Statistical Analysis of Automatic Seed Word Acquisition to Improve Harmful Expression Extraction in Cyberbullying Detection

Suzuha Hatakeyama[1,*], Fumito Masui[1], Michal Ptaszynski[1], Kazuhide Yamamoto[2]

[1] Department of Computer Science, Kitami Institute of Technology, Kitami, Japan.

[2] Department of Electrical Electronics, and Information Engineering, Nagaoka University of Technology, Nagaoka, Japan.

Abstract

We study the social problem of cyberbullying, defined as a new form of bullying that takes place in the Internet space. This paper proposes a method for automatic acquisition of seed words to improve performance of the original method for the cyberbullying detection by Nitta et al. [1]. We conduct an experiment exactly in the same settings to find out that the method based on a Web mining technique, lost over 30% points of its performance since being proposed in 2013. Thus, we hypothesize on the reasons for the decrease in the performance and propose a number of improvements, from which we experimentally choose the best one. Furthermore, we collect several seed word sets using different approaches, evaluate and their precision. We found out that the influential factor in extraction of harmful expressions is not the number of seed words, but the way the seed words were collected and filtered.

References

[1] T. Nitta, F. Masui, M. Ptaszynski, Y. Kimura, R. Rzepka, and K. Araki, "Detecting cyberbullying entries on informal school websites based on category relevance maximization," Proc. of the 6th International Joint Conference on Natural Language Processing (IJCNLP'13), Oct. 2013, pp. 579-586.

[2] A. Kilgarriff, "Googleology is bad science," Computational Linguistics, vol. 33, no. 1, pp. 147-151, 2007.

[3] T. Ishizaka and K. Yamamoto, "Automatic detection of sentences representing slandering on the Web," Proc. of the 17th Annual Meeting of the Association for Natural Language Processing, 2011, pp. 131-134. (In Japanese)

[4] P. D. Turney, "Thumbs up or thumbs down? Semantic orientation applied to unsupervised classification of reviews," Proc. of ACL-02, Jul. 2002, pp. 417-424.

[5] Q. McNemar, "Note on the sampling error of the difference between correlated proportions or percentages," Psychometrika, vol. 12, no. 2, pp. 153-157, 1947.

* Corresponding author. E-mail address: hatakeyama@ialab.cs.kitami-it.ac.jp

Tel.: +81-0157-26-9349