

Twitter Streaming Data Analytics for Disaster Alerts

Syed Attique Shah,
*Department of Information
Technology
BUIITEMS, SDSS
Lab — NCGSA,
Quetta, Pakistan*
attique.shah@buitms.edu.pk
0000-0003-2949-7391

Akhtar Jamil
*Department of Computer Science
National University of Computer and
Emerging Sciences,
Islamabad, Pakistan*
0000-0002-2592-1039

Sadok Ben Yahia
*Department of Software Science
Tallinn University of Technology,
Tallinn, Estonia*
sadok.ben@taltech.ee

Keegan McBride
*Postdoctoral Researcher
Hertie School Centre for Digital
Governance,
Berlin, Germany*
mcbride@hertie-school.org

Dirk Draheim
*Information Systems Group Tallinn
University of Technology,
Tallinn, Estonia*
dirk.draheim@taltech.ee

Abstract—In today’s world, disasters, both natural and man-made, are becoming increasingly frequent, and new solutions are of a compelling need to provide and disseminate information about these disasters to the public and concerned authorities in an effective and efficient manner. One of the most frequently used ways for information dissemination today is through social media, and when it comes to real-time information, Twitter is often the channel of choice. Thus, this paper discusses how Big Data Analytics (BDA) can take advantage of information streaming from Twitter to generate alerts and provide information in real-time on ongoing disasters. The paper proposes TAGS (Twitter Alert Generation System), a novel solution for collecting and analyzing social media streaming data in real time and subsequently issue warnings related to ongoing disasters using a combination of Hadoop and Spark frameworks. The paper tests and evaluates the proposed solution using Twitter data from the 2018 earthquake in Palu City, Sulawesi, Indonesia. The proposed architecture was able to issue alert messages on various disaster scenarios and identify critical information that can be utilized for further analysis. Moreover, the performance of the proposed solution is assessed with respect to processing time and throughput that shows reliable system efficiency.

Index Terms—Big Data Analytics, Disaster Management, Disaster Alert Generation, Twitter, Hadoop, Spark, Geo-social Media Analytics