

# Accepted Manuscript

Detecting Prominent Microblog Users over Crisis Events Phases

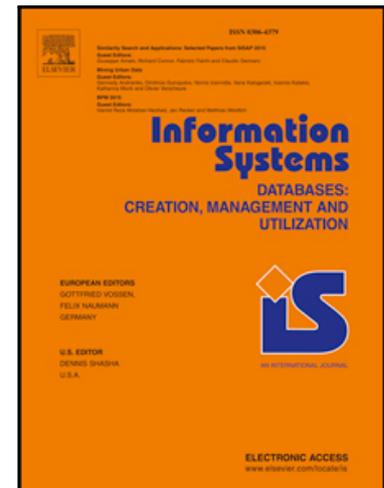
Imen Bizid, Nibal Nayef, Patrice Boursier, Antoine Doucet

PII: S0306-4379(17)30226-0  
DOI: [10.1016/j.is.2017.12.004](https://doi.org/10.1016/j.is.2017.12.004)  
Reference: IS 1269

To appear in: *Information Systems*

Received date: 14 April 2017  
Revised date: 11 December 2017  
Accepted date: 15 December 2017

Please cite this article as: Imen Bizid, Nibal Nayef, Patrice Boursier, Antoine Doucet, Detecting Prominent Microblog Users over Crisis Events Phases, *Information Systems* (2017), doi: [10.1016/j.is.2017.12.004](https://doi.org/10.1016/j.is.2017.12.004)



This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

### Highlights

- A novel phase-aware microblog users modeling approach is proposed
- Prominent microblog users are predicted at an early stage of each event phase
- Our model is experimented using real-world data
- Direct access to the relevant and exclusive information during crisis events

ACCEPTED MANUSCRIPT

# Detecting Prominent Microblog Users over Crisis Events Phases

Imen Bizid<sup>a,1</sup>, Nibal Nayef<sup>a</sup>, Patrice Boursier<sup>a</sup>, Antoine Doucet<sup>a</sup>

<sup>a</sup>*L3i, University of La Rochelle*

---

## Abstract

During crisis events such as disasters, the need for real-time information retrieval (IR) from microblogs becomes essential. However, the huge amount and the variety of the shared information in real time during such events over-complicates this task. Unlike existing IR approaches based on content analysis, we propose to tackle this problem by using user-centric IR approaches with identifying and tracking prominent microblog users who are susceptible to share relevant and exclusive information at an early stage of each analyzed event phase. This approach ensures real-time access to the valuable microblogs information required by the emergency teams. In this approach, we propose a phase-aware probabilistic model for predicting and ranking prominent microblog users over time according to their behavior using Mixture of Gaussians Hidden Markov Models (MoG-HMM). The model utilizes a new user representation which takes into account both the user and the event specificities over time. This user representation comprises the following new aspects (1) Modeling microblog users behavior evolution by considering the different event phases (2) Characterizing users activity over time through a temporal sequence representation (3) Time-series-based selection of the most discriminative features (4) prominent users prediction using probabilistic phase-aware models learned *a priori*. We have conducted experiments during flooding events: we trained our identification models using a dataset relative to the “Alpes-Maritimes floods” and we tested its identification performance using a new dataset relative to another flooding disaster “Herauld floods”. The achieved results show that our model significantly outperforms phase-unaware models and identifies most of the prominent users at an early stage of each event phase.

*Keywords:* Information Retrieval from Microblogs, Prominent Users Prediction, User Behavior Modeling, Phase-Temporal Representation

---

*Email addresses:* [imen.bizid@univ-lr.fr](mailto:imen.bizid@univ-lr.fr) (Imen Bizid), [nibal.nayef@univ-lr.fr](mailto:nibal.nayef@univ-lr.fr) (Nibal Nayef), [patrice.boursier@univ-lr.fr](mailto:patrice.boursier@univ-lr.fr) (Patrice Boursier), [antoine.doucet@univ-lr.fr](mailto:antoine.doucet@univ-lr.fr) (Antoine Doucet)