

Guest editorial: content, concept and context mining in social media

Heng Tao Shen · Xian-Sheng Hua · Jiebo Luo · Vincent Oria

Received: 21 July 2011 /

Accepted: 27 July 2011 /Published online: 10 August 2011

© Springer Science+Business Media, LLC 2011

In the last few years, we have witnessed an exponentially growing amount of social media data from social networks, mainly driven by the huge popularity of Web and the rapid advances of multimedia and Internet technologies. Different from conventional data types, social media data are multi-model in nature, including content such as images, audio, and videos, concept such as discussion topic, tag, and annotation, and context such as links, profile, timestamp, and click-through. All these content, concept and context are essential data sources for mining semantics, user intents, trends, knowledge, etc., from social media data, as well as enabling intelligent applications on social media services.

This rich media type has raised many new research challenges, ranging from large-scale social media content analysis and mining, concept discovery and monitoring, context-based services, to many other exciting new opportunities. The aim of this special issue is to bring out state-of-the-art research in this area and discover directions for future research.

The special issue will serve as a forum for recent advances in the field, with an emphasis on mining the “3 C” (Content, Concept and Context). It received an enthusiastic response. Among 19 submissions, 5 papers were selected after several rounds of rigorous review by the guest editors and the invited reviewers.

H. T. Shen (✉)

The University of Queensland, Brisbane, Australia
e-mail: shenht@itee.uq.edu.au

X.-S. Hua

Microsoft Corporation, One Microsoft Way, Redmond, WA 98052, USA
e-mail: xshua@microsoft.com

J. Luo

Intelligent Systems, Kodak Research Laboratories, Eastman Kodak Company,
1999 Lake Avenue, Rochester, NY 14650, USA
e-mail: jiebo.luo@kodak.com

V. Oria

Department of Computer Science, New Jersey Institute of Technology, Newark, NJ 07102, USA
e-mail: oria@njit.edu

The paper by Zhang et al., titled “Opinion Helpfulness Prediction in the Presence of Words of Few Mouths”, identifies a widely existing phenomenon in social media content called the “words of few mouths” phenomenon and proposes probabilistic approaches for recommender system development in the presence of “words of few mouths”. The advantages of incorporating probabilistic methods in the training of the review helpfulness predictors are demonstrated.

The paper by Angelova et al., titled “Graffiti: Graph-based Classification in Heterogeneous Networks”, addresses the problem of multi-label classification in heterogeneous graphs representing the social data on the Web and proposes a novel approach to classify nodes by exploiting the mutual influence of nodes of the same type based on their connections to nodes of other types.

The paper, by Yao et al., titled “Bursty Event Detection from Collaborative Tags”, aims at monitoring and analysing the temporal patterns of tags for online social resources to trace hot topics on the Web. A new approach to detect bursty tagging event which corresponds to a real life event is proposed, by utilizing time intervals, graph clustering techniques, and tag taxonomy

The paper, by Wu et al., titled “Boosting Web Video Categorization with Contextual Information from Social Web”, explores Web video categorization from a new perspective, by integrating the model-based and data-driven approaches to boost the performance. The boosting comes from query expansion and integration of semantics, relevance and contextual information on social videos.

The paper, by Bisgin et al., titled “A Study of Homophily on Social Media”, studies the homophily issue in existing sociology literature on social media and proposes a systematic approach by studying three online social media sites, Blog Catalog, Last.fm, and LiveJournal. It reports the findings along with some interesting observations. The results reveal that the influence of interest-based homophily is not a very strong leading factor for constructing new ties in social media sites.

The five papers included in this special issue cover several important topics, from content mining, concept classification and detection, to context utilization and homophily study. They present some of the latest results in this active and rapidly expanding research area. We sincerely hope that the set of selected papers provides the community with some insights on the current research and some inspirations on its future trends.