Big Data Management and Analytics for Mobility Forecasting in datAcron

Christos Doulkeridis, Nikos Pelekis, Yannis Theodoridis, George A. Vouros
University of Piraeus
Piraeus, Greece

ABSTRACT

The exploitation of heterogeneous data sources offering very large historical and streaming data is important to increasing the accuracy of operations when analysing and predicting future states of moving entities (planes, vessels, etc.). This article presents the overall goals and big data challenges addressed by datAcron[1] on big data analytics for time-critical mobility forecasting.

Keywords

Big Data, Spatio-temporal Data, Moving Objects, Event detection, Trajectory, Prediction, Analytics.

1. INTRODUCTION – CHALLENGES

datAcron aims to advance the management and integrated exploitation of voluminous and heterogeneous data-at-rest (archival data) and data-in-motion (streaming data) sources, so as to significantly improve the capacities of systems to promote safety and effectiveness of critical operations for large numbers of moving entities in large geographical areas.

Challenges throughout the Big Data ecosystem with special focus on surveillance systems, concern effective detection and forecasting of moving entities' trajectories and events due to these trajectories. These challenges emerge as the number of moving entities and related operations increase at unprecedented scale. In conjunction with the demand for more and more frequent data from many different sources and for each of these entities, this results in generating vast data volumes, of heterogeneous nature, at extremely high rates, whose exploitation calls for novel big data techniques and algorithms that lead to advanced data analytics; this is a core research issue for datAcron.

More concretely, core research challenges in datAcron include:

- distributed management and querying of spatiotemporal RDF data-at-rest (archival) and data-in-motion (streaming), following an integrated approach;
- reconstruction and forecasting of moving entities' trajectories in the challenging Maritime (2D space) and Aviation (3D space) domains ;
- Recognition and forecasting of complex events and patterns due to the movement of entities (e.g. prediction of potential collision, capacity demand, hot spots / paths);
- interactive Visual Analytics for supporting human exploration and interpretation of the above phenomena.

2. OVERALL CONCEPT

The datAcron concept is demonstrated by the overall architecture towards providing a coherent Big Data solution, whose main components are as follows:

D*ata sources* include multiple streaming as well as archival data.

I*n-situ processing* components compress and integrate data at high rates of data compression without affecting the quality of analytics, capitalizing on primitive operators that are applied directly on the data streams.

Data transformation components convert data from disparate data sources as well as analytical results from the datAcron higher-level components to a common representation.

The *data integration*/*interlinking* component interlinks semantically annotated data using link discovery techniques for automatically computing associations between data from heterogeneous sources.

The *spatiotemporal query-answering* component provides parallel query processing techniques for spatio-temporal query languages over interlinked data stored in parallel RDF stores, using sophisticated RDF partitioning algorithms.

The *data analytics* components include trajectory and complex event recognition and forecasting, as well as visual analytics.

3. VALIDATION & EVALUATION

Technological developments in datAcron will be validated and evaluated in user-defined challenges that aim at increasing the safety, efficiency and economy of operations concerning moving entities in the aviation and maritime domains. The main benefit arising from improved trajectory prediction and events forecasting in the aviation use case lies in the accurate prediction of complex events or hotspots, leading to benefits to the overall efficiency of an air-traffic management (ATM) system. On the other hand, discovering and characterizing the activities of vessels at sea are key tasks to Maritime Situational Awareness (MSI) indicators and constitute the basis for predicting vessel activities, towards enhancing safety, detecting anomalous behaviors, enabling an effective and quick response to maritime threats and risks.

4. CONCLUDING REMARKS

datAcron developments are expected to provide advanced methods for the management of big mobility data, as well as analytics methods exploiting such data; many of these methods must comply with operational latency requirements (i.e. in ms) imposed by the target scenarios. Our presentation in this workshop aims to present datAcron current and targeted developments in the big data community and stakeholders from academia and industry.

[1] datAcron project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 687591; **Project Coordinator**: Prof. George A. Vouros, University of Piraeus, Greece (georgev@unipi.gr); **Project web site**: http://datacron-project.eu.