Traffic Congestion Control based In-Memory Analytics: Challenges and Advantages

Aiman Abdul-Razzak Fatehi Al-Sabaawi

School of Electrical Engineering and Computer Science
Queensland University of Technology
Brisbane, Queensland, Australia

ABSTRACT
A method for loading detailed data from more than one source into the main memory in traffic system is presented. The challenges include data volume, data velocity, and Data variation of the traffic. The method uses In-memory analytics to improve query evaluation, to analysis and process reports in the traffic system. This occurs with the development of multicore processors, the loading of data, and the way image and video traffic data are stored and transferred into/from the data centers of different divisions and centralizing access to traffic management facilities, equipment and application systems.

General Terms
Big Data Traffic Congestion, In-memory analytics.

Keywords
Traffic Congestion, In-memory analytics, Real-time, Data volume, Data variety, Data velocity, Big data.

1. INTRODUCTION
Traffic information systems have many complex requirements to be addressed if the database is to be useful on a large scale. Information accuracy and availability in real-time, as well as the variety and volume of data are important elements of traffic information that must be easily interpreted by the user as in [1]. Today, smartphones are being used that carry multiple sensors can be used to retrieve the location, direction and trajectories of the user. There are many sampling techniques that have the ability to collect data from GPS-supporting smartphones. In the Nokia N95 device, the embedded GPS chip-set can produce time-stamped Geolocation from latitude, longitude, altitude every three second. At the same frequency, the phone calculates the immediate velocity from position data and the current time. This vehicle path and velocity information generates the velocity field and a rich history of the dynamics of the vehicle which develops over time [2]. GPS devices are able to communicate with a server using the General Packet Radio Service (GPRS) protocol. Geographical and timestamp information is sent to a server using the remote connection. This technology provides highly accurate traces of the routes taken by cyclists both on and off the street network, providing insight into how and where individuals tend to ride. Collected trip data from bicycle commuters using handheld GPS devices can be used to generate multivariate regression models to identify physical attributes that affect route choice decisions. As a result, GPS sends massive data to the network to make the correct decision on the router as in Figure 1.

This paper is organized as follows: Section 2 discusses memory analytics in information management, including traffic data challenges and advantages of using in –memory analytics. Section 3 potential market. In Section 4, issues in exploitation or implementation. Finally, Section 5 presents the conclusion of the work.

2. USING IN MEMORY ANALYTICS IN INFORMATION MANAGEMENT
In-memory analytics involves quickly analyzing massive volumes of data, such as those from retail and telecommunications companies, traffic networks and GPS trails, the content of web pages, satellite imagery and financial market data. This technique is especially effective for real-time analysis and complex analytics. In [4], the author demonstrated that patterns in the data can be identified quickly using In-memory analytics. In-memory analytics provide get insights as well as determining areas warranting further study in their own right. Traffic monitoring systems perform spatial and time dependent analyses of different types of data, for example, data from GPS, video cameras, microwave radar as well as traditional inductive loop detector data. The aim of such systems is the provision information such as times, average speeds, densities or volumes on a given segment of a roadway, such as from GPS data that are gathered from a smartphone. With the popularity of the application of large-scale traffic sensors and surveillance cameras in traffic network analysis and allocated resources, traffic monitoring systems have collected a large amount of structured/unstructured traffic data as in [4]. Recently, many researches have been investigated related to this subject. Collection and analysis of road traffic data is an important requirement towards establishment of traffic conditions on any given road segments. Global Positioning System (GPS) data receiver application and traffic analysis system was developed in [5] that collected GPS traffic data and provided the ability for monitoring and analyzing traffic scenarios on the roads, for example, the speed of traffic and providing planners on the road usage patterns for decision making. All these aspects can be analysed both in real-time and historically based on the fact that historical data was captured and stored for future use. The system has added the ability to trigger email alerts on speeding vehicles. The results showed that there is great need for real-time traffic information analysis because of the tremendous variability in traffic data in big cities, such as position, speed and directions of vehicles travelling on the main roads with the speed of traffic algorithm developed and effectively put in place.
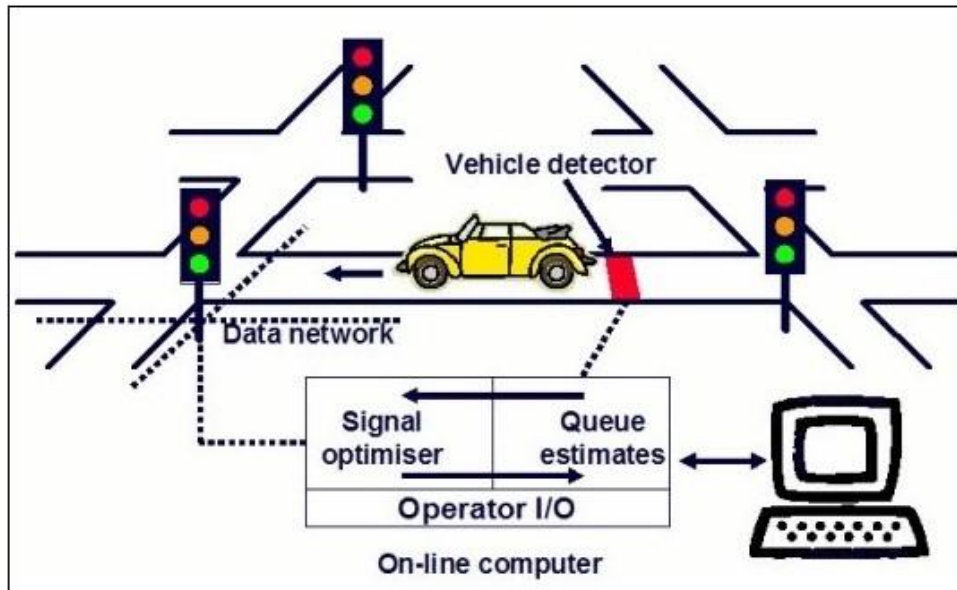
Fig 1: Traffic data network [3]

In [6], they discussed the simulation and evaluation of a traffic congestion detection system which combines inter-vehicular communications, fixed roadside infrastructure and infrastructure-to-infrastructure connectivity and big data. The system gave permission to drivers to identify traffic jam and change their routes to save time. This system monitors, processes and stores large amounts of data, which can detect traffic congestion in a precise way based on Cassandra, which was used in tandem with the OMNeT++ discreet event network simulator, coupled with the SUMO (Simulation of Urban MObility) traffic simulator and the Veins vehicular network framework.

2.1 The traffic Data Management Challenges

Traffic and transport are complex activities and connect locations with flow data. These locations may be predicted, well-connected and display high levels of access – as in many big cities. The flow data between these locations may concern people or goods and can impact on any number of vehicle types as in [7]:

- Data volume: Data involving decisions are generated in a variety of operational systems, for example, bridges, tracks, roads, bus stations, airports, cycle paths and ports. All of these decisions require massive information (terabyte, petabyte) to improve the safety of transport.

- Data velocity: This requires that events and changes need to be dealt with speedily and efficiently. Processing decisions based on incomplete facts can have a lasting negative impact. Consequently, coping with data velocity is vital for making timely decisions and controlling the value chain.

- Data variety: Data can be collected in many diverse forms: Global Positioning System (GPS) as well as process logs, time stamp and Geo-localised spatial data stamps. Metadata regarding device identity, location and status are used by smartphones to stay connected to different networks (Wi-Fi, GSM, etc.). Data that are produced by devices, vehicles and networked objects, public transport card tap-ins or swipes and other data associated with portal access (badges, RFID tags, key cards) or cordon passage (e.g. congestion charging systems, toll roads, etc.). A commercial transaction data (credit card uses and transaction records, bar-code and RFID tag readings) and SMS and Email, metadata relating to phone calls. It is imperative to aggregate and reconcile heterogeneous data to derive maximum value easily as in [8].

These issues can be solved by using In-memory analytics to improve query evaluation and to analyze and process reports in the traffic system. Because In-Memory Analytics has the ability to query the data from Random Access Memory instead of utilizing a physical disk, users can directly load detailed data from more than one source into the main memory. In-Memory Analytics techniques help to support faster business decisions with operations being performed in the memory and performance is improved as storage. The in-Memory Analytics approach has resulted in a fundamental change in storage philosophy. In brief, in computers, data are stored in RAM, but data which are stored in a database as tables through relationships and linked operations among tables, and other database objects. In-memory analytics can provide faster queries and calculations which result in fast response times as in [9].

2.2 The Advantages of Using In-memory Analytics to Process the Traffic Stream Data in Real-Time

In-Memory Analytics are utilized to develop a query evaluation and processing of reports in the traffic system. However, the demands of users and applications on computing resources are clearly understood using data proofing. In-memory analytics is also significant when identifying applications and users who need to process ad hoc and non-routine reports. Data usage models accomplish this and reduce the effort and cost of introducing in-memory analytics in a firm. Currently, there is a dramatic decrease in the cost of processors and memory that is occurring simultaneously with the development of multicore processors. Using In-Memory computing, it becomes possible for storage

and operations to be performed in main memory, avoiding the requirement for hard disks [9].

In-Memory computing becomes useful for two reasons. Firstly, the volume of information is increasing at an alarming rate. Secondly, responses are immediately needed and quick decisions are required in all forward-looking systems. Traditionally, annual and quarterly review reports were used as the basis for decision making. In-Memory computing supports event driven systems that have the capacity to make decisions in real time. Here data are brought closer to the Central Processing Unit (CPU). The querying of the data based on in-memory analytics is a million times faster than disk based access. In-memory analytics and GIS technology provide tools for transportation professionals that are easier to use than ever before and for less than they spend now, because managing traffic data have become easier, more detailed, or more affordable as in [9].

3. POTENTIAL MARKET

In-Memory analytics has given rise to a fundamental change in approach to data management and storage to improve immediate reporting for decision making. A revolution of in-memory analytics techniques, significantly decreases the cost of memory, and direction business intelligence queries have been changed by the development of multicore processors, the loading of data, and the way data are stored and transferred. Adoption of in-memory technology is enabling some companies to use it in traffic and transportation systems [9].

Therefore, managing local traffic provides professional transportation services to the public thorough congestion detection, enhanced traffic performance, and real-time traffic surveillance, where by congestion is detected using optical flow in-memory analytics to process videos and to compute traffic speed as in [9].

4. ISSUES IN EXPLOITATION OR IMPLEMENTATION

There are three issues associated with the implementation of in-memory analytics:

- The cost of hardware and expertise [10]. To reduce cost, there are two alternatives: one is to conduct a pilot study and the other is to purchase to develop the pilot study using the platform's service as in.

- New people are required because of the need for data specialists and technical people who can operate in-memory systems.

- Access to GPS data feeds are required to avoid the investment in infrastructure to collect it. The need for infrastructure is the issue and the solution is to pay for the data as in [11].

Moreover, the average monthly data have now reached 10 terabytes as traffic data continue to grow. Data, pictures and video are stored in various data centers in various divisions as in [12]. Therefore, there are many problems with the volume of big data from traffic stream data as in [13]:

- Real time and cost are a major issue: The important issues for big data traffic are the cost and time required to implement the changes. The system analyses real-time data from its transport networks (for example, road traffic volumes, bus locations, and passenger movements) to monitor traffic light timing and whether buses are running as close as possible to their timetable.

- Cloud Issue: the volume of data may be terabytes of data, and loading it to the cloud could take a great deal of time. Furthermore, the dataset could be changing rapidly.

When using in-memory analytics in traffic systems, traffic problems will occur in gathering traffic data. These include road conditions, passenger flows, transport parameters, locations for temporary detention of vehicles, road accident statistics, parameters of environmental damage from traffic, data for parking lots and effective organization of traffic that needs the integration of numerous sensors which exhibit noise. Noisy data need to be detected early to prevent the unnecessary semantic enrichment that may lead to more significant problems at reasoning time. In the future, a sophisticated filter at the stream processing level is needed to address this problem and to check data validity as in [9].

5. CONCLUSION

This paper has shown the traffic information systems have many issues related to the volume, velocity and variety of data that require solutions. Because of the difficulties faced in attempting to use all these data, there is a need to centralize access to image and video traffic data stored at the data centers of different divisions as well as centralizing access to traffic management facilities, equipment and application systems. As in [14], GPS data are available and accessible and provide information about roads and destinations. At a minimum, this can show where traffic is heading, allowing predictions to be made about where traffic problems will occur, and to create a simulation which can be used for transport system planning and optimization. For future work, In-memory analytics based on GPS data can be provided better information for traffic management in traffic planning.It is recommended that a prototype be developed first. Using these data, big or small prototypes can be built. It is recommended that a pilot study be conducted to examine one segment of the transport system to be able to capture the data correctly. The productive algorithm format can be visualized and developed.

6. ACKNOWLEDGMENT

7. REFERENCES

[1] Herring, R. J. 2010. Real-time traffic modeling and estimation with streaming probe data using machine learning. Retrieved from http://search.proquest.com.ezp 01.library.qut.edu.au/docview/859247684?pq-origsite=summon.

[2] Herrera, J. C., Work, D. B., Herring, R., Ban, X., Jacobson, Q., and Baye, A. M. 2010. Evaluation of traffic data obtained via GPS-enabled mobile phones: The mobile century field experiment. Transportation Research Part C, vol.18, issue 4, (2010), 568-583. doi:10.1016/j.trc.2009.10.006.

[3] How Scoot Works. Retrieved from http://www.scoot-utc.com/HowSCOOTWorks.php

[4] Garber, L. 2012. Using in-memory analytics to quickly crunch big data. Computer, vol. 45, issue 10, (2012),16-18. doi:10.1109/MC.2012.358.

[5] Obuhuma, J. I. and Moturi, C.A., 2012. Use of GPS with road mapping for traffic analysis. International Journal of Scientific and technology research.

[6] Cárdenas-Benítez, N., Aquino-Santos, R., Magaña-Espinoza, P., Aguilar-Velazco, J., Edwards-Block, A. and Medina Cass, A., 2016. Traffic congestion detection system through connected vehicles and big data. Sensors, 16(5), p.599.

[7] OECD, 2015. Big Data and Transport: Understanding and Assessing Options. Retrieved from http://www.internationaltransportforum.org.

[8] IBM, 2016. Traffic Management for a Smarter Planet. Retrieved from http://www.ibm.com/smarterplanet/us/en/traffic_congestion/article/traffic-management-and-prediction.html.

[9] Hota, J. 2013. Adoption of In-Memory Analytics. CSI Communications, 20. Chicago. Retrieved from http://papers.ssrn.com/sol3/papers.cfm?abstract_id=2250411.

[10] Woo, B. 2014. Oracle's 'In-Memory' option may be costly. Retrieved from http://www.neuralytix.com/blog/2014/06/11/oracles-in-memory-option-may-be-costly/

[11] Sheldon, R. 2015. Oracle in-memory database option: Intro and implementation. Retrieved from http://searchoracle.techtarget.com/tip/Oracle-in-memory-database-option-Intro-and-implementation.

[12] Haitao, C. 2013. Improving traffic management with big data analytics. Retrieved from http://www.intel.com.au/content/dam/www/public/us/en/documents/case-studies/big-data-xeon-e5-trustway-case-study.pdf.

[13] Karena, C. 2012. Top 6 big data issues. Retrieved from http://www.smh.com.au/it-pro/business-it/top-6-big-data-issues-20120816-24a6w.html#ixzz3km2jR0OC.

[14] Stelle, G., Olivier, S. L., Stark, D., Rodrigues, A. F. and Hemmert, K. S. 2014. Using a complementary emulation-simulation co-design approach to assess application readiness for processing-in-memory systems. (2014), 64-71. doi:10.1109/Co-HPC.2014.5.