Detecting the Spam Review Using Tri-training

Ji Chengzhang*, Dae-Ki Kang**

*Weifang University of Science & Technology, Weifang 262700, China **Dongseo University, 47 Churye Ro, Sasang-Gu, Busan 617-716, Republic of Korea jcz8888@163.com, dkkang@dongseo.ac.kr

Abstract— Some supervised learning methods were developed to detect spam review and some of them are considerably effective. Some researchers also find that the review spammer consistently produce spam reviews. We observe that the spamming store also consistently produce spam reviews. This provides us two other views to identify review spam: we can identify if the reviewer is spammer and if the store is spamming one. We introduce a three-view semi-supervised method, tri-training, to exploit the large amount of unlabeled data. The experiment results demonstrate that three-view tri-training algorithm can achieve better results than two-view co-training and single-view algorithm.

Keywords- Deceptive reviews, semi-supervised learning, supervised learning, tri-training



Ji Chengzhang He received the B.S. degree in computer science and technology from Linyi University, Linyi, Shandong Province, China in 2005, and received the M.S. degree in computer application technology from Ocean University of China, Qingdao, Shandong Province, China in 2010. He had been working as an Assistant or Lecturer in Weifang University of Science & Technology, China from 2005 to 2015. Currently he is a doctoral candidate in machine learning at Dongseo University, Korea. His research interests include machine learning, data mining, spam review detection.



Dae-Ki Kang is an Associate Professor at Dongseo University in Korea. He was a Senior Member of the Engineering Staff at the Associated Institute of Electronics & Telecommunications Research in South Korea. He earned a PhD in Computer Science from Iowa State University in 2006. He received a Master of Science degree in computer science from Sogang University in 1994 and a BE in Computer Science and Engineering from Hanyang University in 1992.