Architecture of Decision Support System for Crime Visualization

Noor Maizura Mohamad Noor, Siti Haslini Ab Hamid, Rosmayati Mohemad, Masita @ Masila Abdul Jalil and Muhammad Suzuri Hitam
School of Informatics and Applied Mathematics, Universiti Malaysia Terengganu, 21030, Kuala Terengganu, Terengganu, Malaysia

ARTICLE INFO
Article History:
Received: June 16, 2015
Accepted: September 14, 2015

Corresponding Author:
Noor Maizura Mohamad Noor
School of Informatics and Applied Mathematics,
Universiti Malaysia Terengganu, 21030, Kuala Terengganu, Terengganu, Malaysia
Tel: +6016-9806310

ABSTRACT
Recently, crime rate in Malaysia is increasing from day to day. If this situation is not monitored and there is no drastic action taken, it may cause many serious problems. Therefore, crime prevention is one of the important components of an overall strategy to reduce crime and to strengthen public security. However, the main challenge faced by all law-enforcement and intelligence-gathering organizations is to analyze the growing volumes of crime data accurately and efficiently in order to make decision for crime prevention. Due to difficulty in decision making for crime prevention, Decision Support System (DSS) and data mining approach can be used to resolve the problem. Now-a-days, data mining approach has been exposed to be a practical decision-support concept in predicting and preventing crime. As a result, we propose architecture of DSS using visualization technique because it can represent crime data into more comprehensible presentation. The results from this proposed architecture can support the security authority in assessing more suitable law enforcement strategies, increasing the accuracy of selection decision making, as well as improving the use of security authority duty deployment for crime prevention.

Key words: Decision support system, visualization, self-organizing map, crime prevention

INTRODUCTION
Crime refers to an act that violates the rules of law, an act that is considered prohibited or illegal and punishable by law; immoral or deviant actions and behavior as well as a serious offense. Recently, the crime rate in Malaysia is increasing from day to day (Sakip and Abdullah, 2012). If this situation is not monitored and there is no drastic action taken, it may cause many serious problems. Therefore, one of the Malaysia government agenda as stated in National Key Result Area (NKRA) has been introduced to reduce crime index and improve the public safety (Ismail and Ramli, 2013). Crime prevention is one of the important components of an overall strategy to reduce crime and to strengthen public security. Crime prevention refers to all actions and efforts that aim to reduce crime and fear of crime (Marzbali et al., 2012).

A wide range of unstructured data related to crime rapidly cumulate over a long period of time. In addition, the main challenge faced by all law-enforcement and intelligence gathering organizations is to analyze the growing volumes of crime data accurately and efficiently. In Malaysia, the decision making process for crime prevention is done manually by the security authorities. Based on the crime index, the security authorities will determine the level of crime at the level of high, medium or low. Thereafter, they will determine the most appropriate prevention decision according to the level of crime. The selection of the right decision for crime prevention is expected to be more appropriate in the real conditions, where each crime level has a different crime prevention decision. Moreover, they will compare the current year crime volume with the previous year crime volume for each crime type. The increment for each crime type cannot be
techniques which contain information on crime type, data mining class and technique and proposed architecture. Some of the data mining classes can address more than one crime type and therefore we grouped this table based on crime type. For example, Phillips and Lee (2012) presented a framework that allows autonomous exploratory analysis and knowledge discovery in massive real aggregated crime dataset. They used visualization technique to extract and visualize crime patterns efficiently. Li et al. (2010) applied Fuzzy Self-Organizing Map (FSOM) and rule extraction algorithm to analyze crime statistical data for 20 country police bureaus in Taiwan from 2003-2004 and produced results that can help police in criminal incidents prevention. They combined visualization and clustering techniques. But, the limitation of the proposed model is difficult to evaluate the accurate performance. Ku et al. (2012) developed crime report visualization-Textual Analysis of Similar Crimes (TASC) using combination of visualization and clustering techniques to support crime analysts analyze, compare and contrast crime report in timely manner. However, the limitation of TASC is can only handle visualization of text highlighting for simple analytical tasks. Moreover, Palocsay et al. (2000) applied neural network to generate case-by-case results in criminal recidivism but their results are depending heavily on the choice of the network topology. Grubesic (2006) used fuzzy clustering for hot-spot detection but he did not mention specific architecture. Kaikhah and Doddameiti (2006) presented a novel knowledge discovery technique to discover trends using clustering technique for murder, rape, robbery and auto theft crime in United States cities. Furthermore, Yu et al. (2011) applied classification technique for burglary crime. They studied, compared and proposed the best forecasting approach to achieve the most stable results.

Based on Table 1, some of the data mining classes that have been applied in crime researches are visualization, clustering, prediction and classification. For data mining techniques, the researchers used techniques like neural network, fuzzy clustering and naive bayes. Several data mining classes such as visualization, prediction and clustering cover all crime types. However, in term of proposed architecture, some of the crime researches did not propose or mention specific architecture in their works. From the literature (Chen et al., 2004), visualization techniques have been applied in crime researches are visualization, clustering, classification technique for burglary crime. They studied, compared and proposed the best forecasting approach to achieve the most stable results.

Table 1: Several crime researches using data mining techniques

<table>
<thead>
<tr>
<th>Crime type</th>
<th>Data mining class</th>
<th>Data mining technique</th>
<th>Proposed architecture</th>
</tr>
</thead>
<tbody>
<tr>
<td>All crime types</td>
<td>Visualization</td>
<td>Global minimum edge weight, kruskal’s algorithm (Phillips and Lee, 2012)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Visualization, clustering</td>
<td>Word/tag cloud, tree-based graphs, networks (node-link diagrams), circle (radial) graphs (Ku et al., 2012)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td></td>
<td>Fuzzy self-organizing map (FSOM), fuzzy c-means (FCM) (Li et al., 2010)</td>
<td>Yes</td>
</tr>
<tr>
<td></td>
<td>Prediction</td>
<td>Neural network (Palocsay et al., 2000)</td>
<td>No</td>
</tr>
<tr>
<td></td>
<td>Clustering</td>
<td>Fuzzy clustering (Grubesic, 2006)</td>
<td>No</td>
</tr>
<tr>
<td>Burglary</td>
<td>Classification</td>
<td>One nearest neighbor (1NN) and a location constrained variation, decision tree (4J), Support Vector Machine (SVM) with radial basis function, neural network, naive bayes (Yu et al., 2011)</td>
<td>Yes</td>
</tr>
<tr>
<td>Murder, rape, robbery and auto theft</td>
<td>Clustering</td>
<td>Neural network (Kaikhah and Doddameiti, 2006)</td>
<td>No</td>
</tr>
</tbody>
</table>

more than five percents. The manual process for decision making for crime prevention will produce inaccurate and inefficient result. That requires a method to help the security authorities to determine the most appropriate decisions.

Over the past few years, a number of article on crime researches have appeared in conferences and journal publications. Researchers have developed many methods to support decision making for crime prevention over the last two decades. These works originate from disciplines of behavior and psychology, statistic and artificial intelligence (Li et al., 2010). The approach of behavior science and psychology has been aimed at preventing individuals from committing crimes since the 1960s. This traditional approach depends on the expertise and tactic knowledge of specialists, which easily leads to criminal experts’ fatigue, misjudgment and slow response. Moreover, it usually lacks of the real data for verification, thus making it difficult to apply the findings to the actual policy. Therefore, researchers have continued to find new improved methods in this area especially in artificial intelligence.

Due to difficulty in decision making for crime prevention, Decision Support System (DSS) has become increasingly important to the security authority. The DSS can be defined as an integrated, interactive computer system, consisting of analytical tools and information management capabilities, designed to aid decision makers in solving relatively large and unstructured problems (Sieker et al., 2006). In addition, crime analysis can make an important contribution to the delivery of an effective and efficient security authority services and to how the security authorities and their partners tackle crime. It also provides valuable information that is fundamental to internal management decisions and policy making. To handle crime analysis, data mining is a great tool that enables criminal investigators who may lack extensive training as data analysts to discover huge databases quickly and efficiently (Chen et al., 2004). In addition, data mining approach has been exposed to be a practical decision-support concept in predicting and preventing crime. Data mining contains six common classes of tasks which include clustering, classification, outlier detection, prediction, regression and visualization (Ngai et al., 2011). Table 1 analyzes several crime researches using data mining techniques which contain information on crime type, data mining class and technique and proposed architecture. Some of the data mining classes can address more than one crime type and therefore we grouped this table based on crime type. For example, Phillips and Lee (2012) presented a framework that allows autonomous exploratory analysis and knowledge discovery in massive real aggregated crime dataset. They used visualization technique to extract and visualize crime patterns efficiently. Li et al. (2010) applied Fuzzy Self-Organizing Map (FSOM) and rule extraction algorithm to analyze crime statistical data for 20 country police bureaus in Taiwan from 2003-2004 and produced results that can help police in criminal incidents prevention. They combined visualization and clustering techniques. But, the limitation of the proposed model is difficult to evaluate the accurate performance. Ku et al. (2012) developed crime report visualization-Textual Analysis of Similar Crimes (TASC) using combination of visualization and clustering techniques to support crime analysts analyze, compare and contrast crime report in timely manner. However, the limitation of TASC is can only handle visualization of text highlighting for simple analytical tasks. Moreover, Palocsay et al. (2000) applied neural network to generate case-by-case results in criminal recidivism but their results are depending heavily on the choice of the network topology. Grubesic (2006) used fuzzy clustering for hot-spot detection but he did not mention specific architecture. Kaikhah and Doddameiti (2006) presented a novel knowledge discovery technique to discover trends using clustering technique for murder, rape, robbery and auto theft crime in United States cities. Furthermore, Yu et al. (2011) applied classification technique for burglary crime. They studied, compared and proposed the best forecasting approach to achieve the most stable results.

Based on Table 1, some of the data mining classes that have been applied in crime researches are visualization, clustering, prediction and classification. For data mining techniques, the researchers used techniques like neural network, fuzzy clustering and naive bayes. Several data mining classes such as visualization, prediction and clustering cover all crime types. However, in term of proposed architecture, some of the crime researches did not propose or mention specific architecture in their works. From the literature (Chen et al., 2004), visualization techniques have been applied in crime researches are visualization, clustering, classification technique for burglary crime. They studied, compared and proposed the best forecasting approach to achieve the most stable results.
demonstrated an enormous capacity to analyze crime data. Although visualization has the highest analysis capability on crime data, but there is still a lack of researches using this new technique. Because of this, current researchers are more focus on using visualization technique in their researches and studies including researches related to visualization of crime data. Unfortunately, the research in crime visualization is still limited in Malaysia. Thus, this research is conducted to overcome this gap and make a valuable contribution in analyzing crime data and thus support decision making for crime prevention.

Based on the issues that have been discussed as well as our previous crime researches (Noor et al., 2011a-c), we propose architecture of DSS using visualization technique for all crime types. In this study, we choose visualization technique because it has been proven can represent crime data into more understandable presentation. The aim of this research is to discover crime visualization, which will help the security authorities to have a better understanding of crime data. Then, they can use this information to directly support decision making for crime prevention. Therefore, to understand the architecture of DSS for crime visualization, the rest of the study is structured as follows. Section 2 provides information on materials and methods that include information about experimental design and the proposed architecture for crime visualization. Section 3 shows the experimental result and discussion. Lastly, section 4 gives conclusions and future works.

MATERIALS AND METHODS

Data acquisition: This study uses monthly crime volume for index crimes provided by the Royal Malaysia Police (RMP) from year 2008-2009 for Selangor state. Index crime in Malaysia refers to fourteen types of crime that occur with sufficient regularity and significance that they collectively serve as a meaningful index to the overall crime situation (PMDU., 2010). All index crimes are collected: Murder, rape, fire arm gang robbery, gang robbery without fire arm, fire arm robbery, robbery without fire arm, injuring others, stealing, stealing of truck/van, stealing of car, stealing of bike, snatch theft, house breaking and theft by day, house breaking and theft by night. Index crimes are grouped into two main crime categories, which are property crime and violent crime as depicted in Table 2. Crime volume is a basic indicator of the frequency of known criminal activity and it represents the number of reported offenses. Therefore, 336 temporal crime data is constructed, formed by 14 offenses in Selangor state for 24 months.

Data preprocessing: Data preprocessing includes data cleaning, data integration, data transformation and data reduction. In this study, data transformation is used to convert monthly crime volume into appropriate forms for mining. Due to variability existing in crime types, each of the monthly crime volume is standardized into a value between 0 and 1 using min-max normalization in Eq. 1:

\[
\text{Value} = \frac{X - X_{\text{min}}}{X_{\text{max}} - X_{\text{min}}}
\]

where, X is actual value, Xmin is minimum value of crime volume X and Xmax is maximum value of crime volume X in a crime type.

Parameter setting: At the mining phase, the Self Organizing Map (SOM) network is trained to cluster the standardize crime data. Before training, map size needs to be decided in advance. Specification of map size (number of output neurons) in the SOM training process is very much important to identify the deviation of the data. If the map size is too small, the outcome is more towards general patterns and it may not reveal some significant differences that should be detected. In contrast, larger map sizes will produce more detailed patterns where the differences are too small (Chattopadhyay et al., 2012). In addition, to measure the quality of a SOM’s projection, the Topographic Error (TE) can be used as an alternative way. The TE represents the proportion of all data vectors for which the first and second Best Matching Units (BMUs) are not closest to variability existing in crime types, each of the monthly crime volume is standardized into a value between 0 and 1 using min-max normalization in Eq. 1:

<table>
<thead>
<tr>
<th>Crime category</th>
<th>Crime type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Violent crime</td>
<td>Murder, rape, fire arm gang robbery, gang robbery without fire arm, fire arm robbery, robbery without fire arm, injuring others</td>
</tr>
<tr>
<td>Property crime</td>
<td>Stealing, stealing of truck/van, stealing of bike, stealing of car, snatch theft, house breaking and theft by day, house breaking and theft by night</td>
</tr>
</tbody>
</table>

Proposed architecture: The DSS is an information system that participates and supports the human decision making process. Generally, DSS can help decision makers make use of data and model and solve unstructured problems. The DSS makes full use of suitable computer techniques and through the interactive human-machine model, it helps and improves the effectiveness of decision-making about semi-structures and
non-structures (Zhou et al., 2008). The DSS is applied by human in many areas such as oil management, tourism, education and medical to improve the quality of decision making (Ramnarayan et al., 2004). For instance, a DSS model is used for crime prevention in Taiwan (Li et al., 2010). There are also many other researches related to DSS (Santana et al., 2012; Yang et al., 2012). Therefore, we are confident that DSS can be used to help the security authorities in Malaysia to make use of crime data and solve the problem related to the decision making for crime prevention.

This study uses hybrid DSS, a combination of model-driven DSS (MD-DSS) and web-based technology in handling decision making for crime prevention. The MD-DSS models decision problem using optimization and analytical tools and suggests actions. The MD-DSS uses data and parameters provided by decision maker to aid decision maker in analyzing a situation. This hybrid DSS can help retrieve, analyze and display structured data from relational database or large multidimensional, offer access to system and help communication and decision making for multiple users (Power, 2000). In DSS, there are three fundamental components which are Model-Base Management System (MBMS), Database Management System (DBMS) and Dialog Generation and Management System (DGMS) as depicted in Fig. 1. The DBMS serves as a data bank for the DSS while MBMS transforms data from DBMS into information that is useful for decision making. The DGMS or user interface is used to improve the ability of the system user to utilize and benefit from DSS. For instance, monthly crime volume for each crime type that covered for two years was recorded by the security authorities. By using this data, DBMS can store, update, delete and generate report. In this study, we use MySQL as our DBMS. To transform raw crime data into information, MBMS will communicate with DBMS. All information and reports generated will be displayed using DGMS.

In our proposed architecture, we use SOM in MBMS to transform monthly crime data into crime visualization. The SOM has been one of the most unsupervised neural network models to solve a wide variety of problems in clustering, classification, visualization and modeling (Li et al., 2010). The SOM is a visualization and analysis tool for high dimensional data. The SOM algorithm is based on nonlinear projection mapping which reduce the dimensions of high-dimensional data to two-dimensional data (2D) by producing a topology map. Therefore, SOM can be used to detect, analyze and visualize crime data that contains many variables.

RESULTS AND DISCUSSION

In this section, we analyze and discuss the experimental results for the aspects of crime visualization.

Visualization of crime data: Visualization of crime data provides useful information for the security authorities because visualization itself refers to an easily understandable presentation of data and to methodology that converts complicated data characteristics into clear patterns to allow users to view the complex patterns or relationships uncovered in the data mining process (Ngai et al., 2011). This proposed architecture can plot several types of plots such as using color, barplot, lines and radar. Nevertheless, the results of crime visualization are depending on the crime data and the
Fig. 2(a-g): Observation using “color” plot for violent crime, (a) Murder, (b) Rape, (c) Fire arm gang robbery, (d) Gang robbery without fire arm, (e) Fire arm robbery, (f) Robbery without fire arm and (g) Injuring other.

Fig. 3(a-g): Observation using “color” plot for property crime, (a) Stealing, (b) Stealing of truck/van, (c) Stealing of car, (d) Stealing of bike (e) Snatch theft, (f) House breaking and theft by day and (g) House breaking and theft by night.

Parameter settings as mentioned in section 2. In this study, we will show the result for observation using “color” and “barplot” plots only. For both result of observations, there are 4 clusters appeared in each map result. Cluster 1 located at left bottom corner, cluster 2 at left upper corner, cluster 3 at right bottom corner and cluster 4 at right upper corner of the map.

Observation using “color” plot for violent and property crimes are illustrated in both Fig. 2 and 3. By using this plot, it visualizes the result based on list of input parameters (crime types). Red (dark) color represents high level value while the yellow (light) color represents the low level value. Based on Fig. 2, rape and injuring others have shown the...
Cluster 2 Cluster 4(a)
Cluster 1 Cluster 3

Fig. 4(a-d): Observation using “barplot” plot for all crime types, (a) Cluster 2, (b) Cluster 4, (c) Cluster 1 and (d) Cluster 3

Table 3: Result for cluster 1 in Fig. 4

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Crime type</th>
<th>Level of crime value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>Murder</td>
<td>High</td>
</tr>
<tr>
<td>B</td>
<td>Rape</td>
<td>High</td>
</tr>
<tr>
<td>C</td>
<td>Fire arm gang robbery</td>
<td>High</td>
</tr>
<tr>
<td>D</td>
<td>Gang robbery without fire arm</td>
<td>Low</td>
</tr>
<tr>
<td>E</td>
<td>Fire arm robbery</td>
<td>Low</td>
</tr>
<tr>
<td>F</td>
<td>Robbery without fire arm</td>
<td>High</td>
</tr>
<tr>
<td>G</td>
<td>Injuring other</td>
<td>Medium</td>
</tr>
<tr>
<td>H</td>
<td>Stealing</td>
<td>High</td>
</tr>
<tr>
<td>I</td>
<td>Stealing of truck/van</td>
<td>High</td>
</tr>
<tr>
<td>J</td>
<td>Stealing of car</td>
<td>Very high</td>
</tr>
<tr>
<td>K</td>
<td>Stealing of bike</td>
<td>High</td>
</tr>
<tr>
<td>L</td>
<td>Snatch theft</td>
<td>Low</td>
</tr>
<tr>
<td>M</td>
<td>House breaking and theft by day</td>
<td>Low</td>
</tr>
<tr>
<td>N</td>
<td>House breaking and theft by night</td>
<td>Very low</td>
</tr>
</tbody>
</table>

highest crime value compared to other crime types in category of violent crime because three from four clusters displayed red color. Murder and robbery without fire arm show medium crime value while fire arm gang robbery, gang robbery without fire arm and fire arm robbery have shown small crime value. According to Fig. 3, stealing and house breaking and theft by night from category of property crime have shown the highest crime value. Stealing of truck/van, stealing of car, stealing of bike and snatch theft display medium crime value while house breaking and theft by day shows small crime value.

Next, we move to the observation using “barplot” plot for all crime types as depicted in Fig. 4. To understand the result, let us analyze the result for cluster 1 precisely. Table 3 contains detailed result for cluster 1 in Fig. 4. The list of A-N refers to the list of input parameters. For example, “A” refers to murder while “N” refers to house breaking and theft by night. In cluster 1, “J” (stealing of car) shows the highest crime value while “N” (house breaking and theft by night) shows the lowest crime value. Additionally, according to Fig. 4, it illustrates that most of the low crime values are assigned to cluster 2 and most of the high crime values are assigned to cluster 3. However, through observation using “barplot” plot, it can only represent general crime result in each cluster. To make a detailed conclusion on crime results, we need to analyze all four clusters and refer to observation based on input parameter such as by using “color” plot.

In this study, we agree with the capability of SOM as a visualization technique that can demonstrate a great capacity to analyze crime data (Chen et al., 2004). According to the analysis, we can discover four crime types that displayed the highest crime values compared to other crime types in Selangor state. The four crime types are rape, injuring others, stealing and house breaking and theft by night. In addition, based on Fig. 2 and 3, property crime shows higher crime value compared to violent crime. The results obtained are in support of other previous researches and surveys (PMDU., 2010). Sources from the RMP show that property crime makes up most of the overall index crime. Furthermore, although this study is aim to support decision making processes for crime prevention similar to other previous researches and studies, but it produces different way of results. For example, Li et al. (2010) used similar monthly crime volume but the result produced different crime trend patterns in Taiwan. Ku et al. (2012) applied visualization but he analyzed crime text report to facilitate analysis and decision making process. Kaikhah and Doddameti (2006) applied knowledge discovery technique but only focused for murder, rape, robbery and auto theft crime in US.

Therefore, the information we provide is useful for the security authority to determine level of crime for each crime type. The security authority can make a better use of its duty deployment, establish more satisfactory law enforcement policies and focus on the index crimes that affect the Malaysia society.

Strategy making of security authority duty deployment:
Based on the crime visualization result, it can support the security authorities to decide the kind of duty deployment that should be applied for each crime types. In general, there are two kinds of strategies of security authority duty deployment (Li et al., 2010). The first strategy is deploys the routine duties and the second strategy consists of special duties conducted by the security authorities. Based on our results, index crimes like rape, injuring others and stealing show high crime value in Selangor state. According to Sham et al. (2012), there is a high number of women who will experience greater fear while
travelling if public transport supply are not allocated accordingly to secure their journey. This may increase opportunity of crime towards women like rape, injuring others and stealing. In addition, house breaking and theft by night is also display high crime value in Selangor state. This is a reason for the government to take special emphasis for preventing house break-in in Government Transformation Program (GTP) (PMDU., 2013). The government took initiatives like residential police patrols, dedicated house break-in teams, coordinating volunteers through community policing and Crime Prevention Through Environmental Design (CPTED) to improve home security. For the crime types that fall under low and medium levels, the security authority can implement routine duties to prevent the crimes. Furthermore, to increase the safety perception index, several initiatives have been introduced such as Rakan Cop, whitening blackspot initiative, crime awareness day, safety perception survey and women awareness campaign.

CONCLUSION

In this study, architecture of DSS for crime visualization was described. The key contribution of this study is the proposed architecture is not only can be used for crime visualization but it could be valuable to other visualization apart from the crime data. Main feature of the architecture is the ability to visualize crime data by using several types of plots. The advantages of using this proposed architecture are it is a powerful exercise and also provide fastest way to communicate the results to others because it is a web-based application that used visualization technique. Therefore, the proposed architecture can support the security authority in assessing more suitable law enforcement strategies, increasing the accuracy of selection decision making, as well as improving the use of security authority duty deployment for crime prevention.

One limitation of the current study is that, it has computation time issue in handling huge crime data. Therefore, future researches could be aimed to find a better way to handle huge crime data and thus support the process of decision making for crime prevention. In addition, due to the availability of the crime data, the research was conducted using monthly crime volume that covered a period of two years for a specific state in Malaysia. If the period under consideration could be increased, it would provide a better result. In the future, we intend to test the architecture using monthly crime data that covered a period of five years for all states in Malaysia. Hopefully, this study will support one of the Malaysian government agendas known as NKRA, in which to reduce crime in Malaysia.

ACKNOWLEDGMENTS

This study is supported by Fundamental Research Grant Scheme (FRGS) under Ministry of Education (MOE), Malaysia with grant number 59289 and it is gratefully appreciated. Special thanks are also accorded to Criminal Investigation Department from the Royal Malaysia Police for the support and cooperation in providing crime data.

REFERENCES


