# 5 What is the author like?

Estimation of microblog user attributes using microblog structure

Abstract— Word-of-mouth marketing on microblogs (e.g. Twitter) has become more important given the increasing number of users and posts. However, unlike questionnaire surveys, there is a problem in that most users do not specify their attributes (gender, age, etc.) in their profiles. We resolve this issue by employing user attribute estimation. Existing methods use bag-of-words features extracted from manually labeled users' posts to estimate user attributes. We propose a new method that focuses on the microblog structure and the topics of words in the user's posts. This method is very accurate in estimating the major marketing attributes; e.g. it is about 90% accurate for gender. Our method is designed for marketing use, but we intend to enhance it so that it can estimate various and more segmented attributes for personalization use.
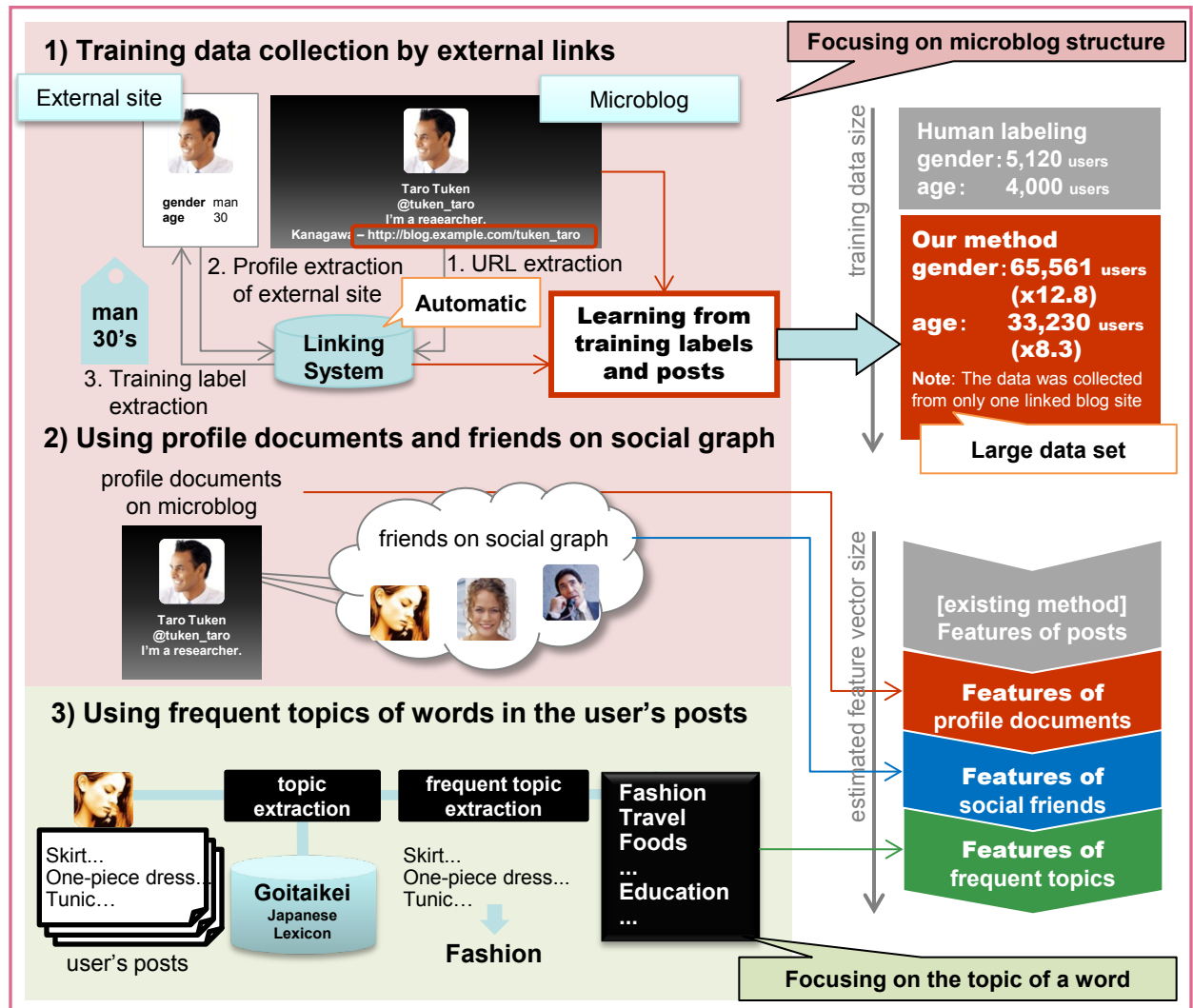
Related works

[1] J. Ito, K. Nishida, T. Hoshide, H. Toda, T. Uchiyama, "Estimation of Twitter User Attributes by Learning from Users who have both Twitter and Blog Accounts and Utilizing User Homophily," in *Proc. the 5th Forum on Data Engineering and Information Management (DEIM)*, 2013. (in Japanese)

Contact

Jun Ito
Service Harmonization Project
NTT Service Evolution Laboratories
E-mail : ito.jun{at}lab.ntt.co.jp

Tohru Hirano
Audio, Speech, and Language Media Project
NTT Media Intelligence Laboratories
E-mail : hirano.tohru{at}lab.ntt.co.jp

(Please replace {at} with @)