Workshop on Multimodal Crowd Sensing (CrowdSens 2012)

Haggai Roitman
IBM Research Haifa
31905 Haifa, Israel
haggai@il.ibm.com

Iván Cantador
Universidad Autónoma de Madrid
28049 Madrid, Spain
ivan.cantador@uam.es

Miriam Fernández
Knowledge Media Institute
MK7 6AA Milton Keynes, UK
m.fernandez@open.ac.uk

ABSTRACT
This paper provides an overview of the 1st International Workshop on Multimodal Crowd Sensing (CrowdSens 2012), held at the 21st ACM International Conference on Information and Knowledge Management (CIKM 2012). This workshop aimed to provide an open forum for researchers from various fields such as fields such as Natural Language Processing, Information Extraction, Data Mining, Information Retrieval, User Modeling and Personalization, Stream Processing, and Sensor Networks, for addressing the challenges of effectively mining, analyzing, fusing, and exploiting information sourced from multimodal physical and social sensor data sources.

Categories and Subject Descriptors
H.4: Information Systems applications

General Terms
Algorithms, Experimentation, Performance, Human Factors.

Keywords
Social networks, social media, web mining, sensor data.

1. MOTIVATION
The ubiquitous availability of computing technology, in particular smartphones, tablets, laptops and other easily portable devices, and the adoption of social networking sites, make it possible to be connected and continuously contribute to a massively distributed information publishing process. By doing so, users are (unconsciously) acting as social sensors, whose sensor readings are their manually generated data. People document their daily life experiences, report on their physical locations and social interactions with others, express opinions and provide diverse observations on both the physical world (sights, sounds, smells, feelings, etc.) and the online world (news, music, events, etc.). Such massive amounts of ubiquitous social sensors, if wisely utilized, can provide new forms of valuable information that are currently not available by any traditional data collection methods including real physical sensors, and can be used to enhance decision making processes.

It has been shown over and over that reports on real world events – such as the Japan's Earthquake and Tsunami, the Arab Spring uprisings, and the England's riots happened in 2011 – are much faster propagated within networks of social sensors (e.g. on Twitter) than traditional information processes (e.g. seismic sensor reading analysis, police emergency reports, and news media coverage). Moreover, in these cases human observers can further enrich and interpret the integrated sensor-derived information. As an example, nowadays both reviewers and journalists make increasing usage of massive data collected from social sensors in order to study public opinions, and discover new perspectives of daily stories.

In this context effective mining, analyzing, fusing, and exploiting information from multimodal physical and social sensor data sources is still an open and exciting challenge. Many factors are included into the complexity of the problem, such as the real-time requirements of the data processing; the heterogeneity of the data sources, from physical sensors data to posts on social media; and the ubiquitous and noisy nature of the human-sensor generated information, which can be written in an informal style, and can be redundant, incomplete, or even incorrect.

2. GOALS AND TOPICS OF INTEREST
The main goal of CrowdSens 2012 is to become a major international forum for researchers and practitioners from different research areas who focus their work on user-generated contents. The workshop thus aimed to stimulate discussions about how the knowledge embedded in human sensor data can be collected, extracted, modeled, analyzed, integrated, summarized, and finally exploited.

The workshop has six main research themes. The first theme is focused on *data acquisition methods for crowd sensing*, and includes topics such as real-time data acquisition methods, massive scale social network monitoring and crawling, and predictive models for social data acquisition. The second theme is focused on *data models for crowd sensing*. It includes topics such as social sensor data representation, social sensor events models, and semantic models for crowd sensing. The third theme is focused on *data processing, analysis and classification methods*. It includes topics such as data mining methods under incomplete and noisy data, crowd behavioral analysis and prediction, and real-time community detection and analysis. The fourth theme is focused on *event detection, fusion and summarization methods*. It includes topics such as detection of developing events, event uncertainty estimation, and multimodal data fusion methods. The fifth theme is focused on *evaluation methods for crowd sensing*. It includes topics such as quality metrics, performance indicators, benchmarks and evaluation methodologies for crowd sensing. Finally, the sixth theme is focused on *applications of crowd sensing*, and includes topics such as government and politics, news mining from social sensors, and public safety and health.

3. PROGRAM COMMITEE

The following is the list of the workshop's PC members:

- **Harith Alani**, Knowledge Media Institute, UK
- **Ching-man Au Yeung**, Hong Kong Applied Science and Technology Research Institute, China
- **Pablo Castells**, Universidad Autónoma de Madrid, Spain
- **Federica Cena**, Universita' degli Studi di Torino, Italy
- **Aba-Sah Dadzie**, University of Sheffield, UK
- **Ido Guy**, IBM Research - Haifa, Israel
- **Vanessa Lopez**, IBM Dublin, Ireland
- **Jonathan Mamou**, IBM Research - Haifa, Israel
- **Yosi Mass**, IBM Research - Haifa, Israel
- **Jérôme Picault**, Bell Labs, Alcatel-Lucent, France
- **Louiqa Raschid**, Maryland University, USA
- **Matthew Rowe**, Knowledge Media Institute, UK
- **Claudia Wagner**, Institute of Information Systems - Joanneum Research, Austria

4. PROGRAM

The workshop program included two invited talks, two sessions for presenting accepted publications, and a final panel for open discussion. Next we provide a few details about the invited talks and accepted papers.

4.1 Invited Talks

The workshop had two invited talks. The first talk was given by Ido Guy, social technology manager at IBM Haifa Research Lab, and was titled "Crowdsourcing in the enterprise." The talk focused on today's concept of Social Business, and how the emerging technologies exploiting social information help increasing the employee's engagement.

The second invited talk, by Manuel Cebrián, senior research fellow at the Media Laboratory at the Massachusetts Institute of Technology, was titled "Using friends as sensors to detect planetary-scale contagious outbreaks." The talk was about an analytical model of the contagious spread of information at a large scale. This model can help yielding not just early detection, but also advancing warning of contagious outbreaks.

4.2 Accepted Papers

A total of 9 papers were submitted to the workshop, out of which 3 were accepted as full (6-8 page-long) papers, and 2 as short (2-4 page-long) papers. The topics of the papers spanned from conceptual modeling principles for crowdsourcing till brand perception in social media. In the following we shortly describe each of the accepted papers:

The full paper by Packer et al. [Packer et al., 2012] presents an approach to detect events on Twitter by expanding queries with additional terms leveraged from structural data. The novelty of this approach is focused on estimating the relevance of a tweet to a particular entity in order to determine whether the tweet can be used as a social sensor for the event.

The full paper by Saeidi [Saeidi, 2012] discusses a transitive vote delegation scheme over social networks. The vote delegation mechanisms are based on graph analysis and processing algorithms such as random walks and clustering.

The full paper by Lukyanenko and Parsons [Lukyanenko & Parsons, 2012] discusses the use of crowdsourcing for conceptual modeling, the challenges posed and potential resolutions. It proposes the use of ontology-based modeling as a solution to counter the specific challenges due to heterogeneous (and potentially conflicting) perspectives on and interpretation of a domain by the different users/contributors encountered when relying on crowdsourcing as a source of information. The connection between the semantic web and crowdsourcing is a key discussion bring by this paper.

The short paper by Roitman and colleagues, [Roitman et al., 2012] presents a prototype for extracting events from tweets. The system is meant to be part of the *smartcity*, to enable officials to learn about events in the city from citizens and eyewitnesses. The paper describes the grand vision and report on some preliminary results from analyzing tweets collected for the city of New York in one weekend.

The short paper by Shmueli-Scheuer and colleagues [Shmueli-Scheuer et al., 2012] describes a system for mining brand perception from twitter. The paper highlights the architecture and functionality of the system, and describes some of the technical and computational steps the system takes when processing tweets for monitoring brands.

5. ACKNOWLEDGMENTS

6. REFERENCES

[1] Packer, H.S., Samangooei, S., Hare, J.S., Gibbins, N. and Lewis, P. Event Detection using Twitter and Structured Semantic Query ExpansionForman, G. 2003. In Proceedings of the 1st International Workshop on Multimodal Crowd Sensing (Crows Sens2012).

[2] Lukyanenko, R. and Parsons, J. Conceptual Modeling Principles for Crowdsourcing. In Proceedings of the 1st International Workshop on Multimodal Crowd Sensing (CrowsSens 2012).

[3] Roitman, H., Mamou, J., Mehta, S., Satt, A. and Subramaniam, L.V. Harnessing the Crowds for Smart City Sensing. In Proceedings of the 1st International Workshop on Multimodal Crowd Sensing (CrowsSens 2012).

[4] Saeidi, Z. Algorithm for Representative Democracy Voting in Social Network. In Proceedings of the 1st International Workshop on Multimodal Crowd Sensing (CrowsSens 2012).

[5] Shmueli-Scheuer, M., Sznajder, B., Cohen, D., Raviv, A. and Konopnicki, D. "Greaaaat bargains starting from just 99p!!!! :-)" Brand Perception in the Social Media. In Proceedings of the 1st International Workshop on Multimodal Crowd Sensing (CrowsSens 2012).