T-Cube Web Interface in Support of Real-Time Bio-surveillance Program

Artur Dubrawski, Maheshkumar Sabhnani, Michael Knight, Michael Baysek, Daniel Neill, Saswati Ray, Anna Michalska and Nuwan Waidyanatha

Abstract—**T-Cube Web Interface is a generic tool to visualize and manipulate large scale multivariate time series datasets. The interface allows the user to execute complex queries quickly and to run various types of statistical tests on the loaded data. We show its utility in an important application scenario: real-time bio-surveillance system designed to support rapid detection and mitigation of bio-medical threats in developing countries.**

Index Terms—bio-surveillance, event detection, interactive analytics, data cubes.

Infectious diseases such as tuberculosis or avian influenza pose a persistent threat to human population all over the world. Their mitigation and containment is especially difficult in the least developed countries, where existing surveillance systems cannot gather and process the necessary information in time for public health authorities to effectively respond to rapidly developing bio-events. For example, the system currently in use in Sri Lanka is paper-based and relies on the postal service to communicate community health reports to the central office for analysis, which may take even ten days. Communication latency is exacerbated by limitations of analytic resources (in terms of their quantity as well as training) available to process the available information. We believe that that affordable, existing information and communication technologies can play the key role in collecting, communicating and analyzing the relevant data, enabling reliable and timely identification of emerging threats and, therefore, effective containment of their consequences.

The International Development Research Centre of Canada supports the Real-Time Bio-surveillance Program (RTBP), a pilot project currently under development in Sri Lanka and Tamil Nadu state of India. It will consist of information gathering system based on mobile handheld devices and wireless networking, and of a specialized IT system designed for automated surveillance of the incoming data and for interactive navigation through data and extracted information. In this demonstration, we present a prototype application of the T-Cube Web Interface in monitoring of public health data in Sri Lanka and India.

T-Cube Web Interface [1] is a generic tool for interactive analysis, visualization and manipulation of large scale multidimensional temporal and spatio-temporal datasets commonly encountered in public health domain. It allows the users to execute complex queries quickly and to perform various types of statistical tests against the loaded data. The underlying data representation technology, the T-Cube [2], is an in-memory data structure designed to improve the response time to ad-hoc time series queries against large datasets. T-Cube has been tested on synthetic and real-world datasets containing millions of records and hundreds of dimensions. Results show that its response time can be orders of magnitude shorter than that of the state-of-the-art commercial database tools. The attainable speedups have been shown to qualitatively change the way the users interact with their data, leading to better situational awareness, comprehension, and interpretability of data and results of analyses. It has already been successfully used in equipment maintenance, logistics and food safety domains. We intend to leverage the benefits of T-Cube technology in RTPB.

Demonstrated capabilities of the tool include geospatio-temporal visualization of syndromic data, navigation through different levels of data aggregation, and selected statistical analyses including spatial scan for rapid detection of emerging outbreaks of diseases. Presentation emphasizes user-perceived utilities due to computational efficiency of the analytic algorithms and the underlying data representation technique.

A. Dubrawski, M. Sabhnani, M. Knight, M. Baysek, D. Neill, and S. Ray are with the Auton Lab, Carnegie Mellon University, Pittsburgh, PA 15213 USA (Corresponding author: A. Dubrawski, phone: 412-268-6233; fax: 412-268-6233; e-mail: awd@cs.cmu.edu) (Other authors email: sabhnani@cs.cmu.edu, mikek@cs.cmu.edu, mjbaysek@cs.cmu.edu, neill@cs.cmu.edu, sray@cs.cmu.edu).

A. Michalska is with the Institute of Computer Science, University of Warsaw, ul. Banacha 2, 02-097 Warszawa, Poland (e-mail: michalsk@mimuw.edu.pl).

N. Waidyanatha is with LIRNEasia, 12 Balcombe Place, Colombo 08, Sri Lanka (e-mail: waidyanatha@lirne.net).

REFERENCES

[1] S. Ray, A. Michalska, M. Sabhnani, A. Dubrawski, M. Baysek, L. Chen, J. Ostlund, T-Cube Web Interface: A Tool for Immediate Visualization, Interactive Manipulation and Analysis of Large Sets of Multivariate Time Series, *2008 AMIA Annual Symposium*, Washington, DC, 2008.

[2] M. Sabhnani , A. Moore, and A. Dubrawski, Rapid processing of ad-hoc queries against large sets of time series. *Advances in Disease Surveillance* 2, 2007.