



Estimation of topics and social influences on social media contents

Department of Management, Graduate School of Economics

Associate Professor Mirai Igarashi



https://researchmap.jp/mirai_igarashi?lang=en

Abstract

This study aims to represent user-generated content (UGC) posted on social media using a mathematical framework known as a topic model. This approach enables the estimation of (1) topic distributions, how much each topic (e.g., music, fashion, technology) is reflected in UGC, and (2) social influence, how such content propagates through social media networks. An empirical analysis using Pinterest image data demonstrates that the proposed model outperforms conventional topic models in both predictive accuracy and topic interpretability. Furthermore, by identifying influential users for each estimated topic, the study shows that the efficiency of Seeded Marketing Campaigns (SMCs) can be significantly improved.

Background & Results

The user-generated content (UGC) posted on social media platforms contains rich information about users' interests, identities, and behavioral patterns. Marketers utilize such data to understand market trends and, at times, to actively intervene in them. However, leveraging UGC data presents several challenges: it requires appropriate handling of unstructured data such as text and images, and it involves disentangling complex social influences embedded in content that reflect not only individual sentiments but also network-driven effects. To address these challenges, this study extends a class of topic models by incorporating social influence effects that vary across topics. This extension enables the simultaneous estimation of (1) topic distributions in UGC (i.e., what topics are present and to what extent) and (2) topic-specific social influence across users within a network. An empirical analysis using Pinterest image data demonstrates that the proposed model outperforms conventional topic models in both predictive accuracy and topic interpretability. As a result, marketers can use an interpretable model to better capture market trends and design efficient production plans. Furthermore, in simulation experiments of Seeded Marketing Campaigns (SMCs), which aim to stimulate word-of-mouth diffusion by sending products to selected users, the proposed approach was shown to enhance campaign efficiency for certain topics.

Traditionally, research on efficiency of SMCs has assumed that users with a large number of followers (i.e., influencers) are the most effective investment targets, a view also common in practice. However, this study suggests that considering topic-specific variations in social influence can further improve the efficiency of marketing interventions, providing new insights for influencer selection and campaign strategy design.

Significance of the research and Future perspective

The significance of this study lies in its potential to contribute to areas such as market trend analysis, demand forecasting, and efficient production planning in manufacturing industries like apparel, by improving predictive accuracy while maintaining interpretability. In addition, enhancing the efficiency of Seeded Marketing Campaigns (SMCs) through targeting strategies that account for social influence provides practical insights for developing strategies to intervene in market trends. Looking ahead, extending the model to handle larger-scale data is expected to further advance user profiling on social media platforms.

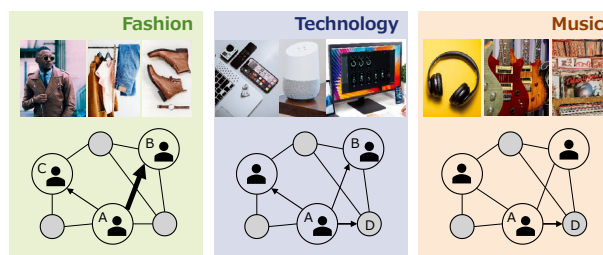


Figure 1 Topic-heterogeneity of social influence

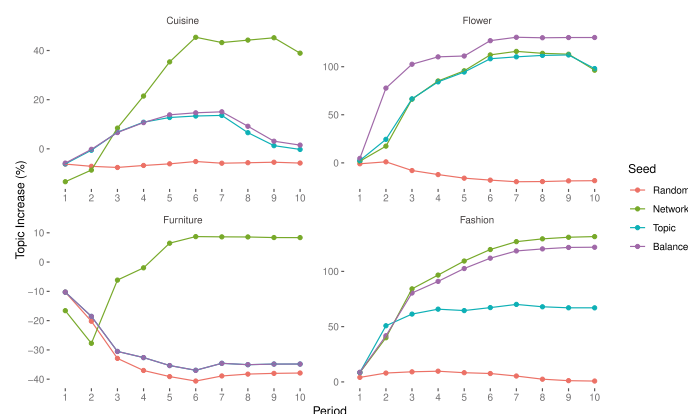


Figure 2 Comparison of topic increase in SMCs simulation

Social innovation

Patent
Treatise
URL
Keyword

Igarashi, Mirai; Zhang, Kunpeng; Terui, Nobuhiko et al. Identifying influential users by topic in unstructured user-generated content. Production and Operations Management. 2025, 34(10), 3267-3288. doi: 10.1177/10591478251332335

user generated content, social networks, social influence, topic models, influencer