Integrating Heterogeneous Urban Data for Emergency Response: The INSIGHT project into Dublin's Traffic Control Room

(Extended Abstract)

Ioannis Katakis (katak@di.uoa.gr), Dimitrios Gunopulos (dg@di.uoa.gr),
National and Kapodistrian University of Athens

The Urban Data Landscape. Today's urban environment is awash with data. Data become available from a wide range of origins, from fixed and mobile sensors to large-scale monitoring infrastructures, and can come from public, private, or industry sources. Making such data useful is an enabler for the development of novel innovative Big Data applications that utilize massive urban data streams. There are large efforts underway on this front, as we are gradually moving towards a smart city era. Technologies that discover knowledge from urban data, such as new techniques for the detection of disastrous events, tracking health issues, monitoring crucial environmental factors, or improving energy efficiency, will have impact in a lot of aspects of the citizen's everyday life.

The importance of data originating in the urban space is going to expand significantly in the future. In fact, the smartamerica.org project reports that United States cities alone are going to invest an estimated sum of 41 trillion USD over the next 20 years with the stated goals to upgrade their infrastructure so that they can improve connectivity and the ability to collect data, in order to improve the efficient use of city resources and enhance the quality of life.

The toughest obstacles in using urban data are that such data are heterogeneous, noisy, and unlabeled. In addition urban data can include massive and high-speed data streams. The combination of monitoring and networking technologies produces a space where a very large numbers of data producers produce high volumes of data that are available in real time.

Integrating Urban Sensors for Emergency Response. The vision of INSIGHT (FP7 project - www.insight-ict.eu) is to develop the capabilities that the combined use of technologies can offer for fundamental advances in the Smart City paradigm, focusing especially on emergency situations. We recognize that a deep integration of data collections, the knowledge discovery and the information sharing capabilities offered by today's technology is the cornerstone for the advances we seek. The goal of INSIGHT is to achieve a significant improvement in the utility of automated systems to manage resources and put new capabilities in the hands of disaster planners and city personnel when responding to emergencies in smart cities and countries. Given the increasing amount and complexity of data, the inherent complexities are immense. To achieve quantum improvement, we have to integrate quantitative data from physical sensor with qualitative data from social networking media.

Technical Contribution. In the proposed presentation we will discuss the challenges we confronted and the solutions we had to provide in order to build a system that will be of great value for citizens and authorities. Research challenges had to be tackled in multiple levels:

- *Architecture:* Massive heterogeneous noisy data sources require special treatment. Data synchronization and storing as well as batch and real time processing were the main elements of the lambda architecture that our infrastructure was based on.
- *Data Integration:* Four sources of information are currently utilized: a) SCATS sensors embedded in the streets of Dublin indicating traffic volume, b) GPS locations of buses moving around the city, c) Geo-located Twitter messages, and d) Crowdsourcing information from the INSIGHT mobile application. Each data source required a unique storing, preprocessing and uncertainty management procedure. Each urban stream is associated with one or more dedicated Intelligent Sensor Agent (ISA). ISAs are INSIGHT's analytic building blocks.
- *Analytics:* A wide range of novel data mining and machine learning algorithms have been employed in order to provide tools that are able to continuously process the data streams and a) recognize incidents, b) predict the situation in the near future, c) identify the event type (traffic jam, flood, etc.)
- *Knowledge Fusion:* Each ISA has its own view of the city and that many times leads in contradictory information on what's happening. Discussions between the ISAs are initiated utilizing an Ontology-based 'common language' and a 'Round-Table' resolution strategy.
- *Usability / Interpretability:* Visual Analytics constitute crucial component since they provide the means that enable situation understanding.
- *Uncertainty Management:* Uncertainty is inherent in all elements of the INSIGHT workflow: from faulty sensors to contradictory ISA predictions. A Crowdsourcing component in the INSIGHT mobile application aims at resolving these issues.

A report from the cockpit. In this presentation we intend to share our experience and lessons learned from the three years of the project. Our research consortium had the chance to closely collaborate with Dublin City Council (Intelligent Transportations Department) towards automatically identifying important issues in the streets of Dublin. Until recently, the current situation in DCC was that human operators monitor a lot of information by looking at CCTV camera's and other data sources. This process is rather inefficient since a large part of this information is intractable by DCC employees. The INSIGHT consortium has already built a system that analyzes traffic flow information, GPS location data from buses moving around the city and real-time twitter data. In addition, a crowdsourcing application is utilized to retrieve information from citizens that are near the location of incidents. Currently, the system is able to identify traffic issues, flooding events, as well as faulty sensors by integrating information obtained from multiple ISAs.

Presentation Content. The presentation will discuss a) the research challenges and the solutions provided by the consortium, b) a report from the use of the prototype, c) live or video demonstration of the system, d) future challenges and opportunities on mining urban data.

Resources

☐ INSIGHT Project Web Site – www.insight-ict.eu
☐ Mining Urban Data Workshops – www.insight-ict.eu/mud & www.insight-ict.eu/mud2
☐ Dublinked – Open Data Repository from the city of Dublin - http://www.dublinked.ie/