

Wrapper Induction of News Information for Feeding to Social Networking Service on Smartphone

Zhong-Liang Xiang*, Xiang-Ru Yu*, Dae-Ki Kang**

**Computer Software Institute, Weifang University of Science & Technology, Shouguang city, Shandong Province, China*

***Division of Computer and Information Engineering, Dongseo University, Busan city, South Korea*

ugood@163.com, vuxiangru1119@163.com, dkkang@dongseo.ac.kr

Abstract—In this paper, we propose NewsFeedAndroid, a novel system that interconnects a social networking service and online newspaper sites in order to extract news articles from the online news sites and to perform feeding of news articles to social network service (SNS) users. In NewsFeedAndroid, news information agents extract news article information from the news and portal sites using Minimum Description Length (MDL) wrapper induction algorithm. The news document collecting module regularly gathers news list information from news list page in the news sites and portals. In the collected documents, the document preprocessing module removes tags that are unnecessary for news information extraction. Lexical analyzer converts the rest text information and tags to a sequence of tokens, and news information is obtained by matching token patterns to the sequence. Those extracted news information from the various sites are integrated in the system and supplied to the end users through the social networking service on a smartphone. NewsFeedAndroid demonstrates a novel usage of integrating social networking services and online newspaper sites.

Keyword—NewsFeedAndroid, Minimum description length, Smartphone, Cellphone, Social network service, Wrapper, News



Zhong-Liang Xiang is a candidate Ph.D. student in computer science at Dongseo University in South Korea. He received a science master degree in computer science at Ocean University of China in 2010 and a Bachelor of Science (BS) degree in computer science at Mudanjiang Normal University in 2003. His research interests include data mining and machine learning.



Xiang-Ru Yu received a science master degree in computer science at Ocean University of China in 2010 and a Bachelor of Science (BS) degree in computer science at Mudanjiang Normal University in 2003. Currently, she is a Lecturer at Weifang University of Science and Technology. Her research interests include data mining and machine learning.



Dae-Ki Kang is a professor at Dongseo University in South Korea. He was a senior member of engineering staff at the attached Institute of Electronics and Telecommunications Research Institute in South Korea. He earned a Ph.D. in computer science from Iowa State University in 2006. His research interests include intrusion detection, security informatics, ontology learning, and relational learning. Prior to joining Iowa State, he worked at a Bay-area startup company and at the Electronics and Telecommunication Research Institute in South Korea. He received a science master degree in computer science at Sogang University in 1994 and a bachelor of engineering (BE) degree in computer science and engineering at Hanyang University in 1992.