Journal of Geotechnical and Transportation Engineering
Volume 1 | Issue 2
Big Data-Driven Smart Policing: Big Data-Based Patrol Car Dispatching
Ben Ali

Received 6/10/2015 Accepted 11/01/2015
Published 12/15/2015
Corresponding Author:
W.Ben Ali; oualid.ali@gutech.edu.om

Big Data-Driven Smart Policing: Big Data-Based Patrol Car Dispatching

Walid Ben Ali
German University of Technology, Muscat, Oman
Oualid.ali@gutech.edu.om

Abstract

Big Data has become one of the buzzwords today. The recent explosion of digital data has led the organization, either private or public, to a new era towards a more efficient decision making. At some point, business decided to use that concept in order to learn what make their clients tick with phrases like 'sales funnel' analysis, 'actionable insights', and 'positive business impact'. So, it stands to reason that Big Data was viewed through green (read: money) colored lenses. Somewhere along the line, however someone realized that collecting and processing data doesn't have to be for business purpose only, but also could be used for other purposes to assist law enforcement or to improve policing or in road safety. This paper presents briefly, how Big Data have been used in the fields of policing order to improve the decision making process in the daily operation of the police. As example, we present a Big Data driven system which is sued to accurately dispatch the patrol cars in a geographic environment. The system is also used to allocate, in real-time, the nearest patrol car to the location of an incident. This system has been implemented and applied in the Emirate of Abu Dhabi in the UAE.

Keywords: Big Data, patrol car allocation, dispatching, GIS

1. Introduction

Law enforcement methodology has evolved tremendously in the past few decades to embrace quantitative analysis and statistics with the rising of data-driven policing. In the current age of Big Data, policing has taken great advantage of this new concept in order to improve its operations. Real-time access to an analysis of vast quantities of information found in criminal records, police databases, and surveillance data may alter policing [1] in the same way that Big Data has revolutionized areas as diverse as presidential elections [2], internet commerce [3], etc. Some have even heralded Big Data's potential to change our assumptions about social relationships, government, scientific study, and even knowledge itself [4]. Most of the literature research work concerning Big Data discussed the benefits of such concept in the business fields. In this paper, we will present how the government and especially, the policing field can take benefit from Big Data. So, we will present from the literature three major uses of Big Data that hint at the future of policing. Then, we will present our research which represents a fourth use of Big Data in spatial decision making process. In this research, we present a Big Data driven practical work which helps the police department dispatching resources such as patrol cars in geographic environment. The structure of the paper is as follow: In the first part, we will present the Big Data and its analytics concepts. Then, in the second part, we will talk about the common uses of Big Data in policing. In the section II, we will present the prototype of a geo-intelligent patrol cars allocation based on Big Data. This prototype has been developed with the support of the Police department of Abu Dhabi Emirate in the UAE. Finally, we conclude the paper.

2. The rise of Big Data

The discussions about Big Data in literature are diverse and so are the definitions of Big Data and how the term is used. In this section, we present a brief overview of Big Data as well as Big Data Analytics.

2.1. Big Data

In one of the largest commercial studies titled "*Big Data: The next frontier for innovation, competition, and productivity*" the McKinsey Global Institute (MGI) used the following definition: "*Big Data refers to datasets whose size is beyond the ability of typical database software tools to capture, store, manage, and analyze. This definition is intentionally subjective and incorporates a moving definition of how big a dataset needs to be in order to be considered Big Data."* [5]
With that definition MGI emphasizes that there is no concrete volume threshold for data to be considered 'big', but it depends on the context. However the definition uses size or volume of data as only criterion. This usage of the term 'Big Data' can be misleading as it suggests that the notion is mainly about the volume of data. In the literature, other definitions of Big Data have been presented, but most of them focus on the size of data in storage. Size matters, but there are other important attributes of Big Data, namely data *variety* and data *velocity*. The three Vs of Big Data (volume, variety, and velocity) constitute a comprehensive definition, and they bust the myth that Big Data is only about data volume [6]. Therefore, most recent publications extend the MGI's definition. One of the

definitions is given in IDC's 'The Digital Universe' study: "*IDC defines Big Data Technologies as a new generation of technologies and architectures, designed to economically extract value from very large volumes of a wide variety of data enabling high-velocity capture, discovery, and/or analysis. There are three main characteristics of Big Data: the data itself, the analysis of the data, and the presentation of the results of the analytics*" [7]
This definition is based on the 3V's model coined by Laney, 2001 [8]. He did not use the term 'Big Data', but he predicted that one trend in e-commerce is, the data management will be more and more important and difficult. He then identified the 3V's- data volume, data velocity and data variety- as the biggest challenges for data management. In the following I will describe the mentioned characteristics or dimensions in more detail:

- *Data volume*: As discussed before handling volume is the obvious and most widely recognized challenge in Big Data world. There is however no clear or concrete quantification of when volume should be considered 'big'. This is rather a moving target increasing with available computing power. While 300 terabyte were considered big 10 years before, today petabyte are considered big and the target is moving more and more to Exabyte and even zettabyte [7].
- *Data velocity*: Velocity refers to the speed of data. This can be twofold. First, it describes the rate of new data flowing in and existing data getting updated [9]. Agrawal et al. 2012, [10], call this the 'acquisition rate challenge'. Second, it corresponds to the time acceptable to analyze the data and act on it while it is flowing, called 'timeliness challenges'. These are essentially two different issues, that do not necessarily need to occur at the same time, but often they do.
- *Data variety*: One driver of Big Data is the potential to use more diverse data sources, data sources that were hard to leverage before and to combine and integrate data sources as a basis for analytics. There is a rapid increase of public available, text-focused sources due to the rise of social media several years ago. This accompanies blog spots, community pages and messages and images from social networks, but there is also a rather new (at least in its dimension) source of data from sensors, mobile phones and GPS [11]. Companies e.g. want to combine sentiment analysis from social media sources with their customer master data and transactional sales data to optimize marketing efforts. Variety herby refers to a general diversity of data sources. This not only implies an increased amount of different data sources but obviously also structural differences between those sources.

2.2. Big Data Analytics

In order to manage the giant volume of data, it has been emerged the 'Big Data' concept. Big Data was seen as a mean to manage to reduce the costs of data management. Now, the companies focus on the value creation potential. In order to benefit from additional insight gained there is the need to access the analytical and execution capabilities of Big Data. To turn Big Data into business advantage and make every decision as desired, businesses have to review the way they analyze it in order to bring the results of knowledge discovery to the business process [13]. This is how Big Data Analytics surged up [12]. So, someone can ask the question "What is Big Data Analytics"? Big Data Analytics is where advanced analytic techniques operate on Big Data [6]. It refers to the skills, technologies, applications and practices for a continuous and iterative exploration and investigation of the Big Data and extracts meaningful information and put these insights into actions for better decision making and transforming the business [12].

3. Use of Big Data in policing: A literature review

Recently, several police departments, especially in the USA, have taken notice that they stand to benefit from Big Data and its analytics concepts. In this section, we present three common uses of Big Data in policing.

3.1. Crime prediction: Predictive policing

Perhaps the most visible use of Big Data by police departments thus far has been predictive policing, which is the *first type* of predictive technology used by the police. Predictive policing is the application of computer modeling to historical crime data to predict future criminal activity [13]. While the police have long tried to find patterns of criminal activity on which to focus their resources [14], predictive policing permits the police to harness thousands of data points to forecast where crime is likely to happen [16]. The most basic models rely on past crimes, but data source can include factors as variable as payday schedules, seasonal variation, liquor store locations, and potential escape routes. What is new about predictive policing is not the use of quantitative data but the application of Big Data and Big Data analytics in order to perform predictive policing [13].
A *second type* of predictive technology focuses on the application of algorithms to Big Data coming out of social media in order to identify likely criminality based on the role of individual plays within a social network. This social network analysis begins with the assumption that social networks under-grid many crimes: an illegal drug-dealing network may loosely follow the hierarchical structure of legitimate business, with suppliers, distributors, buyers, and financiers. The algorithms used in social network software can help police visualizing the density of connections might take the form of exchanges, communications, family ties, participation in crime, or affiliations with an organization.

3. 2. Mass surveillance

If predictive policing harnesses data to predict the future, computer surveillance systems help police create a software-enhanced picture of the present, using thousands of data points from multiple sources within a city 15]. As with predictive policing, computer enhanced mass surveillance grows out of other policing techniques. While surveillance has long been an essential tool for the police, what has changed is its supporting technology. Sophisticated yet inexpensive, the surveillance equipment used by the police today produce enormous amount of data and information, often too much for the police to use in an efficient way without the help of technology. For example the NYPD has database of 16 million license plates captured from its license plate readers, along with the locations of where the plates were photographed [15]. The NYPD has responded to this Big Data problem by creating a software program with Microsoft called "Domain Awareness System or DAS" which collects and analyze information around clock within New York City from sources as disparate as the City's 3,000 public surveillance cameras, over 200 automatic license plate readers, more than 2,000 radiation sensors, and information from police databases. The software's mapping features permit the police to see and understand the information in a way that was not possible

before. Located within the NYPD's lower Manhattan Security Initiative Command Center, the DAS's operators can quickly use the software to identify potential threats. For example, the system can give the police real-time access to information that can reveal connections between persons, items, and places in ways that may not be obvious to individual crime analyst. The DAS employs video analytic software designed to detect threats, such as unattended bags. The NYPD claims that DAS can track where a car associated with a suspect is located, and where it has been in the past days, weeks, or months. The DAS can also check license plate numbers, compare them to watch lists, and provide the police with immediate access to any criminal history associated with the car owner.

3.3. Genetic Big Data: DNA Databanks

Perhaps less obvious but no less important a big data matter is the collection of DNA for criminal justice databases. Nowadays, such database becomes huge. For example, which as of June 2013, the US's database contained DNA profiles for 10.7 million offenders and 1.7 million arrestees. The USA has used this information to amass the largest DNA database in the world. Police agencies around the USA rely on CODIS- the shorthand for the system that links information among the different DNA databases around the country- to match crime scene samples with offender or arrestee DNA profiles. The millions of DNA samples now accessible by the police present another potential use of Big Data.

4. Big Data and spatial decision making: Big Data-driven Patrol Car dispatching in Abu Dhabi City, UAE

In the previous section, we have presented how Big Data can help police to fight crime. In our research, we will present how Big Data can help the police in decision making process. As example, we will show how Big Data helps the police of Abu Dhabi to dispatch efficiently and intelligently the patrol cars in the area of Abu Dhabi, in real time. In order to perform this intelligent real time dispatching, the system uses the following input data:

- The geo-spatial data of Abu Dhabi area;
- The real-time status of traffic in the roads of the regions, including traffic flow, traffic density, etc. This data is huge since it is coming from real-time traffic monitoring systems and devices already deployed in the environment;
- The number of incidents (expected and unexpected) representing accidents, emergency calls and calls for assistance. This data is saved in a huge database, since the system uses the historical records of the accident reports in Abu Dhabi City;
- The patrol vehicles data such as the features, the availability status, etc.;
- The start time and duration of each patrol duty cycle;
- The distance between patrol cars; this data is tracked using special tracking devices installed in each patrol car;
- Typical allocation of patrol officer time (administrative work, reporting work and available time to respond to calls); and
- Other data, such as the population density and distribution across the geographic zone.

As we can see, the input data of the system can be classified as Big Data, since it is coming from several sources, in different formats, and in real-time. The system will use Big Data analytics using the tool spatial-Hadoop [15] with few operational research and optimization algorithm in order to come up with the efficient dispatching result. The result presents the number of needed patrol cars as well as their optimized locations in the city. In Fig. 1, we present the result of the dispatching system in a portion of Abu Dhabi area.

The system is not only used for initial dispatching of the patrol cars, but it is also used to allocated the closest patrol car I an incident occurs. The accurate allocation use Big Data analytics using Spatial-Hadoop software [15]. The allocation can use the distance variable of the traffic time variable. If Fig. 2, we present the allocated result using distance only. But if we consider the traffic time, the allocated car will be different as shown in Fig. 3.

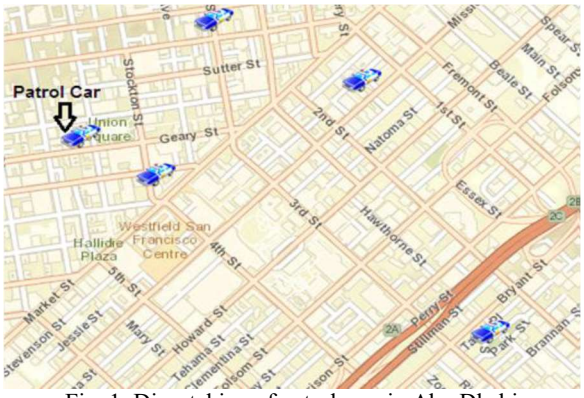

Fig. 1. Dispatching of patrol cars in Abu Dhabi

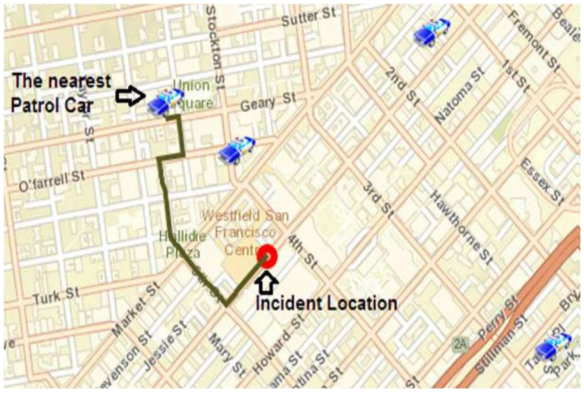

Fig. 2. Patrol car allocation for an incident (by distance)

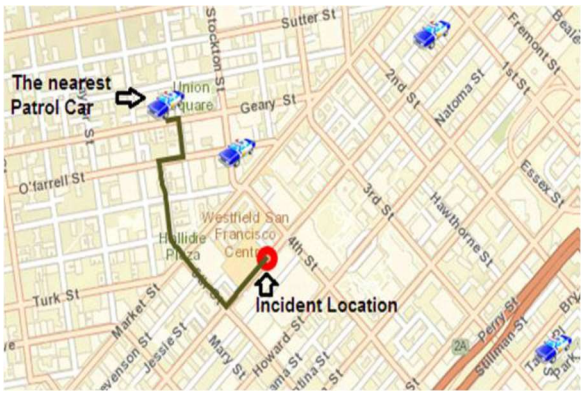

Fig. 3. Patrol car allocation for an incident (by time)

5. Conclusion

Literature has shown that Big Data and Big Data Analytics have been very effective in improving the process of decision making in

several fields and domains such as business, government, etc. In our paper, we have given a brief overview of such buzzwords, Big Data and Big Data analytics. Then, we have presented briefly, how these concepts have been used in the fields of policing in order to improve the decision making process. Afterwards, we have presented a study case which has proved how these concepts can be used to in order to dispatch and allocated patrol cars in a geographic area. By accurate dispatch of patrol cars, we can save time, money and lives. This system has been implemented and applied in the Emirate of Abu Dhabi in the UAE.

References

[1] Beth Pearsall, "Predictive policing: The future of law enforcement." NAT'L Inst. Just. J. 2010.
[2] Michael Scherer, "Inside the secret World of the Data Crunchers Who helped Obama Win." Time, 2012.
[3] Viktor Mayer-Schonberger and Kenneth Cukier, "Big Data: A revolution that will transform how we live, work, and think", 2013.
[4] Adam Frank, Big Data is the steam engine of our time. NPR, Mar 12, 2013.
[5] James Manyika, Michael Chui, Brad Brown, Jacques Bughin, Richard Dobbs, Charles Roxburgh, and Angela Hung Byers. "Big Data: The next frontier for innovation, competition, and productivity". Analyst report, McKinsey Global Institute, May 2011.
[6] Philip Russon. "Big Data Analytics". TDWI Research, Fourth Quarter, 2011.
[7] John Gantz and David Reinsel. "The Digital Universe in 2020: Big Data, Bigger Digital Shadows, and Biggest growth in the far East". Study Report, IDC, December 2012
[8] Doug Laney. "3D Data Management: Controlling data volume, velocity and variety". Technical report, META Group Inc. (Now Gartner Inc.), February 2001.
[9] Jinchuan Chen, Yueguo Chen, Xiaoyong Du, Cuiping Li, Jiaheng Lu, Suyuan Zhao, and Xuan Zhou. "Big Data challenge: A data management perspective". Frontiers of Computer science, 7 (2):157-164, April 2013.
10] Divyakant Agrawal, Philip Bernstein, Elisa Bertino, Susan avidson, Umeshwar Dayal, Michael Franklin, Johannes Gehrke, Laura Haas, Alon Halevy, Jiawei Han, H.V. Jagadish, Alexandros abrinidis, Sam Madden, Yannis Papkonstantinou, Jignesh M. Patel, Raghu Ramakrishnan, Kenneth Ross, Cyrus Shahbi, Dan Suciu, Shiv Vaithyanathan, and Jennifer Widom. "Challenges and opportunities with Big Data". A Community white paper developed by leading researchers across the United States. Whitepaper, Computing Community Consortium, March 2012.
[11] Andrew McAfee and Erik Brynjolfsson. "Big Data" The management revolution". Harvard Business Review, October 2012:60-68, October 2012.
[12] Jaya Singh and Ajay Rana. "Big Data Analysis- From data to insights". International Journal of Scientific & Engineering Research, Vol. 4, No. 6, June 2013
[13] Jennifer Bachner, Predictive policing: Preventing crime with data and analytics 14, 2013.
[14] Mark H. Haller. Historical roots of police behavior: Chicago, 1890-1925, 10 L. & SOC'Y REV. 303, 303, 1976.
[15] Spatial Hadoop. http://spatialhadoop.cs.umn.edu/
[16] Elizabeth E. Joh. "Policing by numbers: Big Data and the fourth amendment". 89 Wash. L. Rev. 35, 2014. 2nd March 2014. Available at SSRN: http://ssrn.com/abstract=2403028.