

**Title:**

Specifying node characteristics by combining social network data and user-generated-content

**Authors:**

Mirai Igarashi (presenter)

Nobuhiko Terui (co-author)

**Abstract:**

[Background]

In modern social media development, the influence of interaction between consumers on their behavior is getting bigger and it is important for companies to know the relationship between consumers through understanding their characteristics on social network.

[Methodology]

Many methods for specifying people's characteristics on the network have been proposed on only the network information. However, not all people are well-connected on the network with the same interest. For example, even if students belong to the same community of "school", they have various hobbies such as music, books, and sports. In addition, people belong to multiple communities such as family, work, and online friends. Then it is more realistic and beneficial for company to identify community according to their interests uncovered by their communications in social media. In this study, by extending mixed membership stochastic block model, we propose a model for estimating characteristics by combining social network data and user-generated-content (UGC) data. The model explores multiple characteristics according to their topics which are identified by using network and UGC data jointly. Also, the model will be extended to accommodate the dynamical change of the characteristics and topics on social network.

[Results]

an empirical analysis using Twitter data clarifies that our model can find realistic and meaningful community structures from large social networks and has a good predictive performance.

**Keywords:**

Social network analysis, Community detection, User-generated-content, Topic modeling, Bayesian inference