Graph-Based Event Detection in Streams: The Twitter Case

Michalis Vazirgiannis

LIX, Ecole Polytechnique, France
mvazirg@lix.polytechnique.fr

Abstract

Due to its instantaneous nature, Twitter has been established as a major communication medium. Among others, people use the service to report latest news and to comment about real-world events. Users show particular interest in social events such as large parties, political campaigns, and sporting events, but also for emergency events such as natural disasters and terrorist attacks. Automated and real-time event detection in this case is an interesting challenge. We present our work on this topic capitalizing on modeling the stream as an evolving graph of words, and then detecting events based on their evolution patterns. To identify important moments, the system detects rapid changes in the graphs edge weights using a convex optimization formulation. Then we need to summarize the event in the best feasible way. We present a method that generates real-time summaries of events using only posts collected from Twitter. The system then extracts a few tweets that best describe the chain of interesting occurrences in the event using a greedy algorithm that maximizes a non-decreasing sub-modular function. Through extensive experiments on real-world sporting events, we show that the proposed system can effectively capture the sub-events, and that it clearly outperforms the dominant sub-event detection method.

Biography

Dr. Vazirgiannis is a Professor at LIX, Ecole Polytechnique in France and leads the Data Science and Mining group. He holds a degree in Physics and a PhD in Informatics from Athens University (Greece), and a Master degree in AI from HerioWatt Univ Edinburgh. He has conducted research in GMD-IPSI, Max Planck MPI (Germany), and in INRIA/FUTURS (Paris). He has been teaching in AUEB (Greece), Ecole Polytechnique, Telecom-Paristech, ENS (France), Tsinghua, Jiaotong Shanghai (China) and in Deusto University (Spain). His current research interests are on machine learning and combinatorial methods for Graph analysis (including community detection, graph clustering and embeddings, influence maximization), text mining including Graph of Words, word embeddings with applications to web advertising and marketing, and event detection and summarization. He has active cooperation with industrial partners in the area of data analytics and machine learning for large scale data repositories in

different application domains. He has supervised fifteen completed PhD dissertations. He has published three books and more than 160 papers in international refereed journals and conferences. He has organized large scale conferences in the area of Data Mining and Machine Learning (such as ECML/PKDD) while he participates in the senior PC of AI and ML conferences, e.g., AAAI and IJCAI. He has received the ERCIM and the Marie Curie EU fellowships, the Tencent "Rhino-Bird International Academic Expert Award" in 2017 and since 2015 he leads the AXA Data Science[1] chair. More information can be found at: `http://www.lix.polytechnique.fr/dascim`.

[1] http://www.lix.polytechnique.fr/dascim/dascis