

Chapter 4.3

Anomaly Detection in Streaming Sensor Data

Alec Pawling

University of Notre Dame, USA

Ping Yan

University of Notre Dame, USA

Julián Candia

Northeastern University, USA

Tim Schoenharl

University of Notre Dame, USA

Greg Madey

University of Notre Dame, USA

ABSTRACT

This chapter considers a cell phone network as a set of automatically deployed sensors that records movement and interaction patterns of the population. The authors discuss methods for detecting anomalies in the streaming data produced by the cell phone network. The authors motivate this discussion by describing the Wireless Phone Based Emergency Response (WIPER) system, a proof-of-concept decision support system for emergency response managers. This chapter also discusses some of the scientific work enabled by this type of sensor data and the related privacy issues. The authors describe scientific studies that use the cell phone data set and steps we have taken to ensure the security of the data. The authors also describe the overall decision support system and discuss three methods of anomaly detection that they have applied to the data.

INTRODUCTION

The Wireless Phone-Based Emergency Response System (WIPER) is a laboratory proof-of-concept,

DOI: 10.4018/978-1-61350-101-6.ch403

Dynamic Data Driven Application System (DD-DAS) prototype that uses cell phone network data to identify potential emergency situations and monitor aggregated population movement and calling activity. The system is designed to complement existing emergency response man-

agement tools by providing a high level view of human activity during a crisis situation using real-time data from the cell phone network in conjunction with geographical information systems (GIS). Using cell phones as sensors has the advantages of automatic deployment and sensor maintenance; however, the data available from the network is limited. Currently only service usage data and coarse location data, approximated by a Voronoi lattice defined by the cell towers, are available, although cell-tower triangulation and GPS could greatly improve the location data (Madey, Szabó, & Barabási, 2006, Madey et al., 2007, Pawling, Schoenharl, Yan, & Madey, 2008, Schoenharl, Bravo, & Madey, 2006, Schoenharl, Madey, Szabó, & Barabási, 2006, Schoenharl & Madey, 2008).

The viability of using cell phones as a sensor network has been established through the use of phone location data for traffic management (Associated Press, 2005). WIPER applies this finding to fill a need in emergency response management for a high level view of an emergency situation that is updated in near real-time. Tatomir and Rothkrantz (2005) and Thomas, Andoh-Baidoo, and George (2005) describe systems for gathering on-site information about emergency situations directly from response worker on the ground via ad-hoc networks of PDAs. While these systems can provide detailed information about some aspects of the situation, such as the location of victims and environmental conditions, the information is limited to what can be observed and reported by the responders. This provides valuable but local information, though there may be observations from different, geographically dispersed locations. In contrast, WIPER provides less detail, but instead gives an overall view of population movements that may be valuable in refining the response plans or directing response workers to gather more detailed information at a particular location.

Dynamic data driven applications systems (DDDAS) provide a framework in which running

simulations incorporate data from a sensor network to improve accuracy. To achieve this, the simulations dynamically steer the measurement process to obtain the most useful data. The development of DDDAS applications is motivated by the limited ability to predict phenomena such as weather and wildfire via simulation. Such phenomena are quite complex, and the correct simulation parameterization is extremely difficult. The goal of DDDAS is to provide robustness to such simulations by allowing them to combine sub-optimal initial parameterizations with newly available, real-world data to improve performance without the expense of rerunning the simulations from the beginning (Douglas & Deshmukh, 2000).

In this chapter, we focus on one component of WIPER: the detection and alert system. This module monitors streaming data from the cell phone network for anomalous activity. Detected anomalies are used to initiate an ensemble of predictive simulations with the goal of aiding emergency response managers in taking effective steps to mitigate crisis events. We discuss methods for anomaly detection on two aspects of the call data: the call activity (the number of calls made in a fixed time interval) and the spatial distribution of network usage.

The remainder of the chapter is organized as follows: we discuss background literature related to mining data from a cell phone network. We start with a discussion of methods for detecting outliers in our data, with a focus on using data clustering to model normality in data. Those clusters of outliers in the streaming data could be indicators of a problem, disaster or emergency in a geographical area (e.g., an industrial explosion, a civil disturbance, progress of a mandated evacuation prior to a hurricane, a terrorist bombing). We then give an overview of the data set and the WIPER system, followed by descriptions of algorithms used in the detection and alert system. Finally, we discuss some of the privacy issues related to this work and our plans for future work

17 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/anomaly-detection-streaming-sensor-data/58823

Related Content

Traffic-Based S-MAC: A Novel Scheduling Mechanism for Optimized Throughput in Mobile Peer-to-Peer Systems

Odysseas Shiakallis, Constandinos X. Mavromoustakis, George Mastorakis, Athina Bourdena and Evangelos Pallis (2015). *International Journal of Wireless Networks and Broadband Technologies* (pp. 62-80).

www.irma-international.org/article/traffic-based-s-mac/125819

Traffic-Based S-MAC: A Novel Scheduling Mechanism for Optimized Throughput in Mobile Peer-to-Peer Systems

Odysseas Shiakallis, Constandinos X. Mavromoustakis, George Mastorakis, Athina Bourdena and Evangelos Pallis (2015). *International Journal of Wireless Networks and Broadband Technologies* (pp. 62-80).

www.irma-international.org/article/traffic-based-s-mac/125819

A Mobility-Based Routing Protocol for CR Enabled Mobile Ad Hoc Networks

Yan Sun, Jingwen Bai, Hao Zhang, Roujia Sun and Chris Phillips (2015). *International Journal of Wireless Networks and Broadband Technologies* (pp. 81-104).

www.irma-international.org/article/a-mobility-based-routing-protocol-for-cr-enabled-mobile-ad-hoc-networks/125820

Communication Issues in Pervasive Healthcare Systems and Applications

Demosthenes Vouyioukas and Ilias Maglogiannis (2012). *Wireless Technologies: Concepts, Methodologies, Tools and Applications* (pp. 984-1014).

www.irma-international.org/chapter/communication-issues-pervasive-healthcare-systems/58827

Localizing Persons Using Body Area Sensor Network

Cheng Guo, R. Venkatesha Prasad, Jing Wang, Vijay Sathyanarayana Rao and Ignas Niemegeers (2012). *Developments in Wireless Network Prototyping, Design, and Deployment: Future Generations* (pp. 273-289).

www.irma-international.org/chapter/localizing-persons-using-body-area/67014