Chapter 8 User Driven Query Framework of Social Networks for Geo-Temporal Analysis of Events of Interest

Gloria Bordogna CNR IREA, Italy

Simone Sterlacchini CNR IDPA, Italy

Paolo Arcaini *Charles University in Prague, Czech Republic*

ABSTRACT

In this chapter we propose a framework for collecting, organizing into a database and querying information in social networks by the specification of content-based, geographic and temporal conditions to the aim of detecting periodic and aperiodic events. Our proposal could be a basis for developing context aware services. For example to identify the streets and their rush hours by analyzing the messages in social media periodically sent by queuing drivers and to report these critical spatio-temporal situations to help other drivers to plan alternative routes. Specifically, we rely on a focused crawler to periodically collect messages in social networks related with the contents of interest, and on an original geo-temporal clustering algorithm in order to explore the geo-temporal distribution of the messages. The clustering algorithm can be customized so as to identify aperiodic and periodic events at global or local scale based on the specification of geographic and temporal query conditions.

INTRODUCTION

The widespread use of social media by smartphones and other smart devices with GPS sensors is fostering a novel era of spatio-temporal computing applications.

The US red cross has reported that the US citizens are increasingly relying on social media and mobile devices to get information on ongoing critical situations, such as traffic jams, spread of pandemies,

DOI: 10.4018/978-1-4666-8767-7.ch008

User Driven Query Framework of Social Networks for Geo-Temporal Analysis

and to seek assistance and safety information as well as to report their health and safety status during or after emergencies (Adam et al., 2012).

New location-based services and context-aware services can exploit social information sources for the most diverse applications in smart cities context, such as leisure recommendation, healthcare and safety, disaster management, critical periodic crisis identification, and so on.

Such services can exploit the information contents provided by users of social media, which can be in the form of free text, pictures and video, coupled with the timestamp and the geolocation as acquired by the GPS sensor of their device, to identify events occurred in specific regions at specific dates, and, depending on the kind of the event, to target either planning or alert responses.

In this chapter, we propose a framework for events' exploration based on querying a collection of social media messages collected by a focused crawler and organized into a database. The queries allow to specify spatio-temporal conditions to the aim of filtering and then analyzing periodic and aperiodic events at global or local scale. Our proposal could be a basis for developing context aware services. For example, to identify the streets and their rush hours by analyzing the messages in social media periodically sent by queuing drivers and to report these critical spatio-temporal situations to help other drivers to plan alternative routes.

Specifically, we rely on a focused crawler to periodically collect messages in social networks related to contents that may be of interest to categories of users, such as urban planners, who might need to know the streets in their city where most often traffic jams occur, territorial administrators and managers, who might need to identify in their territory of competence the periods when critical situation like floods usually occur, or cultural operators, who might be interested in planning the tour of musicians in specific regions and periods of the year more suitable to attract much audience.

In order to explore the collected messages, a query framework is proposed, consisting of two subsequent phases in order to allow the user to drive the exploration whose aim is to verify some "a priori" hypothesis he/she has on an event of interest.

In the first phase, the user can formulate queries specifying content, spatial and temporal conditions on the textual indexes and metadata parts of the messages, i.e., their geotags and timestamps, in order to filter out a subset of interesting messages from the collection. For example, an urban planner of "Bangkok" may want to select messages dealing with "traffic jams" in his/her city.

In the second phase, the user can specify some criteria in order to drive an original geo-temporal analysis of the selected messages to verify an "a priori" hypothesis. For example, the urban planner may hypothesize that traffic jams in "Bangkok" occur periodically in specific hours of the day, but may lack to know the streets and the hours of traffic jams. Thus, his/her criteria of analysis could be to consider a periodic geo-temporal distance so as to identify the most crowded hours of the day in each specific street.

An Italian cultural operator, in order to plan the tour of a musician, may be interested in selecting messages dealing with music events in Italy and in analyzing in which city in Italy and in which season the events had more resonance. In this case, he/she may formulate a query selecting messages dealing with entertainment events in Italy and then ask to group them in order to identify the cities and months where and when the most popular music events took place.

The grouping is performed by a geo-temporal clustering algorithm that can be flexibly customized to use a specific geo-temporal distance measure in order to identify aperiodic or periodic events at global or local scale. 24 more pages are available in the full version of this document, which may be purchased using the "Add to Cart" button on the publisher's webpage:

www.igi-global.com/chapter/user-driven-query-framework-of-social-networks-

for-geo-temporal-analysis-of-events-of-interest/138699

Related Content

The KnowledgeStore: A Storage Framework for Interlinking Unstructured and Structured Knowledge

Francesco Corcoglioniti, Marco Rospocher, Roldano Cattoni, Bernardo Magniniand Luciano Serafini (2018). *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications (pp. 686-721).*

www.irma-international.org/chapter/the-knowledgestore/198572

Named Entity Based Ranking with Term Proximity for XML Retrieval

Abubakar Roko, Shyamala Doraisamy, Azreen Azmanand Azrul Hazri Jantan (2018). *International Journal of Information Retrieval Research (pp. 57-77).* www.irma-international.org/article/named-entity-based-ranking-with-term-proximity-for-xml-retrieval/198965

Document Retrieval Using Efficient Indexing Techniques: A Review

Shweta Gupta, Sunita Yadavand Rajesh Prasad (2018). *Information Retrieval and Management: Concepts, Methodologies, Tools, and Applications (pp. 1745-1764).* www.irma-international.org/chapter/document-retrieval-using-efficient-indexing-techniques/198623

Distributed Database Management With Integration of Blockchain and Long Short-Term Memory

Siddesh G. M., S. R. Mani Sekhar, Vighnesh S., Nikhila Sai, Deepthi Saiand Sanjana D. (2021). International Journal of Information Retrieval Research (pp. 18-33).

www.irma-international.org/article/distributed-database-management-with-integration-of-blockchain-and-long-short-termmemory/280524

SAR: An Algorithm for Selecting a Partition Attribute in Categorical-Valued Information System Using Soft Set Theory

Rabiei Mamat, Tutut Herawanand Mustafa Mat Deris (2013). Information Retrieval Methods for Multidisciplinary Applications (pp. 266-280).

www.irma-international.org/chapter/sar-algorithm-selecting-partition-attribute/75912