Abstract no. 060-0403

The construction of judge system events based on Data Mining Technology

Sien Chen

Xiamen Xindeco Ltd,

School of Economics, Xiamen University

Xiamen, China

sandy80@vip.sina.com

Zhenyu Liu

School of Management, Xiamen University

Xiamen, China

zhenyliu@xmu.edu.cn

POMS 26th Annual Conference

Washington D.C., U.S.A.

May 8 to May 11, 2015

Abstract

Since 2012, the term“big data”has frequently been mentioned and used to describe and define the huge amount of data in the information explosive era and to name related technological development and innovation. As to the police work, the coming of big data era is not only a challenge but also an opportunity. Police agencies should go with the tide of development to start with such aspects as work thinking, top design, public information sharing and application and talent provision so as to promote the new development and progress of police work. This paper will expound the practical effect and significance of police big data application by cases happened in some areas

Keywords： Big data, Police work, Suspect locking, Informatization

Background

The data storage in public security department usually has the following features: TB level volume storage; varied and interactive systems; very complex data structure; the rapid growth of the quantity(Q, C, J,K & W, 2009).

If take the traditional method of data processing, it is definitely a demanding exercise, both time and cost-consuming-which may or may not yield results. So the big data research and construction in public security is quite necessary (Fortunatos, 2010). In recent years, many scientific and efficient big data application platforms emerge as the times require, “locking suspect in 3 seconds” is no longer a fairy tale.

Figure1-Criminal Investigation Tool for Public Security

Data filtering and Fast Suspects Locking

There are a lot query requirements（for the population, crime, immigration, the hotel internet information, vehicle / driver, fugitives, stolen vehicles and other information） in the daily work of law enforcement or security agencies. But, after many years' accumulation, these data is not only very large, and may be stored in different regions and systems. It is very difficult for the statistical work (Cloud Security Alliance, 2009). So, in actual combat operations, the police urgently need to achieve quick information query across different areas and departments, and realize the statistical analysis and comprehensive relevance utilization through timely, comprehensive and accurate information support (Perlroth&Rusli, 2012).

Study on the Law of Crime

Here is an example, "two guys who take the same train and live in the same hotel may be accomplices" (Bitterer, 2011). In the past, if policemen wanted to prove this, they need to piece together the different clues, and it takes them many hours even days to process multi-source data on the traditional architecture.However, on the big data platform, the police could analyze suspects' behavior rules before the fact (such as Internet record, hotel record, driving record) and rate their guilt possibilities. This assists the police to narrow the scope of suspects and reduce the workload.

Analysis of Vehicles on Checkpoint and Their Positions

Analysis of the vehicles on checkpoint and their position information, it not only contributes to the supervision of illegal traffic behavior, but also can find the fake -licensed car, car thefts and other illegal acts (Bitterer, 2011).

There is such a case in Beijing. A car owner sold his car to someone else while he installed GPS in advance. After the deal, he used mobile phone to track the car' location in real time and took the occasion to steal the car back. The police got the report and searched on all major online trade platforms of second-hand car. They tried all kinds of key words, locked some similar cars and analyzed these car's data one by one. Suddenly, they found that a seller's contact number was exactly the same as the former owner's. This man wanted to change a place and sell the car again! The police checked that day's video camera nearby, and found the suspect appeared by driving his own car. It's easy to finish the evidence taking （H&B, 2011）

Video Camera and Rapid Investigation

In an Urban environment, airport or mass transit hubs, there could be thousands of cameras recording the comings and goings of people. Imagine trying to locate and track a lost child, active shooter, criminal or any other person of interest across time and in a location as big as those

mentioned (Brantingham, 2011). It could take hours if not days.

But, owing to the comparison techniques of big data platform, at least 95% of irrelevant images now can be filtered out quickly(Habegger, 2010). The police could be able to track a target individual's movements from location to location, and access all relevant associated recording.

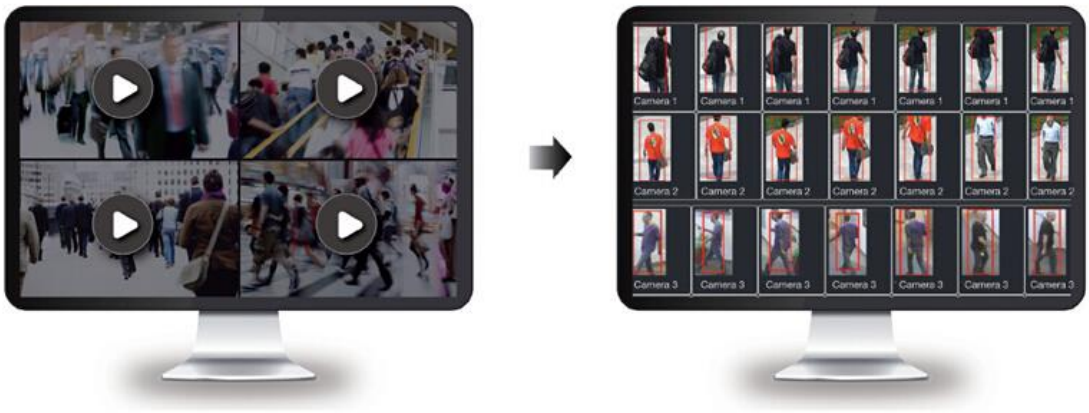

Figure2-Instead of watching hours of video, look for your target in an image album that is pre-sorted, with the top-ranked matches presented first

Information Mining and Crime Prediction

IBM (2011) said that big data can gather a great deal of information and process them, just like "the microscope and magnifier in industrial age". It helps us regain a sense of the outside world, via the neural network of big data, matrix and distributed platform (Chen,Reid, Sinai, Silke&Ganor, 2008). The flexible application of big data can not only lock suspect quickly, but also can prevent crimes.

Trends and Seasonality of Different Types of Cases

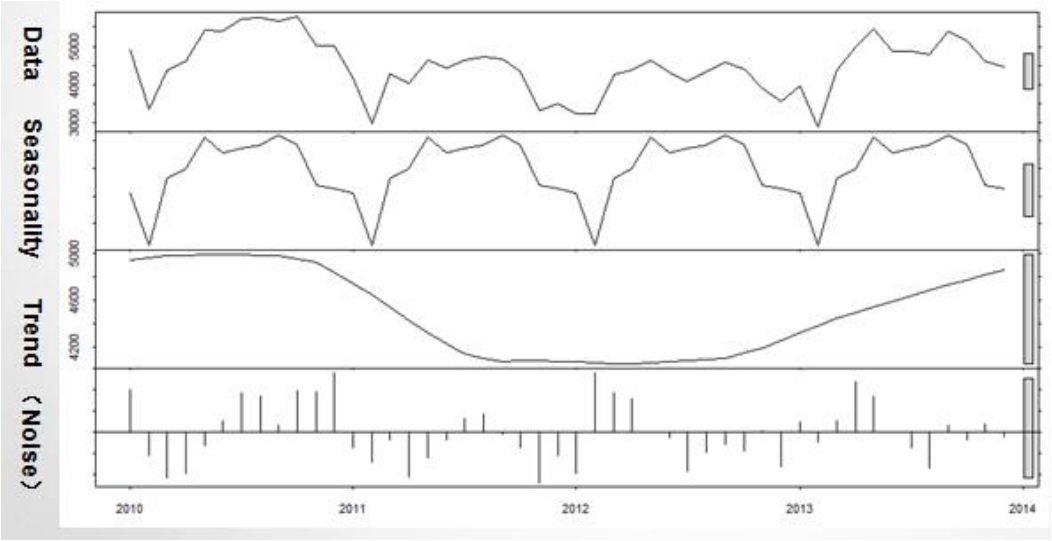

Figure3-The number of brawl and physical injury in Shenyang

Trends and seasonality are calculated by time series model. The actual value is the superposition of seasonality, trend and noise; the seasonality means the rules of 12 months within one year, and it won't change year and years. The future value can be predicted by seasonality and trends (Adomavicius&Tuzhilin, 2005).

Quantitative Forecast of Different Types of Cases

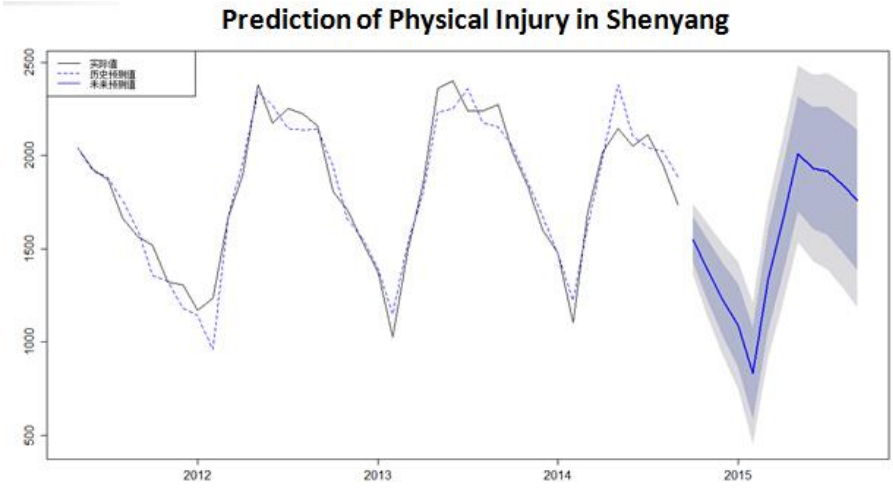

Figure4-Prediction Physical Injury in Shenyang

The black solid line in the figure is the actual value. The blue dotted line is the predicted monthly value which is just calculated by historical values. And the light blue region and light grey region on the both sides of blue solid line respectively represent 80% and 90% confidence intervals. (According to the algorithm, there will be 80% and 95% actual value fall into the predicted

interval.)

Target Prediction

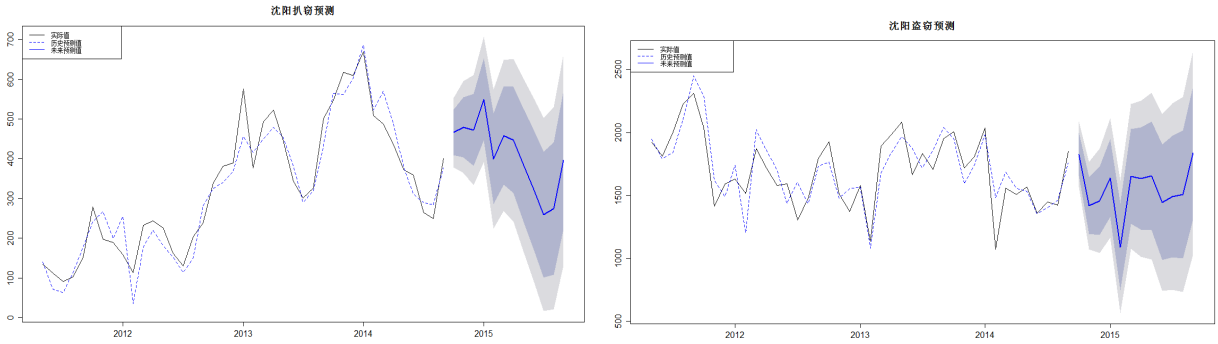

Figure5-Target Prediction in Shenyang

First of all, police resources can be rationally adjusted and allocated through the prediction of future (McAfee& Brynjolfsson, 2010). Secondly, the police can understand event causes and take manual interventions once they find the abnormal value (if the actual value is much more or less than the predicted value). Thirdly, the security services may formulate measures to slow or stop the increase when they find the cases with growth trend. Fourthly, any rules that made crimes reduced can be promoted to other regions as the experiences. At last, it helped to understand the cause of seasonal regularity and do the manual intervention (Stonebrake, Abadi, DeWitt, Maddens, Pavlo& Rasin, 2012).

Data Connection and Improving of Disaster Relief System

The application of big data can also improve the relief system for city's public security. July 21, 2012, it was pouring down in Beijing, a lot of citizens chose microblog to transmit messages. They opened the additional coordinate function, then brought about map locations, and provide convenience for the timely relief.

In the Ya'an earthquake, the missing people platform launched by Microblog, Wechat group and other major internet companies also provided many channel support for relief. Except text and images, a micro-blog or a wechat message still includes information such like time, location… National Information Society Agency (2012) shows disaster area scenes can be restored by data integration, it's also a epitome of big data application. But because of the data disconnection between websites and social media, the greatest value hasn't be developed.

Conclusion

Talking about the prospects of big data in the field of public security, data must be online, open, sharing, Internet and related (Zikopoulos, P, Eaton, DeRoos, Deutsch & Lapis, 2012). It is believed that the data standard and sharing mechanism will be more perfect in the future. The data source and the quality also will be more reliable. At that time, the big data platform which can process unstructured data (pictures, micro-blogs, the character's relationships…) will have a bigger promotion.

References:

Adomavicius, G., and Tuzhilin, A. 2005. "Toward the Next Generation of Recommernder Systems: A Survey of the State-of-the-Art and Possible Extensions," *IEEE Transcations on Knowledge and Data Engineering* (17:6), pp. 734-749.

Bitterer, A. 2011. "Hype Cycle for Business Intelligence," Gartner, Inc., Stamford, CT

Brantingham, P. L. 2011. "Computational Criminology," Keynote Address to the European Intelligence and Security Informatics Conference, Athens, Greece, September 12-14

Chen, H., Reid, E., Sinai, J., Silke, A., and Ganor, B. (eds.). 2008. *Terrorism Informatics: Knowledges Management and Data Mining for Homeland Security*, New York: Springer.

Cloud Security Alliance, "Security guidance for critical areas of focus in cloud computing," 2009, http://www.cloudsecurityalliance.org.

Fortunatos, S. 2010. "Community Detection in Graphs," Physics Reports (486:3-5), pp. 75-174

H. Shacham and B. Waters, "Compact proofs of retrievability," in Proc. of Asiacrypt 2012, vol. 5350, Dec 2011, pp. 90-107.

Habegger, B. Strategic foresight in public policy: Reviewing the experiences of the U.K., Singapore, and the Netherlands. Futures 42, 1 (Feb. 2010), 49--58.

IBM. IBM's Smarter Cities Challenge: Syracuse. Dec. 2011; http://smartercitieschallenge.org/city_syracuse_ny.html

McAfee, A. and Brynjolfsson, E. Big data: The management revolution. Harvard Business Review (Oct. 2012), 61--68.

National Information Society Agency. Evolving World on Big Data: Global Practices. May 2012; http://www.koreainformationsociety.com/2013/11/koreas-national-information-society.html

Perlroth, N., and Rusli, E. M. 2012. "Security Start-Ups Catch Fancy of Investor," New York Times, Technology Section, August 5.

Q. Wang, C. Wang, J. Li, K. Ren, and W. Lou, "Enabling public verifiability and data dynamics for storage security in cloud computing," in Proc. of ESORICS'09, Saint Malo, France, Sep. 2009.

Stonebrake, M., Abadi, D., DeWitt, D.J., Maddens, S., Pavlo, A., and Rasin, A. 2012. "Map Reduce and Parallel DBMSs: Friends or Foes," Communications of the ACM (53:1), pp. 64-71.

Zikopoulos, P.C., Eaton, C., DeRoos, D., Deutsch, T., and Lapis, G. Understanding Big Data: Analytics for Enterprise-Class Hadoop and Streaming Data. McGraw-Hill, New York, 2012.