A Proposed Bi-layer Crime Prevention Framework Using Big Data Analytics

Shu Wei Tao[1], Ooi Chong Yang[2], Mohamed Sahil bin Mohamed Salim[3], Wahidah Husain[4]

[1, 2, 3, 4] School of Computer Sciences, Universiti Sains Malaysia, USM, 11800 Penang, Malaysia
E-mail: {[1]fantasticswt, [2]c.y.world, [3]msahil.ucom14}@student.usm.my, [4]wahidah@usm.my

Abstract— **The future of science and technology sounds very promising. The need to adopt new technologies while navigating towards industry 4.0 has changed the perceptions of law enforcement agency to contend against criminal minds. It is sad but true that the conventional crime prevention system followed by government agencies is not effective for long-term implications. With advanced technologies that continuously generate and exchange data, big data analytics can be applied to predict and prevent crime from happening. However, dealing with the overwhelming amount of complex and heterogeneous crime-related data is never an easy task. There are many data analysis techniques, and each of them has its strengths and weaknesses. In order to identify the most efficient techniques, recent literature is reviewed to spotlight the trend as well as to shed light on the research gaps and challenges in various areas. The areas include crime data collection and preprocessing, crime data analysis, crime prediction and crime prevention. These techniques are further analyzed by considering the advantages and disadvantages, which then provides insight to propose a bi-layer crime prevention framework. The first layer intends to support the law enforcement agency's daily operation while the second layer serves as a countermeasure for the first layer. Both layers aim to reduce the crime rate by involving law enforcement agency through the utilization of various significant data sources and techniques effectively. The proposed crime prevention framework will progressively collect data to deter criminal behavior for the city's environmental design. Ultimately, a safe and secure city is molded shortly.**

Keywords— **data analytics; big data; crime prevention; crime prediction**

I. INTRODUCTION

Many challenges and opportunities are waiting On the eve of the fourth industrial revolution (4IR),. The advent of new technology not only improves the quality of life but also increases the criminals' abilities. The ability of one criminal to affect many is scaling exponentially. In the old days, criminals with a gun could rob a hundred people on a train. With the existence of technology, it expands the criminals' ability. According to the recent November 2015 Paris attacks, 3 bombers had made a city of 10 million people standstill [1]. Internet of Things (IoT) as one of the critical aspects of 4IR is estimated to have 30 billion devices connected in 2020 [2]. With the device connected, the exposure of data makes everyone vulnerable [3]. Undeniably, IoT even enlarges the opportunity for criminals to show their mettle.

Manual crime recordings and analysis system is no longer efficient and give little support for crime prevention. Conventional crime prevention system is not made at the same speed as new technology is developed. Whoever started the trouble should end it. It is time to utilize the technologies available in preventing crime. With billions of connected devices estimated, the amount of data generated serves as valuable sources to be fully exploited. As evidenced by 4IR, including IoT and smart city, big data analytics plays a significant role in predicting and preventing crime.

Big data analytics is a process of analyzing a variety of big dataset with the aim of uncovering hidden patterns, hidden correlations, and trends [4]. Typically, big data analytics encompasses 4 stages. The first stage in big data analytics emphasizes data collection. The second stage involves preprocessing of data collected by transforming raw data into the structured format. Consequently, the third stage of big data analytics makes use of preprocessed data by inspecting and modeling with the objective of discovering useful information. The last stage is where techniques and algorithms are applied to analyzed data in order to predict and forecast about future events.

In this paper, a crime prevention framework with the use of big data analytics is presented. To most effectively explore the way big data helps in preventing crime, the following objectives have been pursued.

- To find and discover efficient ways of current techniques used in big data analytics.

- To propose a crime prevention framework by involving law enforcement agency through big data analytics.

II. MATERIAL AND METHOD

This section introduces the techniques studied throughout the big data analytics processes and crime prevention strategies.

A. Crime Data Collection & Preprocessing

All collected data will be preprocessed before storing in the database. Except for the structured data source from crime-related government organizations, published crime open data and call data records, the collection of data from social network services, newspaper articles, and public information repository are unstructured and incomplete [5]-[11]. The data collection techniques used to collect unstructured data is mainly web crawler [5], [9], [10]. After the collection of unstructured data, Isuru et al. used a classifier, entity extractor, and duplicate detector techniques to preprocess the data [10]. Likewise, Nasution et al. and Jeon et al. used map-reduce while Zhang used intelligence analysis for data preprocessing [5], [9], [11]. Despite various techniques used during data preprocessing, all the techniques are aimed to reduce and eliminate the duplication of data during data collection. For structured data, preprocessing techniques are used to transform the data into the desired format. The techniques used are XML parser, metadata, map reduce and GIS raster data analysis [5]-[8]. Table I is the summary of researches in crime data collection and preprocessing.

B. Crime Data Analysis

Each research papers to analyze the preprocessed data use various data analysis techniques. Preprocessed data that involves the usage of crawler when collecting, data especially from web environment commonly implements text mining technique for crime analysis. As evidenced by text mining techniques used by Jeon et al., it helps in analyzing natural languages using a statistical algorithm [5]. Text mining technique is also used together with opinion analysis to classify the polarity of text-mined, which results

in identifying the 10 highest number of counts for crime danger areas [5].

There are many data analysis techniques, which can be used to analyze crime data effectively. This includes, but not limited to spectral analysis, SQL queries analysis, correlation analysis and forensic analysis [7], [8], [11], [12]. Luca et al. applied spectroscopy principle to highlight the seasonal components of the crime dataset. Spectroscopy represents data in spectrum format allowing the identification of significant periods of crime to happen. The significant periods can be acknowledged from periodogram produced which drives towards the inspection of weekly patterns using a heat map [7]. Subsequently, SQL queries analysis also intends to unearth the suspicious individual which results in an increment of overall efficiency using the proposed solution [8]. Zhang used correlation analysis in exercising rules of data association in order to identify the hidden relationship between crimes [11]. Besides, Umar et al. utilized forensic analysis to investigate crimes by extracting information in a repeatability and reproducibility manner [12].

K-mean algorithm, a traditional clustering algorithm is also suitable for data analysis. Teng improved the existing k-mean algorithm by combining genetic algorithm [13]. Similarly, Turki et al. modified the algorithm by featuring an alternative way of recognizing the optimal number of clusters [14]. Both have modified k-mean algorithm in different methods to cluster the crime pattern. Despite the different approaches taken, both ideas have been successfully implemented where the performance of the modified k-mean algorithm is improved. The only drawback was that the absence of crime prediction and prevention model using the algorithm proposed. Table II is the summary of researches in crime data analysis.

C. Crime Prediction

Crime prediction is to forecast an event that might happen anytime in the future. Crime Pattern Theory (CPT) states that crime can only occur if it involves two people's personal spaces intersecting at one point or another [15].

Soumya et al. implemented an artificial neural network (ANN) to predict crime depending on a series of input [15].

TABLE I
SUMMARY OF RESEARCHES IN CRIME DATA COLLECTION & PREPROCESSING

Studies	Data Types	Data Sources	Data Collection Techniques	Data Preprocessing Techniques
[9]	Unstructured	Online and social media	Web crawler	Map-reduce
[5]	Unstructured	Crime-related Social Network Services	Web crawler	Map-reduce
[10]	Unstructured	Sri Lanka English Newspapers (Daily Mirror, The Island, and Ceylon Today)	Web crawler	Document classifier, entity extractor, and duplicate detector
[11]	Unstructured	Public security information system database and public information repository	Database connection	Intelligence analysis
[5]	Structured	Crime-related government organizations	System connection	XML parser
[6]	Structured	Crime data from the Seattle Police Department, Build environment data (King Country GIS datasets), U.S Census data	A web crawler, Database connection	GIS raster data analysis
[7]	Structured	All crime events in San Francisco from January 2013 to August 2016	Portal acquisition	Metadata
[8]	Structured	Call Data Records (CDR) from telecommunication companies	Request and authorization	Map-reduce

TABLE II
SUMMARY OF RESEARCHES IN CRIME DATA ANALYSIS

Studies	Techniques	Tasks	Results	Research Gaps	Research Challenges
[5]	Text mining and opinion analysis	To analyze natural languages and classify polarity of text for crime danger analysis	10 highest number of counts of crime danger areas are identified	Lack of integration between big data and IoT	Privacy invasion, private sector data copyright, big data quality, and structure performance
[7]	Spectral analysis and heatmap	To highlight seasonal components	Periodogram of crime with detrending is produced	Need for consistent observations	Analyze finer granularities of time and space
[8]	SQL queries analysis	To identify most suspicious individual	Better in technical, economic and management efficiency	Gloss over in analysis procedure	Improves the efficiency in battling crime
[11]	Correlation analysis	To evident relationship and chain hidden	No result	No crime model and visualization	The large-scale database, structured data, and privacy issues
[12]	Forensic analysis	To extract information from WhatsApp	Meet validation test of repeatability and reproducibility	No crime prediction and crime prevention model	Extraction abilities in non-Android platforms
[13]	Improved k-means based on genetic algorithm	To analyze the abnormal capital flow	Achieve a 91.3% of average accuracy (Improved 2%)	No crime prediction and crime prevention model	Improves the accuracy of parameter tuning
[14]	Modified k-means	To quickly identify and initialize centroid point for better clustering	Faster in identifying some clusters accurately	No crime prediction and crime prevention model	Predict new crime pattern

The termination of training data is determined through a gamma test, which is used to reduce the noise from the data and to determine the exact terminating point [15]. The challenge is inclined towards the quality of data and unable to do prediction if the input is the source of the forecasted result.

Contrast to ANN that is suitable for long-term prediction, Autoregressive integrated moving average (ARIMA) and exponential smoothing gives a more accurate forecast of crime over short-term of crime series [16], [17]. The main advantage of ARIMA over exponential smoothing is that it can predict the sudden changes based on the input given [17].

TABLE III
SUMMARY OF RESEARCHES IN CRIME PREDICTION

Studies	Techniques	Tasks	Results	Research Gaps	Research Challenges
[15]	Artificial neural networks	To analyze the collection of different processing elements	Graph of crime which shows the actual crime and the predicted crime	Unable to identify which factor of the input provide the predicted output	An inconsistent and incomplete form of the data to be processed
[16]	Exponential smoothing	To fit the trend of series into two types, SES and HES	SES and HES provides more accurate results compared to real sample data	Does not forecast with a large dataset and sudden changes in crime pattern well	Value of smoothing coefficients and the series will provide an ideal prediction
[17]	Autoregressive integrated moving average (ARIMA)	To identify the ARIMA structure, estimate the coefficient value that fit the test to forecast the crime	Fit the series well, able to pick up the turning points of the series of crime and accurate results	Only short-term forecast of crime series but requires at least 50 or more dataset	Identification of the ARIMA structure and estimation of the coefficient value
[18]	Nonlinear autoregressive with exogenous (external) input (NARX)	To predict the next crime suspect's biography and time	The plot of next serial cases for suspect age by google earth	Does not forecast well if incomplete suspect's biography dataset	The result of prediction highly depends on the correctness of suspect's biography
[19]	Point-process model	To analyze the relationships of time and space to get a specific point of view	The result of data-driven support to predict future crimes	Narrow the research for future deployment	Specific time and duration in crime dataset can affect the results of prediction
[20]	Linear regression	To forecast the status of crime in a yearly manner	Graph of number of crimes forecasted in the following year for different regions	No crime prediction on the location, only the number of crimes	Specialized software required which can work with the statistics well for a proper output
[21]	Apriori algorithm	To reduce the number of item set of crimes	Prediction of crime where the individual will commit next in future	Cannot predict a new type of crime the individual can commit	High cost and time consuming to reveal the crime pattern

There is another neural network approach used by Ghazvini et al. [18]. Ghazvini et al. applied nonlinear autoregressive with exogenous input (NARX) to predict the next crime suspect's biography and time [18]. Similarly, Xu uses Point-process model to predict the future crime from the spatial-temporal dataset of crime [19]. As can be seen, both make use of time attribute from the dataset to drive the crime prediction.

Awal et al. implemented another crime prediction technique known as linear regression method [20]. The primary objective of linear regression is to minimize the cost function so that the hypothesis of direct regression is aligned with the output of the training dataset [20].

Apriori algorithm is another crime prediction technique where the itemset is used for a more successive generation, which predicts the next crime an individual might commit [21]. The central concept of the apriori algorithm is that if an itemset is frequent, then the other subsets in the dataset will also be frequent [21]. Table III is the summary of researches in crime prediction.

D. Crime Prevention

Crime prediction is the attempt to reduce and deter crime and criminals. Crime prevention embraces strategies and actions that seek to reduce the rate of crimes [22].

There are several methods found in recent research about crime prevention which is improving police patrol routes, crime prevention based on agent-based simulation model and integrating crime prevention through environmental design [6], [23]-[26].

Xu et al. used cross-entropy (CE) algorithm to design randomized patrol routes schedule in dynamic patrol environment [23]. Randomize strategy has been applied in the CE algorithm to avoid criminals to be sure about the patrol routes. Likewise, Li et al. aimed to improve the performance of CE algorithm by integrating Getis-ORD Gi*

algorithm that calculates the likelihood for each segment and set up transition matrix to obtain the optimal set of patrol routes [24]. Though both strategies have the similarity where police patrol routes are wavered according to the predicted crime hotspot, enhanced CE algorithm can be used in real time application where the transition matrix update is based on real-time crime likelihood, and the patrol routes are wavered accordingly [24].

In the same way, Danilo et al. implemented the genetic algorithm (GA) to achieve optimal police patrol route by identifying a set of patrol routes to reduce the occurrence of crime [25]. The GA algorithm uses the Darwinian theory of evolution by natural selection to find the optimal patrol route [25]. The main achievement of this paper is using a GA to prove that the best route sets found has higher probability where crime will be highlighted.

Bosse et al. practiced agent-based technique to address the concern on the emergence of crime hot-spots and prevents them accordingly [26]. The main objective of this paper is to find out the best strategy of the movement of the guardians based on the situation catered by the criminals and passer-by. The safety of location with an estimated increment in the crowd will be enforced by increasing the number of guardians [26]. Vividly, crime prevention methods discussed utilizes the forecasted results which then enforce the safety of the community in a dynamic approach.

However, an entirely different approach is worth mentioning. Instead of utilizing forecasted results, Crime Prevention through Environmental Design (CPTED) studied by Sohn depends entirely on results from data analysis over a long period of observation on the relationship between the occurrence of crime and the building design [6]. It makes use of the data and found out the culprit which encourage the opportunity for crime to happen. Table IV is the summary of researches in crime prevention.

TABLE IV
SUMMARY OF RESEARCHES IN CRIME PREDICTION

Studies	Techniques	Tasks	Results	Research Gaps	Research Challenges
[6]	Crime Prevention Through Environmental Design (CPTED)	To increase the controllability of crime attractors and preventers in the neighborhood	Residential crime density negatively related to the density of bus stop but positively related to the street intersection density	No crime prevention model	Data collected in a short period may affect the long-term impact of CPTED strategies
[23]	Cross-Entropy (CE)	To design randomize patrol routes schedule strategy to fit into dynamic patrol environment	CE algorithm offer 20% - 40 % higher reward than greedy strategy and Hamilton strategy	Lack of explanation on the application of CE in real time application	Multiple patrol unit planning in a single patrol area
[24]	Enhanced cross entropy with Getis-ORD Gi* algorithm	To improve the police patrol service performance and apply it in real-time applications	Patrol routes suggested help in achieving near-optimal coverage of the patrol area	Lack of explanation on how it can be applied in the real-time application	Ensure the frequency of edge visited is proportional to the crime likelihood
[25]	Genetic algorithms	To automatically discover crime hotspots	GAPatrol could identify hotspots well	Lack of parameter tuning	Criminals' non-deterministic behavior
[26]	Agent-based simulation	To compare the effectiveness of different strategies for guardian movements	Movement of the guardian to the new place where the density of passers-by increases in future	Lack of justification for the simulation results	Some parameters not always be feasible and exact leads to the simulation result cannot be generalized

III. RESULT AND DISCUSSION

To effectively contend with crime from happening, a Bi-layer Crime Prevention Framework (BCPF) is proposed. BCPF aims to reduce the crime rate by involving law enforcement agency (LEA) through the utilization of various significant data sources and techniques effectively. The main concept is to maximize the LEA's operational efficiency and in the meantime, gradually deter criminal behavior through environmental design.

A. Framework Overview

BCPF, as the name implies, is a framework that prevents crime in two layers.

- Layer 1: It discerns crime hotspot areas by analyzing big data through text mining. Short-term and long-term crime prediction techniques are used jointly to predict crime, which supports the planning of LEA's daily operation.
- Layer 2: It selectively activates for further crime analytics when the first layer fails. It performs crime analytics on additional data sources from the telecommunication company to allocate the most suspicious individual, which enables LEA's operation planning in a dynamic environment.

Simultaneously, both layers will disclose valuable information with the goal of deterring crime through environmental design [6]. Fig. 1 illustrates the overview of BCPF framework.
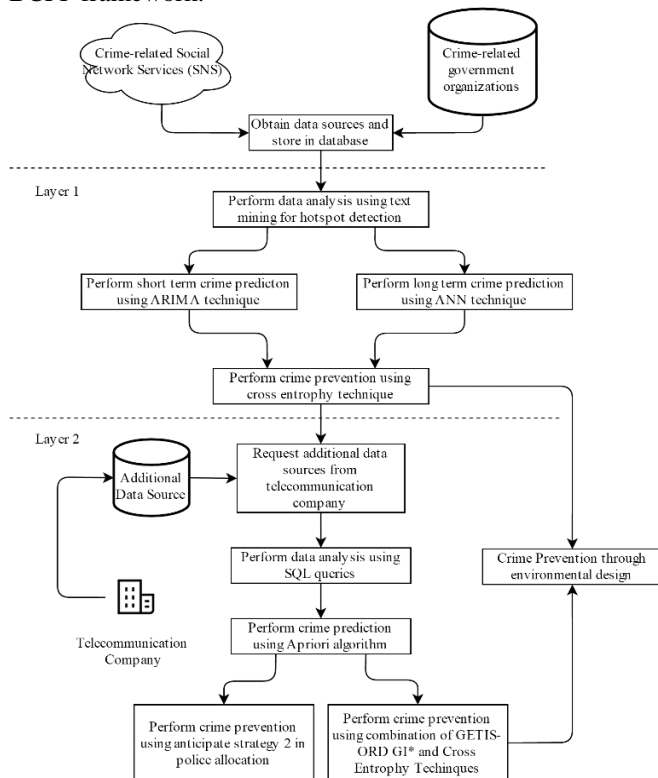
Fig. 1 BCPF Framework Overview
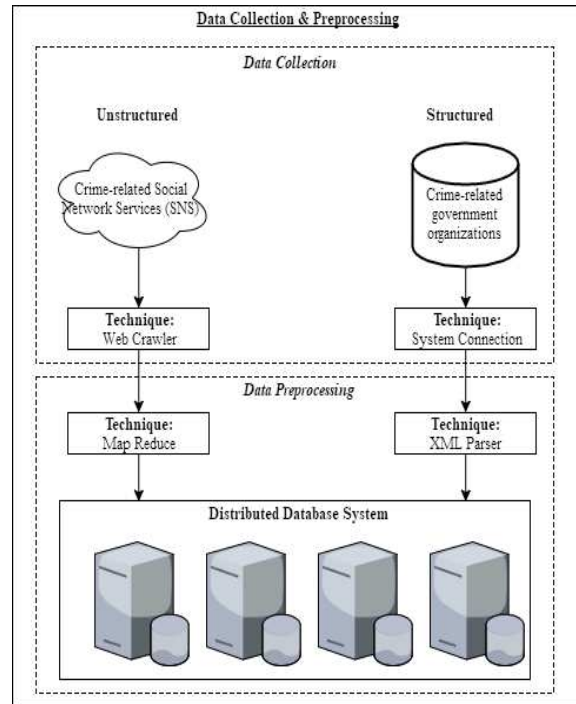
B. Data Collection & Preprocessing

Fig. 2 Data Collection & Data Preprocessing of BCPF

Fig. 2 illustrates the data collection and preprocessing processes and techniques used in BCPF. Crime data is collected from Social Network Services (SNS) and government organizations. The formal crime data is unstructured, and the recent crime data is structured. The techniques used are depending on its data types. SNS data is collected by a web crawler while the government organization data is collected by system connections [5].

The collected SNS data is preprocessed using map reduce technique, which aims to reduce and eliminate the duplication of data. This can be reasoned by the design of crawler, which iteratively crawls data from the World Wide Web where different web pages with the same content are crawled and duplicated [27]. As for collected government organization data, the data is parsed with XML parser to the desired format before storing in the database.

C. BCPF Layer 1

From the crime data collected and preprocessed, text mining analysis technique is used for hotspot detection. Text mining is used to omit words and to determine crime and location of related words. With the crime and location classified, the hotspot detection will be identified according to the frequency of the words appearing throughout the data in text mining. The highest frequency of crime appearing will serve as a benchmark for assigning color on the map. The input will be the preprocessed data above for the crime prediction. Predictions fall into two categories: long-term and short-term. For the long-term, Artificial Neural Network (ANN) method will be used. The input is the crime cluster dataset consisting of crime type, the number of occurrences, and the cluster hotspots map with centroid [15]. This will

output accurate annual prediction of crime, which can help the LEA to focus on dividing the task among them.
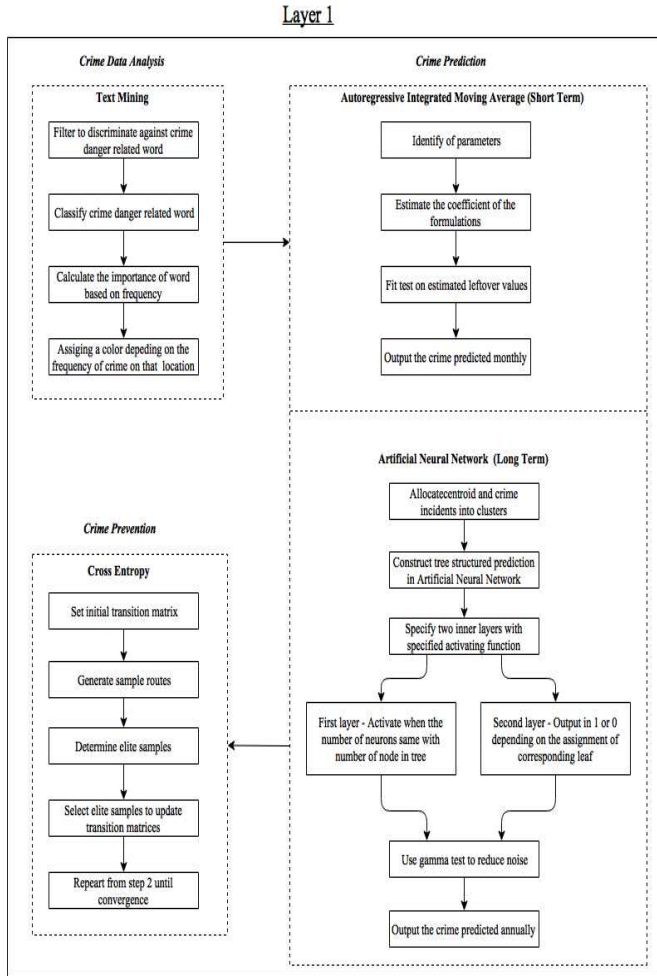

Fig. 3 BCPF Layer 1

LEA cannot depend on the annual prediction and keep the plan unchanged throughout the year. Changes might happen to affect the prediction, which leads to short-term prediction. For the short-term prediction, Autoregressive integrated moving average (ARIMA) method is used. ARIMA model with the input of the previous month of crime cluster data will predict the sudden changes in the dataset which will provide a more accurate prediction for the following month [16], [17].

Predicted data will be utilized by integrating them as the guidance for crime prevention. Crime prevention strategy will encompass two components, which focus on daily operational planning and long-term environmental design. The formal engages cross-entropy method and randomize strategy with the data predicted in crime prediction for patrol routes planning, while the latter will utilize the data collected in daily operation that observe the attributes of environment design of the crime scene and use it for city planning.

D. BCPF Layer 2

Fig. 4 illustrates the detail techniques and the steps proposed in big data analytics as well as some prevention strategies in Layer 2. For further data analysis, the LEA requests the higher authority for the Call Data Records

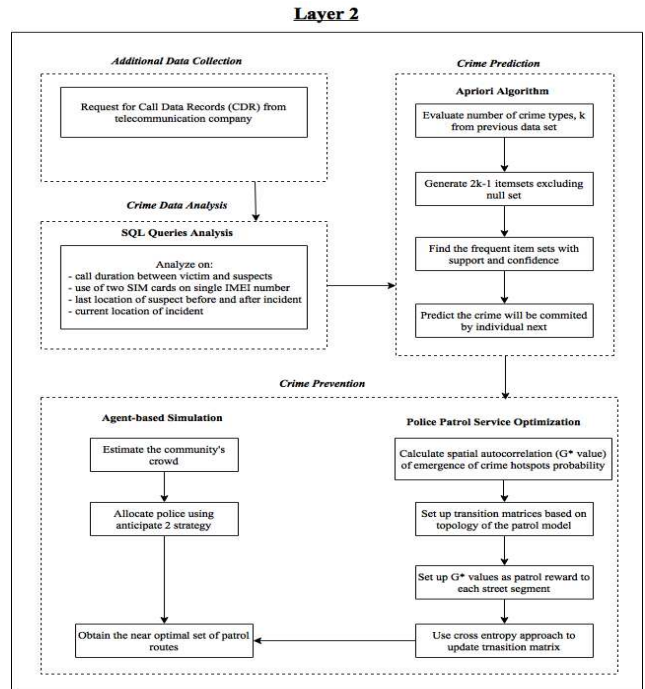(CDR) of the mobile numbers of the area where the crime is committed [8].

Fig. 4 BCPF Layer 2

SQL queries analysis is applied to the CDR collected. The details of CDR are shown in Table 1.

TABLE I
INFORMATION IN CDR

Call Data Record (CDR)
Phone Number of Subscriber (on who's name the Number is Registered)
Phone Numbers of the parties on the other end of the Calls.
Time and Date of Calls
Call Durations
Sequence Number
Route by which Call Entered the Exchange
Route by Which the call left the Exchange

With the required criterions, the result of the suspicious target is then given as output. The output will then be used to predict the possibility of crime an individual will be committing in future through an Apriori algorithm. Apriori algorithm is an algorithm used to reduce the number of the item set in the sense that if the item set is frequent for the individual, then the individual must also be frequent to all the subsets [21].

In order to enhance the crime prevention strategy, layer 2 will have a broader focus area, which involves the allocation of police. Some police will be allocated according to the crime hotspots predicted. Meanwhile, for the patrol routes planning, a combination of Getis-ORD Gi* score and cross entropy methods will be used and applied in a dynamic environment. Getis-ORD Gi* score will be given according to the predicted probability of crime emerging in each street [24]. Turnout data from the Apriori algorithm will not only be used to predict crime but also be used to provide insight for the community's estimation crowd. The safety of location with an estimated increment in the crowd will be enforced by increasing the number of police [26]. Data collection will continue to feed and exploit for city planning.

E. Discussion

This BCPF is proposed after courteous consideration from each of the techniques used by other researchers in big data analytics. It envelopes from the crime data collection and preprocessing stage, to crime data analysis, followed by crime prediction and prevention.

The rationale of having the second layer in the BCPF is to aid the LEA operational planning in an immediate and dynamic environment. There is no absolute prediction when the prediction is based on the historical dataset. As long as there is a new criminal or crime-taking place, the prediction is no longer feasible. At this moment, the second layer serves as a countermeasure where it performs additional data analytics to counter this problem.

IV. CONCLUSION

Crime is expected to increase by the improvement of technology that provides the convenience for the criminal to commit crime. There is always a challenge to prevent crime using big data analytics by handling the enormous volume of data generated from various sources.

In this paper, a bi-layer crime prevention framework using big data analytics is presented. The uniqueness of this framework is the proper strategy designed to address the challenges, whereby the first and second layers not only ease the LEA's operational planning, but also continuously deliver valuable data for city's environmental design. Hence, a safe and secure city is molded in the future.

Further investigations would need to drive artificial intelligence techniques into big data analytics. To contend with criminal minds, big data analytics with artificial intelligence might shape in the future. Besides, future research will be considered to focus on preventing cybercrime as recent WannaCry ransomware attack that affects more than 230,000 individuals across 150 countries is believed to be a nightmare and might resurface in the future [28].

REFERENCES

[1] (2017) UrbiStat AdminStat in Demography [Online]. Available: https://ugeo.urbistat.com/AdminStat/en/fr/demografia/dati-sintesi/paris/75/3

[2] (2017) "IoT: number of connecting devices worldwide 2012-2025 | Statista", Statista [Online]. Available: https://www.statista.com/statistics/471264/iot-number-of-connected-devices-worldwide/

[3] G. Marc, *Future crimes: Everything is connected, everyone is vulnerable and what we can do about it,* 2015.

[4] (2017) TechTarget Search Business Analytics [Online]. Available: http://searchbusinessanalytics.techtarget.com/definition/big-data-analytics

[5] J. Jeon, and S. Jeong, "Designing a Crime-Prevention System by Converging Big Data and IoT," in *Journal of Internet Computing and Services (JICS)*, 2016, pp. 115-128.

[6] D. W. Sohn, "D. (2016). Residential crimes and neighborhood built environment: Assessing the effectiveness of crime prevention through environmental design (CPTED)," 2016). Cities, 52, pp.86-93.

[7] V. Luca, B. Elena, and T. Politecnico, "A spectral analysis of crimes in San Francisco," 2016.

[8] K. S. Er., A. Firoz, A. D. Hamza, and Sabigua, "Criminal Investigation Using Call Data Records (CDR) through Big Data Technology," in *International Conference on Nascent Technologies in the Engineering Field*, 2017.

[9] F. B. B. Nasution, N. E. N. Bazin, Daliyusmanto, and A. Zulfikar, "Big Data's Tools for Internet Data Analytics: Modelling of System Dynamics," *International Journal on Advanced Science, Engineering and Information Technology,* vol. 7, no. 3, pp. 745-753, 2017. [Online]. Available: http://dx.doi.org/10.18517/ijaseit.07.3.1088

[10] J. Isuru, S. Chamath, L. Sampath, W. Tharindu, P. Indika and W. Adeesha, "Crime Analytics: Analysis of Crimes Through Newspaper Articles, 2015.

[11] Y. Zhang, "Analysis of Crime Factors Correlation Based on Data Mining Technology," in *International Conference on Robots & Intelligent System*, 2016.

[12] R. Umar, I. Riadi, and G. M. Zamroni, "Mobile Forensic Tools Evaluation for Digital Crime Investigation," *International Journal on Advanced Science, Engineering and Information Technology,* vol. 8, no.3, pp. 949-955, 2018. [Online]. Available: http://dx.doi.org/10.18157/ijaseit.8.3.3591

[13] L. Teng, "Criminal Behavior Analysis Method based on Data Mining Technology," in *International Conference on Smart City and System Engineering*, 2016.

[14] A. Turki, S. Daming, W. David, and W. William, "Criminal Pattern Identification Based on Modified K-means Clustering," in *International Conference on Machine Learning and Cybernetics, Jeju, South Korea*, 2016.

[15] Soumya and A. S. Baghel, "A Predictive Model for Mapping Crime using Big Data Analytics," vol. 04, no. 04, pp. 344–348, Apr. 2015.

[16] P. Chen, H. Yuan, and X. Shu, "Forecasting Crime Using the ARIMA Model," in *Fifth International Conference on Fuzzy Systems and Knowledge Discovery*, Nov. 2008.

[17] N. M. M. Noor, A. Retnowardh, M. L. Abd, and M. Y. M. Saman, "Crime Forecasting using ARIMA Model and Fuzzy Alpha-cut," I in *Journal of Applied Sciences*, vol. 13, no. 1, pp. 167–172, Feb. 2013.

[18] A. Ghazvini, M. Z. B. A. Nazri, S. N. H. S. Abdullah, M. N. Junoh and Z. Abidin bin Kasim, "Biography commercial serial crime analysis using enhanced dynamic neural network," *2015 7th International Conference of Soft Computing and Pattern Recognition (SoCPaR)*, Fukuoka, 2015, pp. 334-339.

[19] J. Xu, "Predict Future Events with Point-process Modeling," IBM Big Data & Analytics Hub, 18-Sep-2015. [Online]. Available: http://www.ibmbigdatahub.com/blog/predict-future-events-point-process-modeling

[20] M. A. Awal, J. Rabbi, S. I. Hossain, and M. M. A. Hashem, "Using Linear Regression to Forecast Future Trends in Crime of Bangladesh," in *5th International Conference on Informatics, Electronics and Vision (ICIEV)*, 2016.

[21] A. Agarwal, D. Chougule, A. Agarwal, and D. Chimote, "Application for Analysis and Prediction of Crime Data using Data Mining," in *International Journal of Advanced Computational Engineering and Networking*, ISSN: 2320-2106., vol. 4, no. 5, pp. 9–12, May 2016.

[22] (2017) "United Nations Office on Drugs and Crime [Online]. Available: https://www.unodc.org/unodc/en/justice-and-prison-reform/CrimePrevention.html

[23] C. Xu and T. S. P. Yum, "Cross Entropy approach for patrol route planning in dynamic environments," *in 2010 IEEE International Conference on Intelligence and Security Informatics, Vancouver*, BC, 2010, pp. 114-119.

[24] L. Li, Z. Jiang, N. Duan, W. Dong, K. Hu, and W. Sun "Police patrol service optimization based on the spatial pattern of hotspots," in *Service Operations Logistics and Informatics (SOLI) 2011 IEEE International Conference*, on, pp. 45-50 July 2011.

[25] R. Danilo, M. Adriano, L. V. C. André and F. Vasco, "Towards Optimal Police Patrol Routes with Genetic Algorithms," in *S. Mehrotra et al. (Eds.): ISI 2006*, 2006.

[26] T. Bosse and C. Gerritsen, "Comparing Crime Prevention Strategies by Agent-Based Simulation," in *2009 IEEE/WIC/ACM International Joint Conference on Web Intelligence and Intelligent Agent Technology*, Milan, Italy, 2009, pp. 491-496.

[27] (2012) "Web Crawling: Data Scraping vs. Data Crawling [Online] Available: https://www.promptcloud.com/data-scraping-vs-data-crawling/

[28] CERT-MU, *The WannaCry Ransomeware: White Paper*, May 2017.