \times 4 = 400)$. Aggression levels of all comments displayed in the experiment system (Fig. 3) were pre-calculated, since they were necessary to compute the exact aggression level of participants, and then to calculate the titles exposed to participants using algorithms explained in the previous section (Eq. (1)). Comments with four levels of aggression levels were used, and 100 comments were prepared for each of four aggressions levels, totaling 400 comments.
We arranged the aggression level of comments by combining four aspects related to their content: Slanderous, Critical, Declarative and Affirmative. Slanderous is when the context and its words cause others uncomfortable. Critical indicates the words and context are contradictory. They can be used as a broader indicator of aggression. Declarative indicates the words and context do not include any softener expression, which is defined as the ending of sentences to make them milder, such as “in my humble opinion” or “for what it’s worth” [8]. Affirmative indicates the words and context are not belonging to the previous three content elements. In addition, the aggression level of the four example comments from other users (two agreeing and two disagreeing opinions) presented at Step 1 (Fig. 2) are level D, designed to persuade subjects to select aggressive comments from the eight candidate comments at Step 2 and 3.

Prior to the experiments, we executed quantification of aggression levels of all 400 comments used in our system. Each comment was reviewed by five persons. Note that no reviewers participated in the experiments. Reviewers rated the comments into five levels, 1–5, where 1 is not aggressive and 5 is extremely aggressive. The average of the scores assigned by five reviewers was defined as the aggression level of a comment. After the first rounds of aggression level assignment, those comments with a large discrepancy between the measured and intended aggression levels had their expression modified and measured again by the five reviewers until the discrepancy was smaller than 0.5.

### 2.5 Performance Indicators

The following indicators were used to measure the performance of the proposed method.

**Indicator 1:** The average aggression level of comments selected by participants.

**Indicator 2:** The average time duration to select (formulate) the comments to post.

The second indicator is used because displaying the attached title to the users might affect their self-consciousness to remind that their comments are assessed by other people, affecting the time to select the comments to post.

### 3. Results and Discussions

The number of participants was 72 (age: 18 to 27 years old), and Table 1 shows the ratio of male to female. They were divided into the two experimental groups, Algorithm A (N = 24, the title was presented) and Algorithm B (N = 24, the title was presented), and the control group (N = 24, the title was hidden). The experiments were approved by the University of Tsukuba ethics committee.

| Indicator Level | Control | Algorithm A | Algorithm B |
|-----------------|---------|-------------|-------------|
| Aggression Level| 0.09    | 0.11        | 0.08        |
| Comment Selection (s) | 0.39 | 0.37         | 0.37        |

Since the average aggression levels of Algorithms A and B were approximately the same, it suggests that simply displaying the title effectively reduces the aggression level of the posted comments. Although we tested only two methods to calculate the title of the user, these two methods can be positioned as the two extremes; one is totally unpredictable or counterintuitive (Algorithm B), and the other one is totally predictable or algorithmic (Algorithm A). Exhausting testing of all possible title generation algorithms would be necessary to obtain the complete picture of the influence of the algorithm. However, the results of the two extreme cases (Algorithms A and B) suggest that displaying the title to the user does reduce the aggression level, independent of the algorithm to calculate the title.

The three titles used in the experiment (rabid, clement, and normal) have meanings related to aggressiveness, and this paper measured the participant’s actions related to aggressiveness. For titles without meanings related to aggressiveness, new experiments are necessary to measure their effects. However, we predict that other titles have a similar effect, which is the attenuation of extreme actions.

### 3.2 Comment Selection Time

Figure 7 illustrates the average time duration of comment selection. No difference was found between Algorithm A and the
control group. On the other hand, the time of Algorithm B was 18.3% shorter ($p < 0.01$) than the control group. The results of the Indicator 2 are interesting as the algorithm to calculate the title influences the user’s behavior. All participants executed only one experiment type, so the learning factor due to the familiarization to the experimental environment can be excluded. The standard deviation of the comment selection time was also constant among three experimental groups, although they were larger than the standard deviation of the aggression level (Table 2). These values indicate that the existence of the title did not affect the distribution of comment selection time.

The analysis, however, indicated that no correlation exists between the BAQ score and the average aggression levels of selected comments. Therefore the reduced aggression levels in Algorithms A and B are probably due to showing their titles to the users.

The ratio between male and female was approximately the same in Algorithm A and B groups (Table 1). In the control group, however, male participants were approximately double of females. This is not a problem because the gender difference seems to have no effect on cyber aggression [2]. Moreover, our experiment treats pseudonym user account cases, so we assume that difference of gender ratio did not affect the average aggression of selected comments.

Aggressive comments, those above 4.0, were hardly selected in both the experimental and control groups. It is possible that the participants do not tend to select highly aggressive comments, or the experiment environment repressed the selection of aggressive comments. To verify this point, the experimental design should be improved. The participants selected their comments in Step 3 (Fig. 3), where comments with different aggression levels were shown simultaneously. This might have helped the participants to choose comments with appropriate aggression level by comparing the displayed comments. This experiment setting is different from actual CMC usage conditions, where users write comments using character input devices. Although the experiment setting could have affected participants’ behavior, it influenced equally on all three experiment groups (Control, Algorithms A and B), thus this factor is negligible. Furthermore, the average aggression levels were lower than that of the control group, suggesting the validity of the experiment design.

### 3.4 Comparison to Existing Mechanisms

Many CMC services implement user moderation, trying to attenuate aggressiveness of the comments posted by the users. For instance, Slashdot1, textream2, and YouTube3 are CMC services which provide mechanisms where users moderate each other’s comments, and every user can see the assessments of their comments. The drawback of these CMC services is that anonymous user accounts cannot be used because the titles are assigned by other users, which automatically requires the disclosure of the user account. Slashdot, for instance, displays users’ behaviors evaluated by other users and the titles (‘carma points’) which reflect users’ behavior. The titles in these CMC services are generated using algorithms developed by the service providers, usually based on other users’ evaluations and the contents of comments written by the user. The title is a reminder to users that their comments are observed by other users.

Basically, the mechanisms employed by the CMC services belong to the two categories. (1) User confirmation of the posted comments. When the user posts a comment, the user is requested to review her comment before actually submitting. This mechanism is employed in 2-channel4. (2) Automatic detection of non-desirable words. When the user posts a comment, the service invokes a natural language processing module

---

1. https://slashdot.org  
2. http://textream.yahoo.co.jp  
3. https://www.youtube.com/  
4. http://www.2ch.net
in the background to check the comment [10]–[14], and if inappropriate or aggressive terms are detected, it blocks the comment. For instance, textstream employs this mechanism. These techniques are, however, ineffective because the users can easily avoid them. If the user is familiar with these methods, the user clicks the button automatically in the first method, or avoid using inappropriate words that are detected by the system.

A considerable amount of people prefer using anonymous social network services (SNS) than autonym SNS [15], with the varying range depending on the nationality. The lowest is the UK (33%), and the highest is Japan (66.3%). Therefore, an effective mechanism to alleviate aggressions in anonymous CMC services is necessary.

The proposed method is easily implemented as it requires no algorithm to calculate the title and is suitable for anonymous CMC services, solving the deficiencies of conventional CMC services.

4. Conclusion

The purpose of this study is to investigate if simply displaying the title to the users can alleviate aggressions on CMC services, and to measure the effects on reduction of aggressiveness compared to an algorithmic calculation of titles, which is the mechanism used in currently available CMC services. Displaying the title to users is a mechanism to increase the public self-consciousness and to reduce private self-consciousness of users of pseudonym type CMC services.

Our analysis indicates that a simple title assignment is sufficient to reduce aggressiveness, presenting reduction effects equivalent to the accurate title calculated based on the users’ behavior. It was possible to use accurate aggressiveness values in our experiments, but it is difficult in actual CMC services, and no definitive algorithm exists to precisely calculate the aggressiveness of comments posted by the users. Furthermore, the development of new algorithms to estimate aggressiveness might be unnecessary.

Acknowledgments

This research was supported in part by JSPS Grants-in-Aid for Scientific Research Grant Number JP26280118.

References

[1] P.F. Braine: Hormonal aspects of aggression and violence, Understanding and Control of Biobehavioral Influences on Violence, Vol. 2, pp. 177–244, National Academy Press, 1994.
[2] C.L. Schoffstall and R. Cohen: Cyber aggression: The relation between online offenders and offline social competence, Social Development, Vol. 20, No. 3, pp. 587–604, 2011.
[3] S. Hinduja and J.W. Patchin: Cyberbullying: An exploratory analysis of factors related to offending and victimization, Deviant Behavior, Vol. 29, pp. 129–156, 2008.
[4] M.L. Ybarra and K.J. Mitchell: Youth engaging in online harassment: Associations with care-giver child relationships, Internet use, and personal characteristics, Journal of Adolescence, Vol. 27, pp. 319–336, 2004.
[5] M.L. Ybarra, M. Diener-West, and P.J. Leaf: Examining the overlap in Internet harassment and school bullying: Implications for school intervention, Journal of Adolescent Health, Vol. 41, pp. S31–S41, 2007.
[6] H. Yamaguchi and T. Maeshiro: A mechanism to control aggressive comments in pseudonym type computer mediated communications, HIMI 2016, Part II, LNCS 9735, pp. 90–100, 2016.
[7] H. Yamaguchi and T. Maeshiro: Effect of random title assignment to control aggressive comments in pseudonym type computer mediated communications, Proceedings of the SICE Annual Conference 2016, pp. 842–847, 2016.
[8] E. Matsuda and K. Okada: Interpersonal relationships on computer-mediated communication: The psychological analysis of interpersonal trust, aggression and emotion-regulation skills, Journal of Multimedia Aided Education Research, Vol. 2, No. 1, pp. 159–173, 2005.
[9] A. Ando, S. Soga, K. Yamasaki, S. Shimai, H. Shimada, N. Usuki, O. Ohashi, and A. Sakai: Development of the Japanese version of the Buss-Perry aggression questionnaire (BAQ), The Japanese Journal of Psychology, Vol. 70, No. 5, pp. 384–392, 1999.
[10] E. Spertus: Automatic recognition of hostile messages, Proceedings of the Conference on Innovative Applications of Artificial Intelligence, pp. 1058–1065, 1997.
[11] S.D. Durbin, J.N. Richter, and D. Warner: A system for affective rating of texts, Proceedings of OTC-03, 3rd Workshop on Operational Text Classification, 2003.
[12] A. Esuli and F. Sebastiani: SentiWordNet: A publicly available lexical resource for opinion mining, Proceedings of LREC 2006, 5th Conference on Language Resources and Evaluation, pp. 417–422, 2006.
[13] Z. Fei, J. Liu, and G. Wu: Sentiment classification using phrase patterns, Proceedings of the Fourth International Conference on Computer and Information Technology (CIT’04), pp. 1147–1152, 2004.
[14] G. Lesched and J. Kaye: Understanding how bloggers feel: Recognizing affect in blog posts, Extended Abstracts of CHI 2006, pp. 1019–1024, 2006.
[15] Ministry of Internal Affairs and Communications, White Paper Information and Communications in Japan 2014, 2014.

Tetsuya Maeshiro (Member)

He received a Doctor of Information Science degree from Japan Advanced Institute of Science and Technology. He is currently an Associate Professor of the School of Library, Information and Media Studies, the University of Tsukuba. His research interests include bioinformatics, kansei, and wisdom science.

Hiroki Yamaguchi

He received M.S. degree from the University of Tsukuba in 2016. His research interests include human-system interactions and socio-informatics.