Additional Feet-on-the-Street Deployment Method for Indexed Crime Prevention Initiative

(Kaedah Tambahan untuk Pengerahan Rondaan Kaki bagi Inisiatif Pencegahan Jenayah Berindeks)

Mohammed Ariff Abdullah
Siti Norul Huda Sheikh Abdullah
Md Jan Nordin
(Faculty of Information Science and Technology, Universiti Kebangsaan Malaysia)

ABSTRACT

Under the National Key Result Area (NKRA) Safe City Program's (SCP) Safe City Monitoring System (SCMS) initiative, the Royal Malaysian Police (RMP) manages the deployment of feet-on-the-street via the indexed crime hotspots. Working on an approach known as the Repeat Location Finder (RLF), the RMP determines the displacement of indexed crime on the hotspots and may deploy feet-on-the-streets at the identified displacement areas as crime prevention measures. This paper introduces another deployment capability by shifting the focus from the hotspots to the identified serial suspects. Displacement models work on the concentration of crime incidents and the propensity location where the concentration might shift to the surrounding immediate hotspots. This additional method on the other hand, works on the identified suspects and identifies the next location where the suspects might surface, which may take place beyond the distance and boundaries of the hotspots. The objective of this paper is to identify the spatial features that positively contribute towards this new method. The solutions to the objective have been tested on a dataset made available by the RMP comprising 74 serial criminal suspects around the areas of Selangor, Kuala Lumpur and Putrajaya, spanning from Jan 1st to Dec 31st 2013. The identification capability moves as high as 92.86%. The RMP has been presented with the results of this paper and it was concluded that this method may be applicable as another capability in managing the deployment of feet-on-the-street resources.

Keywords: Safe city monitoring system; indexed crime; repeat location finder; hotspots; displacement; deployment.

ABSTRAK

Melalui Sistem Pemantauan Bandar Selamat (SCMS) di bawah Program Bandar Selamat (SCP) Bidang Keberhasilan Utama Negara (NKRA), Polis DiRaja Malaysia (RMP) menguruskan pengerahan rondaan kaki berpandukan tempat-tempat jenayah berindeks yang berketumpatan tinggi. Melalui pendekatan yang dikenali sebagai Pencarian Lokasi Berulang (RLF), RMP menentukan anjakan jenayah berindeks dari tempat-tempat ketumpatan jenayah tinggi dan mengerahkan rondaan kaki ke kawasan-kawasan anjakan yang dikenalpasti untuk tujuan pencegahan jenayah. Makalah ini memperkenalkan satu lagi keupayaan pengerahan dengan memindahkan tumpuan daripada kawasan berketumpatan jenayah berindeks tinggi kepada suspek bersiri yang dikenalpasti. Model RLF mencari lokasi ketumpatan insiden jenayah berindeks dan melakarkan lokasi kecenderungan ketumpatan jenayah beranjak mengelilingi lokasi terdekat ketumpatan asal, manakala kaedah tambahan ini memberi tumpuan kepada suspek yang dikenalpasti dan mengenalpasti anjakan lokasi seterusnya di mana suspek mungkin muncul, di mana lokasinya mungkin berada di luar jarak dan sempadan lokasi ketumpatan. Objektif makalah ini adalah untuk mengenalpasti fitur spasial yang menyumbang kepada kaedah baharu ini. Penyelesaian kepada objektif ini telah diuji ke atas set data jenayah berindeks yang disediakan oleh RMP yang terdiri daripada 74 suspek jenayah indeks bersiri di sekitar kawasan Selangor, Kuala Lumpur dan Putrajaya bermula dari 1 Jan hingga 31 Dis 2013. Puncak kemampuan mengenalpasti lokasi anjakan adalah setinggi 92.86%. Hasil kertas kerja ini telah dibentangkan kepada RMP dan kesimpulan dibuat bahawa kaedah ini boleh digunapakai sebagai keupayaan tambahan dalam menguruskan sumber pengerahan rondaan kaki.

Kata kunci: Sistem pemantauan bandar selamat; jenayah indeks; pencarian lokasi berulang; lokasi panas; anjakan; pengerahan.

INTRODUCTION

The Department of Federal Town & Country Planning Peninsular Malaysia (JPBD) in collaboration with the Royal Malaysian Police (RMP) have launched the Safe City Monitoring System (SCMS) in 2007 (Valentine 2012). As an expert in geo-spatial applications and spatial data acquisition, JPBD provides the platform for geo-spatial data and analysis capabilities while the RMP provides the crime data for indexed crimes. Through automation, the indexed crime data from a central national repository system at the RMP known as the Police Report System (PRS) (Goh 2003) is mapped into the SCMS as indexed crime hot spots, hence enabling the SCMS to churn out

maps visualizing the spread of indexed crimes. Through the indexed crime hot spots, the RMP manages the deployment of their feet-on-the-street manpower as part of their crime prevention initiatives. To predict a repeat-location of the hot spots, a buffer distance of 400m from the edge of the hot spots is created. The movements of the indexed crimes are predicted to take place within the boundaries of the buffer areas and hence the name Repeat Location Finder (RLF) (Valentine 2012). The RMP uses the RLF as a preventive tool to curb indexed crimes by deploying their feet-on-the-street within the areas identified by the RLF (Junoh 2014). SCMS is now implemented at every police station throughout Malaysia (Kasim 2017).

LITERATURE REVIEW

The author looks for early existence of pattern recognition involving the practical use of maps. The author also looks for previous works that associate behavioral inclination with spatial spaces. The author believes that establishing location-based behavior is important as it might explain why a concentration of crime happens at specific places.

Pattern recognition that uses mapping capabilities was first made successful by Dr John Snow back in 1849. Dr Snow was given the daunting task of addressing the cholera outbreak in which he discovered that the outbreak cases were cluttered near to each other: it was here that spatial density was first discovered. Figure 1 shows how Dr Snow used spatial density to correctly identify the location of the outbreak cases. Via analogy, Dr Snow proceeded with the identification of the source of the outbreak – the water pump at Broad Street. Dr Snow then ordered for the pumps in the area to be removed and the cholera outbreak was successfully curbed (Snow 1854). No one knew for sure back in 1850 how the outbreak

was successfully curbed by removing the water pumps, but a spatial association was certainly identified. It is significant to note that by plotting the cholera cases on a map, Dr Snow had managed to perform the early versions of pattern recognition involving spatial analysis.

As for the law enforcement domain, the author is interested in discovering whether geographical features have any influence on crime, whether a given location can be linked to a peculiar type of crime, and whether the same location may encourage, discourage or influence the behavior of the criminals. In a study by Willis (2009), it was found that crimes are indigenous and that geographical location affects the behaviour of the criminals. This means that certain types of crimes only happen at certain types of places and do not happen or are less likely to happen elsewhere. This supports the author's theory that the spatial properties of a location tend to attract certain types of suspects to commit certain offences. Spatial types and their features could be the factor that attracts or repels the suspects in choosing a certain location.

Since crimes are indigenous, the author tries to implicate whether the economic standing of the dwellers has any bearing towards the occurrence and concentration of crime. Is it possible that a neighborhood at lower cost housing might experience higher crime rates as compared to a neighborhood at higher cost housing? A landmark study was carried out and published by the Kuala Lumpur Royal Malaysia Police College on tried and closed-for-filing criminal cases which spanned over the period of seven years. The study summarizes that the suspects from different demographic backgrounds tend to behave differently and can therefore be clustered according to their demographic profiles. The statistics from the study indicate that the crimes are more likely to be committed by people from the lower economic class (Goh 2004). Hence, it can be assumed that if a location is clustered by lower

FIGURE 1. Clusters of cholera outbreak as mapped by Dr Snow. (a) Cases were marked and the hot spot area was identified within the perforated line. (b) Zoomed-in view of the area where the water pump was identified as the common denominator and was acknowledged as the cause of the outbreak (MATRIX 2017).

economic class inhabitants, then there is a probability that the cluster will have a higher concentration of suspects. This is because the housing projects in Malaysia are classified as either low-cost or high-cost, which means that the housing areas of the rich are separated from that of the poor. Therefore, the low-cost housing areas will probably have a higher concentration of crimes.

Besides the economic status of the inhabitants, the concentration of crimes is also speculated to be linked to other factors. For example, is it possible for a petrol station to become a hotspot for crime as the criminals are likely to grab a drink and be energized before committing an offence? Or is it possible for highways to become a hotspot for crime since the criminals may need a quick getaway route? It has been established in another study that a set of spatial temporal patterns may be associated with crimes. Yu (2016), for instance, mentioned that if there is a pattern of robberies occurring at a place near the banks, then the location of the banks in relation to the place of the robberies can be considered as a spatial pattern. It is thus speculated that if there are 10 spatial features at Site A and the same spatial features also exist at Site B, then there might be an underlying pattern linking the spatial features at both sites.

Another similar study on crime location suggests that crime spaces are not linear to the urban areas and may as well occur outside of any known hotspots. Tayebi (2015) mentioned that for every criminal, there exists a workspace. This workspace is a location familiar to the particular criminal and it is highly unlikely that a criminal shall venture outside of his workspace. Also known as activity space, a workspace is established to have its own spatial behavior. The author concludes that crimes do not necessarily move in an orderly finite distance and may move in any direction from its point of origin.

Visualizing crimes via the electronic media shall be beneficial for the community and the author discovers that plotting the crimes onto maps has been a practice in recent days. The Chicago Police Department, for example, makes available an interactive online map which displays crime data and the location of the police stations. Dubbed as CLEARmap, which stands for Citizen and Law Enforcement Analysis and Reporting Map Application, the purpose of this data sharing is to enhance public safety (City of Chicago Police Department 2017). With CLEARmap, the public may have a visual presentation of the crimes on a city map and therefore may take precautions when entering areas with high crime rates. By empowering the public with knowledge of the crime-prone areas, the City of Chicago Police Department expects to promote better awareness towards crime prevention. However, with respect to the Malaysian crime-mapping initiative, SCMS is made available only to the RMP and local city councils.

Venturing into the mathematical model of crime prediction, the Rossmo model (Rossmo 1995) sets out to find the location where the next crime is likely to occur. Given a known starting point, the Rossmo model would calculate the distance or buffer zone from the starting point. It is within this buffer zone, which surrounds the starting point at 360 degrees, that the next crime is most likely to occur (Zheng 2011). The Rossmo model works by focusing on the displacement of crime. In other words, if a crime happens at spot A, then similar crimes might take place at spot B within a predetermined distance not far away from its point of origin. In addition, spot B may be situated at any direction or bearing from spot A.

The use of some complex formula is also stipulated as a probable method in performing predictions. Coined as the Hybrid Approach, it suggests the use of various contributing factors which can be utilized in considering a prediction (Azeez 2015). The contributing factors are visual analytics approach, social sentiments via the likes of Twitter and other forms of social media, and statistical technique. The output for this method is the Graph Database which leads to the ability to predict the most probable crime which may happen at a given time and space frame. The hybrid approach focuses on the hotspots, contrary to this paper which focuses on the movement of the suspects.

In another recent work, Bohani (2018) came up with a probability table which shows the likelihood of thievery and robbery happening at selected spatial features such as residential areas, banks and eateries. This method however neither tracks the movements of a serial suspect nor the movements of crime hot spots as it only stipulates the probabilities of the crimes taking place at the locality of the identified spatial features.

With the boom of the internet and the usage of mobile devices, crowd sourcing might be another tool for crime-mapping (Malleson 2015). There are two methods in determining locations which may have high concentrations of crime rate, namely the residential population and the ambient population. The author feels that a higher population density might likely produce a higher crime rate. While residential population works on the density of the actual population, ambient population on the other hand, works on the number of social media tweets sent out from the same area of the population. Though it can be concluded that a high residential population density and a high ambient population density may have some spatial patterns related to areas with high crime rates, the opposite cannot be said the same. It seems that high crime areas on the other hand, may neither display the existence of high residential nor ambient populations. Hence, no conclusion is reached.

From the previous works, the author concludes that some form of prior knowledge is needed when addressing crimes. For example, Goh (2004), Azeez (2015) and Malleson (2015) mention about social demography while the City of Chicago Police Department (2017) and Safe City Monitoring System (Valentine 2012; Junoh, 2014) work on crime locations and crime hot spots. None of these works mentioned about using previous knowledge to apprehend the specific suspects of wrong doings as they only focus on the general movement of hot spots. Rossmo

(1995) on the other hand uses the crime displacement method to curb crimes by predicting crime movements. However, displacement models fail because as mentioned by Zheng (2011) and Tayebi (2015), the movements of crimes are not linear and they move in any direction outside of any known hot spots.

Taking cues from Willis (2009) and Yu (2016), the author is inspired to learn more about spatial association. The author explores the association of distance between spatial features and crime locations. Spatial distances are radials from their points of origin and would create multiple and overlapping areas when we work with multiple spatial features. Hence, there is a need to keep the areas of interest and to omit the areas that are not associated with any crime locations. The author believes that the Zaermelo-Fraenkel set theory is best suited because it is able to identify areas of junctions and disjunctions (Lian 2011) and by coupling with pattern recognition via neural network, we are able to predict the next areas of junction.

It is the objective of this paper to move away from the paradigm of working with prior knowledge and hotspots as currently being deployed in SCMS via RLF, and to propose a new dimension by utilizing coupling models that are capable of generating the next indexed crime location of a suspected serial criminal.

This research introduces the creation of workspace and the significance of set theory coupled with pattern recognition in identifying the spatial features that give positive influence to the direction of this new method. This new method adds a new dimension to the existing management of the feet-on-the street deployment method to combat indexed-crime by moving beyond the confines of RLF.

METHODOLOGY

The following are the scopes of this research:

i.   Indexed crime categories: a) Robbery and b) Burglary & Theft.
ii. Indexed crime specimens: Jan 1st to Dec 31st, 2013.
iii. Geographical coverage: State of Selangor and the Federal Territories of Putrajaya & Kuala Lumpur.

The spatial data used in this research was obtained from JPBD, while the geo-referenced crime data was obtained from RMP. The spatial data features consist of Point of Interest (POI), Polygons and Polylines, while the crime data feature is made up of Point of Interest only. The entire dataset for this research is made up of different features and attributes. The data from JPBD consists of 33 spatial features while the data from RMP consists of 74 indexed crime suspects with confirmed arrests.

Figure 2 describes the elements used in representing the types of data in this research:

i.   **Polygons** as shown in Figure 2(a) are used to represent administrative boundaries. Polygons mark State lines, Police Contingents and Local Governments.

ii. **Point-of-Interest** (POI) as shown in Figure 2(b) and (c) are points which are used to mark the location of spatial features by JPBD while RMP uses POI to identify the location of crimes. POI data type consists of Bank establishments, Bus and Taxi stops, Cemeteries, Colleges and Universities, Commercial sites, Embassies, Fire & Rescue departments, Food eateries, Golf courses, Government establishments, Hospital & Medical facilities, Hotels, Libraries, Markets, Museums, Parks, Petrol Stations, Places of interests, Places of Worship, Police establishments, Ports & Jetties, Postal offices, Rail & LRT lines, Residential areas, Schools, Shopping Centers, Sports & Fitness facilities and Toll plazas.

iii. **Polylines** as shown in Figure 2(d) are used to represent road networks. Polylines draw Transit Roads, State Roads, Municipal Roads, Federal Roads, Expressways and Highways.

Figure 3 outlines the research framework.

i.   Prepare data by translating data from map format to neural network format
     a.   For all suspects, determine all of Location 1.
     b.   For all suspects, determine all of Location 2.
     c.   For all Location 1 for all suspects, calculate the nearest distance to every spatial feature.
     d.   For all Location 2 for all suspects, also calculate the nearest distance to every spatial feature.
ii. Create Workspace value
     a.   Determine a workspace value for workspace creation.
     b.   Add workspace value to all distances to all suspects in relation to Location 1.
     c.   Add workspace value to all distances to all suspects in relation to Location 2.
iii. Run network
     a.   Input layer:
          ●   Each node represents a feature where the number of nodes equals the number of features.
          ●   Each node carries the distance value plus the Workspace value of the feature from Location 1.
     b.   Output Layer:
          ●   Each node represents a feature where the number of nodes equals the number of features.
          ●   Each node carries the distance value plus the Workspace value from Location 2.
iv. Verify training
     a.   Use data which is not part of the training data for verification purpose.
     b.   Input to the network is Location 1 with the Workspace value matching the training set.
     c.   Output (t) to the network shall be converted into raster data.
v.   Run set theory and re-translate data from neural network format to map format

a. Distance values churned out by the network shall be fed into the GIS.

b. For all feature values, run Python script to churn out polygons
- For each feature from its center, create a radial distance with the distance value taken from the Output (*t*) values churned out and an area is created.
- For features of the same type, create union to have a single layer of the same feature (*u*).
- For features of different types, take all of the union layers (*u*) and perform intersection.
- Discard areas of disjunction.
- Remaining areas, if any, are the areas where the new crime Workspace are predicted.

Representation of the spatial data is carried out to enable machine learning (Nordin 2009; Ali 2009; Kahaki 2011; Ghazvini 2015). The relationship between crime data and spatial feature attributes is defined as the distance between the crime location of interest and the spatial features. GIS is used to calculate the distance, and the neural network learns these attributes from the last two known locations and predicts new attributes. The resultant attributes are then converted back to GIS and mapped to their respective spatial features. GIS is then used to identify the junctions and disjunctions of the spatial features. The operations of junctions and disjunctions are shown in Figure 4. By eliminating the disjunctions, features that do not contribute to the prediction are hence eliminated, leaving behind the predicted locations.

Figure 4 provides a simple illustration on the use of the Zaermelo-Fraenkel set theory. Figure 4(a) and Figure 4(d) depict a finite geo-fence which contains geo-features
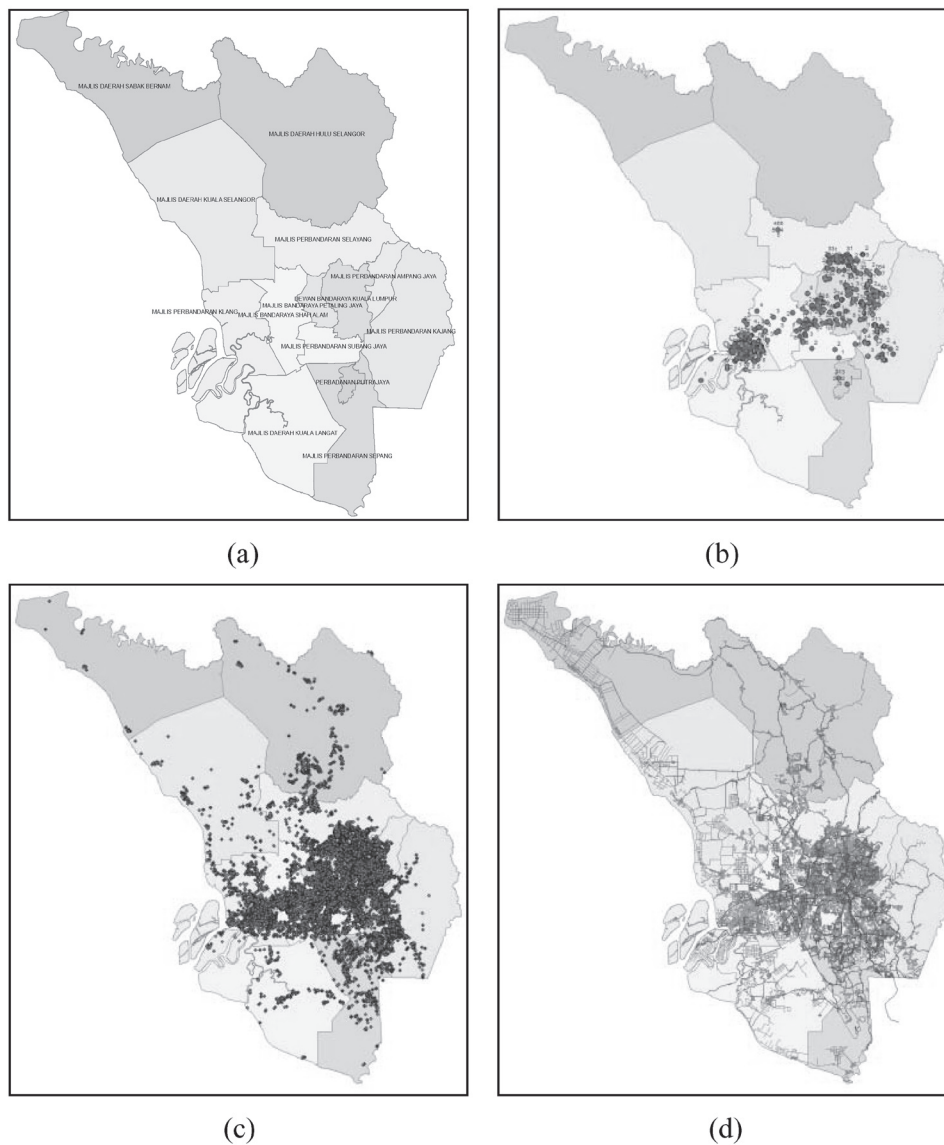

(a)

(b)

(c)

(d)

FIGURE 2. (a) Polygon data type showing the base map of the State of Selangor and the Federal Territories of Kuala Lumpur and Putrajaya. (b) Actual locations of crimes as committed by serial suspects are represented in the shape of points. (c) Point-of-Interest (POI) data type representing spatial features. (d) Poly-line data type representing the network of roads and land transportation.
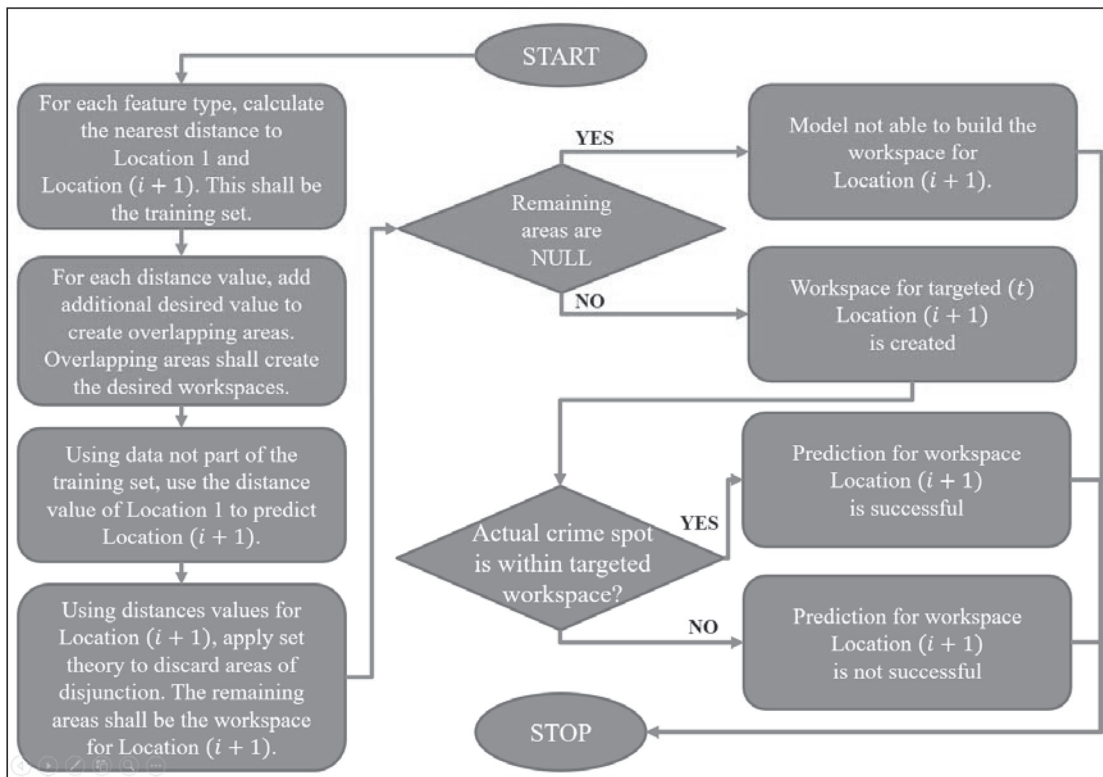
FIGURE 3. The Conceptual Framework showing the creation of Workspace by translating data from the map format into the Neural Network format, which is then re-translated into the map format which undergoes the Spatial Elimination Process via the Zermelo-Fraenkel Set Theory.

x_1, x_2 and x_3, where each geo-feature is different. The circles which each geo-feature creates represent the workspace distance of each geo-feature. Figure 4(b) and Figure 4(e) represent the intersection areas of x_1, x_2 and x_3. The area of intersection is known as the area of junction, while the non-intersecting area is called the area of disjunction. The area of junction is the predicted area of the location where the next crime shall be committed. Figure 4(c) and Figure 4(f) represent the union of areas x_1, x_2 and x_3.

When a suspect commits a crime at Location 1, similar spatial qualities at Location 1 might likely be applied at Location 2 (Junoh 2014; Mun 2015). Hence, after building the workspace for Location 1, the information is then used in building the workspace for Location 2.

Figure 5 provides a simple explanation for better understanding of the creation or definition of Workspace. Figure 5(a) is the location of the crime. From the
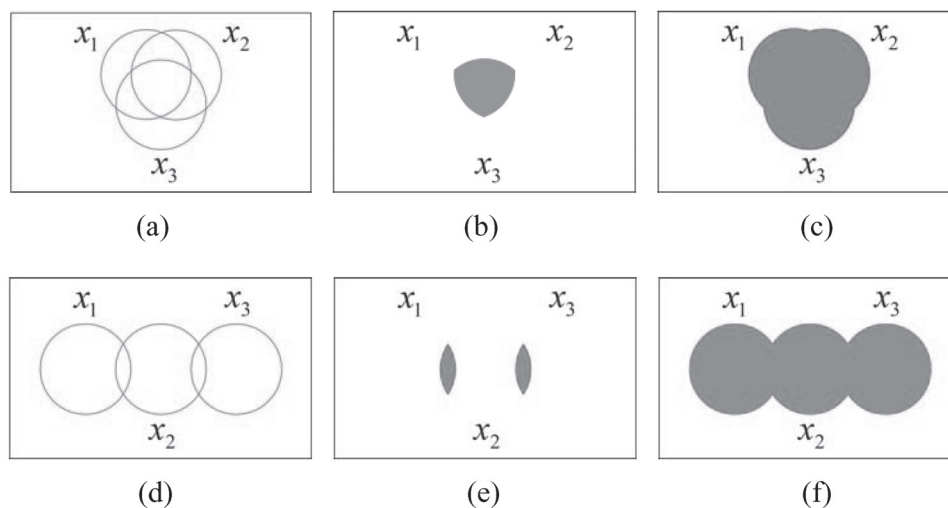

FIGURE 4. Samples of the set operations as laid out by Zaermelo-Fraenkel (a) and (d), x_1, x_2 and x_3 exist in a finite space, (b) x_1, x_2 and x_3 intersect each other ($x_1 \cap x_2 \cap x_3$), (c) and (f) union of x_1, x_2 and x_3 ($x_1 \cup x_2 \cup x_3$), (e) x_1 intersects with x_2 ($x_1 \cap x_2$) and x_2 intersects with x_3 intersect ($x_2 \cap x_3$).

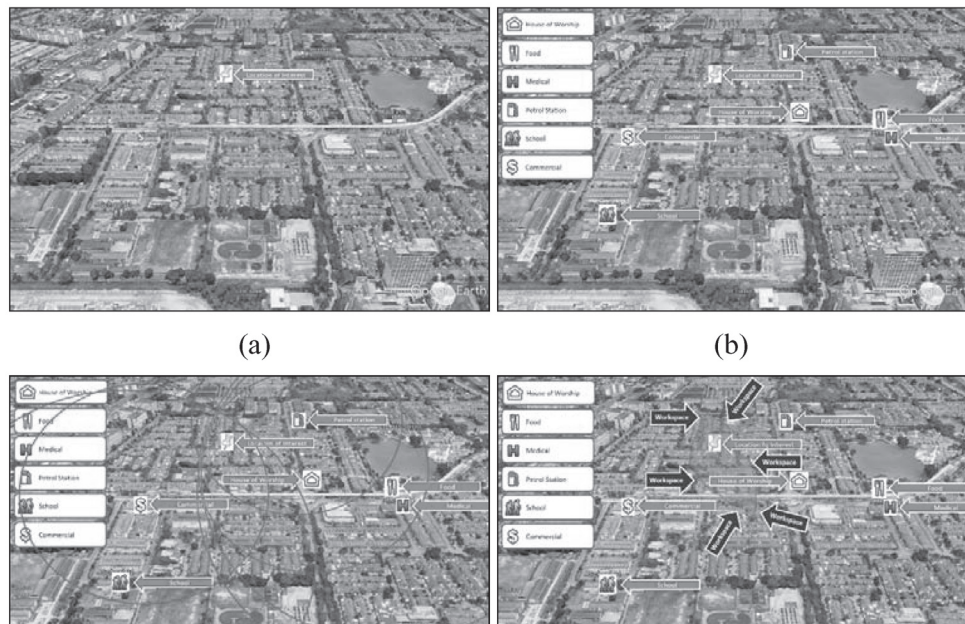(a)   (b)

FIGURE 5. Creating a Workspace (a) Location of crime. (b) Spatial features nearest to the location of crime. (c) A value is added to the distance of each spatial feature which is known as the Workspace Value. (d) Intersecting area is kept and is called the Workspace.

location of the crime, all spatial features surrounding the crime location would then be located and only the features nearest to the crime location would be retained. In the example in Figure 5(b), five features have been identified, namely the Places of Worship, Food eateries, Hospital & Medical facilities, Petrol Stations, School and Commercial sites. The radial distance from each feature is calculated and a distance value known as the Workspace Value is added, as shown in Figure 5(c). In Figure 5(d), intersection is performed and the remaining area is called the crime Workspace.

RESULTS

Two sets of experiments are executed:

i.   Set 1: Spatial Distance Elimination Method - DEM:
   - In this set of experiment, the author wishes to introduce the use of spatial features based on their association with the distance of the crime locations. This method calculates all crimes in Location 1 and the distance against all spatial features. Spatial features which have the closest distance to the crime scene in Location 1 are retained while spatial features with further distances are systematically pruned. By using this method, the author wishes to explore the possibility of justifying the use of spatial features based on how near they are to the crime scene – nearer distance makes more sense than further distance.
   - This set is made up of six runs. The first run uses all spatial features as described in Table 1, while the subsequent runs eliminate spatial

features based on their distance from the first crime location (Location 1) as provided by the training set. Table 2 (to be read together with Table 1) shows the spatial features which are pruned at each run. Run 2 has spatial features with the longest distance pruned, and progresses towards Run 6 which has spatial features with the shortest distance. The results of this set are shown in Figure 7.

ii. Set 2: Spatial Frequency Elimination Method - FEM:
   - For this set of experiments, the author wishes to introduce the use of spatial features based on their frequency of occurance. Spatial features which have higher frequency in relation to the first crime location are kept, while spatial features with lower frequency are systematically pruned. In this method, the author wishes to explore the relevance of spatial features based on their occurance – spatial features which appear more create higher relevance as compared to spatial features with lower appearance in the spatial space.
   - This set is also made up of six runs. The first run uses all spatial features as described in Table 1, while the subsequent runs eliminate spatial features based on their frequency or availability across the spatial space as described in the geographical coverage from the first crime location as provided by the training set. Table 3 (to be read together with Table 1) shows the spatial features which are pruned at each run. Run 2 has spatial features with the least frequency pruned, and progresses towards Run 6 which has among the highest spatial

TABLE 1. Complete listing of features and feature types used in this paper

Feature Type				
Point	01. Bank establishments	2. Bus and Taxi stops	03. Cemeteries	04. Colleges and Universities
	05. Commercial sites	6. Embassies	07. Fire & Rescue departments	08. Food eateries
	09. Golf courses	10. Government establishments	11. Hospital & Medical facilities	12. Hotels
	13. Libraries	14. Markets	15. Museums	16. Parks
	17. Petrol Stations	18. Places of Interest	19. Places of Worship	20. Police establishments
	21. Ports and Jetties	22. Postal offices	23. Rail & LRT lines	24. Residential areas
	25. Schools	26. Shopping Centers	27. Sports & Fitness facilities	28. Toll Plazas
Polylines	29. Transit Roads	30. State Roads	31. Local Roads	32. Federal Roads
	33. Highways			
Polygons	34. Selangor State	35. Kuala Lumpur	36. Putrajaya	

TABLE 2. Pruned features for predicting location via distance elimination method

Run #	Pruned Spatial Features			
1	No pruning			
2	1. Embassies	02. Ports and Jetties		
3	1. Embassies	02. Ports and Jetties	03. Museums	04. Libraries
4	1. Embassies	02. Ports and Jetties	03. Museums	04. Libraries
	5. Bus and Taxi stops	06. Golf courses	07. Toll Plazas	08. Places of Interest
	9. Parks			
5	1. Embassies	02. Ports and Jetties	03. Museums	04. Libraries
	5. Bus and Taxi stops	06. Golf courses	07. Toll Plazas	08. Places of Interest
	9. Parks	10. Rail & LRT lines	11. Fire & Rescue departments	
6	1. Embassies	02. Ports and Jetties	03. Museums	04. Libraries
	5. Bus and Taxi stops	06. Golf courses	07. Toll Plazas	08. Places of Interest
	9. Parks	10. Rail & LRT lines	11. Fire & Rescue departments	12. Postal offices
	13. Highways	14. Police establishments	15. Hotels	16. Shopping Centers
	17. Hospitals & Medical facilities	18. Colleges & Universities	19. Sports & Fitness facilities	20. State Roads
	21. Government establishments	22. Transit Roads	23. Cemeteries	24. Markets

TABLE 3. Pruned features for predicting location via frequency elimination method

Run #	Pruned Spatial Features			
1	No pruning			
2	01. Embassies	02. Ports & Jetties		
3	01. Embassies	02. Ports & Jetties	03. Federal Roads	04. Highways
	5. State Roads	6. Transit Roads		
4	01. Embassies	02. Ports & Jetties	03. Federal Roads	04. Highways
	05. State Roads	06. Transit Roads	07. Markets	08. Toll Plazas
	09. Golf courses	10. Fire & Rescue departments	11. Parks	12. Bus & Taxi stops
	13. Libraries	14. Museums		
5	01. Embassies	02. Ports & Jetties	03. Federal Roads	04. Highways
	05. State Roads	06. Transit Roads	07. Markets	08. Toll Plazas
	09. Golf courses	10. Fire & Rescue departments	11. Parks	12. Bus and Taxi stops
	13. Libraries	14. Museums	15. Sports & Fitness facilities	16. Cemeteries
	17. Places of Interest	18. Rail & LRT lines		
6	01. Embassies	02. Ports & Jetties	03. Federal Roads	04. Highways
	05. State Roads	06. Transit Roads	07. Markets	08. Toll Plazas
	09. Golf courses	10. Fire & Rescue departments	11. Parks	12. Bus Taxi
	13. Libraries	14. Museums	15. Sports & Fitness facilities	16. Cemeteries
	17. Places of Interest	18. Rail & LRT lines	19. Shopping Centers	20. Hospitals & Medical facilities
	21. Police establishments	22. Postal offices		

(a)

(b)

(c)

FIGURE 6. Sample of prediction results where "1" denotes the first crime location, "2" denotes the second crime location and the striped areas denote the prediction areas. (a) All features used (b) Distance Elimination Method (c) Frequency Elimination Method
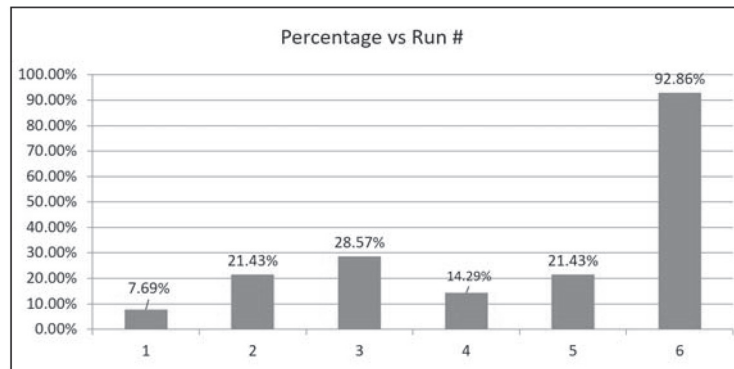

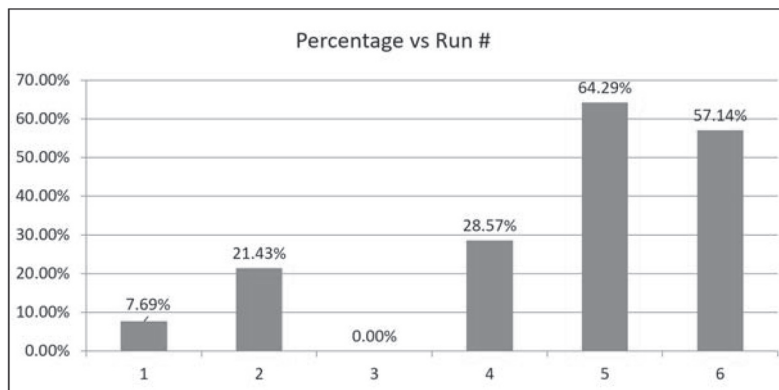FIGURE 7. Success rate of location prediction via the distance elimination method

FIGURE 8. Success rate of location prediction via the frequency elimination method

frequencies. The results of this set are shown in Figure 8.

The results of Set 1 and Set 2 are shown in Figure 7 and Figure 8 respectively. The y-axis represents the percentage of successful predictions, while the x-axis represents the discarded spatial features. Both tables show an upward trend towards pruning, and the predictions have indeed improved when the spatial features closer to the crime location are selected (Set 1) and when the spatial features of greater availability are used (Set 2).

The prediction samples are shown in Figure 6 involving three separate suspects. The goal of the experiments is to identify the areas where the next crime locations will be, as tagged to the serial suspects. The predicted areas are marked with stripes, and when the second location falls within the striped areas then the predictions are successful. Figure 6(a) to 6(c) also show that the predictions will fail if RLF is used since the movements of the suspects may be in any direction and may cover any distance, which is beyond the capability of the RLF.

DISCUSSION

This paper looks at the capabilities of the existing SCMS and explores to introduce a new dimension in the deployment of feet-on-the-street. The current method of deployment is by identifying the hotspots and creating a buffer zone surrounding the hotspots. Working on the displacement of the crime, the next hotspot location is said to be within the buffer zone, which is set at 400m from any hotspot. The exact location is however not known since the displacement may move in any direction from its point of origin.

Hence, the author wishes to further improve the capability of SCMS beyond the displacement of the hotspots by focusing on the movement of specific suspects. This is achieved by working on the capability of predicting the next location where the suspects may surface. In other words, given the first location that a suspect committed a crime, the author wishes to locate the next location in which the same suspect will commit another crime.

Run 1, as shown in Figure 7 and Figure 8, is a base run intended to show whether this new dimension is applicable as a new method. The new method of developing a Workspace by creating an empirical association based on the location seems to work. Based on the creation of the predicted workspace, this initial base run is able to create workspaces that coincide with 7.69% of the suspects. Once a base run is established, the author looks at the possibilities to further improve this new dimension by exploring new methods in selecting the spatial features.

After exploring the base run which validates the area of workspace creation, the next step is to explore and determine whether the distance of the spatial features from

Location 1 may create unique associations leading to the betterment of the base run – Distance Elimination Method (DEM). According to DEM, features that are nearer to all of Location 1 of all suspects are to be retained, while those that are further away will be systematically eliminated. Based on Figure 7, it can be seen that when more spatial features are eliminated, workspaces could be created in which their footprints contain Location 2. It seems difficult to establish an upward trend just by looking at the Run 1 to Run 5 experiments since the percentages oscillate between 7.69% and 28.57%. However, Run 6 being the last of them shows the possibility of an upward trend. In another method, the spatial features are eliminated via the Frequency Elimination Method (FEM) based on their frequency of occurrence over the finite geo fence of Selangor, Kuala Lumpur and Putrajaya. Working similarly to DEM, FEM also displays uncertainty as the percentages seem to oscillate between 0% and 28.57%. With this oscillation, it is hard to determine any upward trend. However, from Run 5 and Run 6 as shown in Figure 8, an upward trend is made possible.

Unlike RLF, workspace creation via FEM and DEM does not rely on the displacement method for two reasons: 1) displacement does not determine the next location since displacement may move in any direction from its point of origin; and 2) displacement works within a preset buffer value.

From the dataset received, the distance travelled from Location 1 to Location 2 can be seen. From Table 4, it is evident that the displacement method does not make the cut and this confirms what Zheng (2011) and Tayebi (2015) have said. From Location 1 to Location 2, the suspects seem to have travelled an average of 8.18 km. With a minimum distance of 0.24 km and a maximum distance of 39.937 km, the displacement method should be less than accurate as it has a preset distance of 400m.

TABLE 4. Straight line distance traversed by serial suspects

	Distance (km)
Min	00.245
Max	39.937
Avg	08.183

Based on additional scrutiny on the dataset as shown in Figure 9, only a small percentage of the suspects, which is 8.82%, have travelled the distance of less than 1 km while the remaining 91.18% have moved more than 1 km from the first location. Based on Table 4 and Figure 9, it is evident that the RLF method should fail if the objective is to track the movement of the serial suspects since they are highly unlikely to move within a 400 m buffer zone from their current crime location. Furthermore, serial suspects are very likely to traverse outside of or beyond the jurisdiction of the police station of their place of origin.
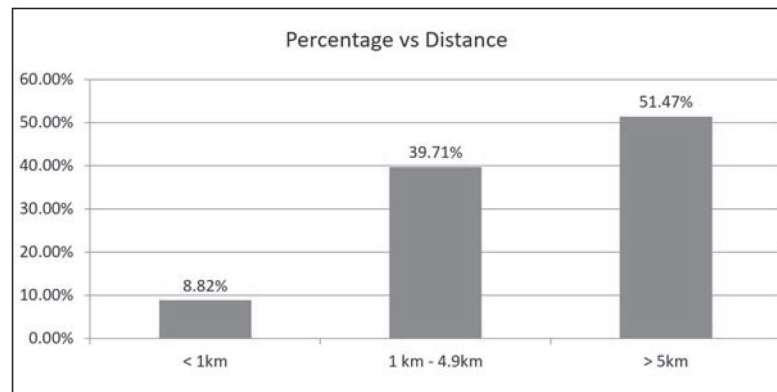
FIGURE 9. Breakdown of the percentages of serial suspects against the distances they travelled from Location 1 to Location 2

The RLF serves its purpose since it is a good tool in predicting the displacement of hotspots. However, RLF is not a heuristic tool and could not be used to track down serial suspects who had evidently traversed beyond the 400 m buffer set by the RLF. Thus, to increase the capability of SCMS, a heuristic method has been introduced. Through the process of workspace creation and later on with the use of the spatial elimination methods, it is determined that the attribute characteristics of any given feature can be measured; therefore, the features that are more likely and less likely to be associated or compelling for a suspect in selecting a crime location can be ascertained.

The ability to predict the suspects' next location will add a new dimension to the management of resources. Since RLF works by identifying hotspots and their displacement, it is thus unable to detect the movement of specific suspects and is also bound by the radial restrictions of the RLF model. This new method on the other hand, adds a new dimension of deployment by specifically pin-pointing the predicted area of crime well beyond the boundaries of the RLF, and hence may add to the capabilities of the RMP in deploying its resources.

In summary, the Safe City Monitoring System (SCMS) crime mapping facility is a shared facility concerted by two agencies that are in great cooperation – Federal Town & Country Planning Department (JPBD) and the Royal Malaysian Police (RMP). The role of JPBD is to make the SCMS platform as well as the spatial data available, while the role of the RMP is to plot the location of the crimes at source as they are being recorded by the attending officer.

i. The Crime Mapping: When multiple layers of data are arranged together as one visualization, crime mapping can be visualized as a map. From this map, one can visualize the locations and concentrations of the crimes.

ii. The Machine Learning: To make machine learning possible, raster information must be translated from the map format into the neural network friendly vector data. The vector data contains the utmost important dimension for this thesis which provides the coordinate of the locations. Since this thesis is about predicting locations, the coordinate feature is a must. Once the coordinates of the spatial features are obtained, then naturally the calculation of the distance shall be made possible.

iii. The Spatial Elimination Process: Even after the values of the predicted distance are obtained, they are basically useless since distance values are merely displacements from their points of origin. To make sense, the vector data must then be converted back into the map format. Via the set theory, spatial areas that are of importance are eliminated, leaving intact the exact locations of the prediction.

For future researchers who wish to take up from where the author has left off, there shall always be room for improvement. The concept of workspace is the concept of creating an area in which a suspect is said to be comfortable in undertaking his or her criminal activities. Based on Figure 6, it can be seen that the creation of workspaces can be of various sizes. The author therefore suggests for future researchers to look into enhancement or refinement methods to reduce the size and location of the areas of junction.

MANAGERIAL IMPLICATION

The Safe City Monitoring System (SCMS) under the Safe City Program (SCP) initiative is an excellent example to depict the use of crime visualization for the purpose of keeping the cities safe. Crime mapping is not possible without the collaboration between JPBD and RMP. SCMS revolutionizes the existing PRS and the way how crime incidents are recorded at the PRS and later on visualized for the purpose of crime mapping.

As shown in Figure 10, the location of a crime is plotted in accordance to the information given in real time i.e. as the incident is being reported and the location is being confirmed by the investigating officer. The following process is a string of sequence which describes how data integration between PRS and SCMS happens:

i. Criminal incident happens.
ii. The victim or his/her representative makes a police report.

iii. Information on the scene as well as information on the incident is recorded.
iv. If the nature of the crime is covered under SCMS, then the attending officer shall display a map. The victim or his/her representative is then required to identify the location of the crime.
v. The Investigating Officer (IO) is then asked to verify the validity of the report. Once verified, the identified location shall be recorded in SCMS.

By tagging the report of the incident to a GIS system, the data recorded from the report of the incident will be added to the existing spatial data. The electronic visualization of the crime maps or hotspots of a specific type of incident can be churned out easily and is of great benefit for law enforcement as it enhances the capability to visualize and analyze the spread of crime over a multitude of domains and variables. In addition, it enables the easy modification of the maps such as the change of scale, the addition or removal of map details, the addition or removal of incidents and also the cross reference with other types of incidents.

The SCMS is unique and has won the ESRI International Special Achievement in Geography Award 2012, with the following key business cases solved (ESRI 2012):

i. SCMS has managed to integrate with the Police Reporting System (PRS), and hence has enabled real-time mapping and updating of crimes at all police stations;
ii. SCMS has managed to resolve issues on data redundancy by enabling its users to perform real-time geo-processing analysis;
iii. SCMS has managed to enable its users to perform real-time analysis on crimes such as Hot Spots and Timeline.

SCP, as it is, has been deploying feet-on-the-street based on the data which supports the need for the deployment. This method of deployment is effective as it allows RMP to effectively manage the utilization and concentration of the use of its resources, personnel and assets in areas known to be infested with crimes, or also known as hotspots. Alongside RMP, JPBD may deploy its resources for target hardening. With the capabilities of RLF, the management of RMP and JPBD are better off at guessing the displacement of the crimes.

Presentations were then made to SCP and RMP. According to Kasim (2017) in his function as the Deputy Director of The Department of Crime Prevention & Community Safety of the RMP, the SCP has also improved the capabilities of police stations throughout Malaysia. At present, each police station has a good mapping tool which allows data entry at source, which in turn allows the Officer In-Charge of the Police Station (OCS) to identify the crime hotspots within the jurisdiction of the station. Each OCS now can better manage the deployment of police assets such as the Mobile Patrol Vehicle (MPV), Motorcycle Patrol Unit (URB) and of course, the feet-on-the-street. The RLF however, shall not be replaced by the new concept introduced in this thesis for these two reasons: 1) the objective of the RLF is to deal with hotspots, and 2) the objective of predicting the next workspace is to deal with the serial suspects. As established in Table 4 and as shown in Figure 9, the serial suspects may traverse beyond the boundaries out of the jurisdiction of any OCS. However, the State police and the Federal police shall be interested to investigate. Kasim (2017) further elaborates that the new concept of predicting the workspace may be further explored and may be brought forward into the domain of national policing. Having identified the location of a would-be crime spot, RMP may choose to deploy the appropriate amount of resources and the suitable policing methods by either making their presence visible or by choosing to be low profile by applying discrete policing methods such as observation and electronic surveillance. In fact, a specialized police division with pre-crime or foreknowledge capabilities may be explored further. According to Valentine (2018) in his function as the head of the Safe City Program, the Safe City initiative has won many local and international awards and is recognized as the world's first of its kind; adding the predictive capability shall propel SCP onto higher grounds. SCP has tremendously increased the efficiency of RMP, and the capability to predict the location of a suspect may add another dimension to policing and police efficiency.
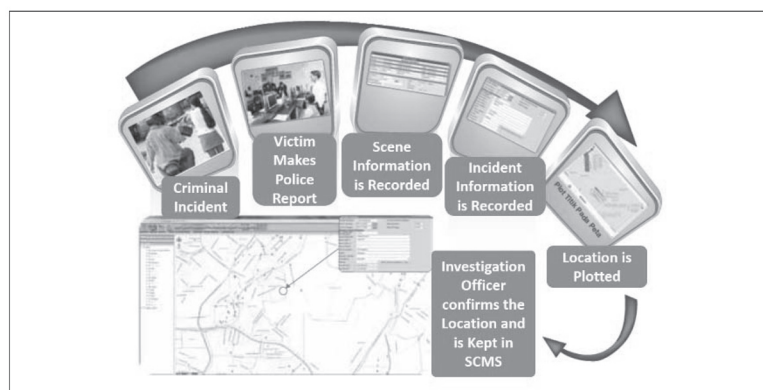

FIGURE 10. Crime mapping process

ACKNOWLEDGEMENT

REFERENCES

Ali, A.M. & Nordin, M.J. 2009. Vision based reconstruction multi-clouds of scale invariant feature transform. *Journal of Computer Science* 5(12): 951-958.

Azeez, J. & Aravindhar, D.J. 2015. *Hybrid approach to crime prediction using deep learning*. Paper presented at the international conference on Advances in Computing, Communications and Informatics, 10-13 August, Kerala, India.

Bohani, F.A. Abdullah, S.N.H.S. Zakree, A.N. Jeffry, Y. Abdullah, M.A. Junoh, M.N. & Kasim, Z.A. 2018. Amenities surrounding commercial serial crime prediction at greater valley and Kuala Lumpur using K-Means clustering. *Jurnal Teknologi (Sciences & Engineering)* 80: 43-53.

City of Chicago Police Department, 2017. CLEARmap. Available at http://home.chicagopolice.org/online-services/i-clear-application-for-law-enforcement/

ESRI. 2012. ESRI international user conference: 2012 SAG award winner. Available at http://events.esri.com/uc/2012/sag/list/?fa=Detail&SID=1506

Ghazvini, A. Abdullah, S.N.H.S. Abdullah, M.A. Junoh, M.N. & Kasim, Z.A. 2015. Distance prediction for commercial serial crime cases using time delay neural networks. *International Journal of Advanced Computer Science and Applications* 11: 1-11.

Goh, B.K. 2003. First information report. *Journal of the Kuala Lumpur Royal Malaysia Police College* 1: 57-74.

Goh, B.K. 2004. Confession, denial or silence? A case study on the exercise of right to silence by accused persons during police investigation. *Journal of the Kuala Lumpur Royal Malaysia Police College* 2: 17-31.

Junoh, M.N. 2014. *Introduction to the capabilities of the Safe City Monitoring System*. Discussion, 29 Oct, Kuala Lumpur, Malaysia.

Kahaki, S.M.M. & Nordin, M.J. 2011. Highway traffic incident detection using high-resolution. *Journal of Computer Science* 7(6): 949-953.

Kasim, Z.A. 2017. Introducing the capabilities to predict the locations of serial criminals into the Safe City Monitoring System. Discussion, 11 Aug, Kuala Lumpur, Malaysia.

Lian, T. 2011. Fundamentals of Zermelo-Fraenkel set theory. Available at https://www.math.uchicago.edu/~may/VIGRE/VIGRE2011/REUPapers/Lian.pdf.

Malleson, N. & Andresen, M.A. 2015. The impact of using social media data in crime rate calculations: Shifting hot spots and changing spatial patterns. *Cartography and Geographic Information Science* 42(2): 112-121.

MATRIX: The center for digital humanities & social sciences 2017. Available at http://matrix.msu.edu/~johnsnow/images/online_companion/chapter_images/fig12-5.jpg.

Mun, K.K. 2015. *Crime Displacement to Measure the Effects of Law Enforcement*. Discussion, Mar 9, Kuala Lumpur, Malaysia.

Nordin, M.J. & Ali, A.M. 2009. Indoor navigation and localization for visually impaired people using weighted topological map. *Journal of Computer Science* 5(11): 883-889.

Rossmo, K.R. 1995. *Geographic profiling: Target patterns of serial murderers*. PhD. Diss., Simon Fraser University, British Columbia, Canada.

Snow, J. 1854. *On The Mode of Communication of Cholera*. 2nd edition. London: Wilson & Ogilvy.

Tayebi, M.A. Glausser, U. & Brantingham, P.L. 2015. *Learning where to inspect: Location learning for crime prediction*. Paper presented at IEEE international conference on intelligence and security informatics, 27-29 May, Baltimore, USA.

Valentine, P.A. 2012. *Introduction to Safe City Program*. Discussion, 13 Feb, Kuala Lumpur, Malaysia.

Valentine, P.A. 2018. *Contribution of the Safe City Program towards safer cities*. Discussion, 13 Aug, Kuala Lumpur, Malaysia.

Yu, C.H. Ding, W. Morabito, M. & Chen, P. 2016. Hierarchical spatio-temporal pattern discovery and predictive modeling. *IEEE Transactions on Knowledge and Data Engineering* 28(4): 979-993.

Willis, C.L. Evans, T.D. & LaGrange, R.L. 2009. "Down Home" criminology: The place of indigenous theories of crime. *Journal of Criminal Justice* 27(3): 227-238.

Zheng, X.F. Yang, C. & Ma, Z.Y. 2011. *A mathematical modeling approach for geographical profiling and crime prediction*. Paper presented at IEEE international conference on software engineering and service science, 15-17 July, Beijing, China.

Mohammed Ariff Abdullah (corresponding author)
Center for Artificial Intelligence Technology
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor, MALAYSIA.
E-Mail: riffref@yahoo.com

Siti Norul Huda Sheikh Abdullah
Center for Cyber Security
Faculty of Information Sciences and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor, MALAYSIA.
E-Mail: snhsabdullah@ukm.edu.my

Md Jan Nordin
Center for Artificial Intelligence Technology
Faculty of Information Science and Technology
Universiti Kebangsaan Malaysia
43600 UKM Bangi, Selangor, MALAYSIA.
E-Mail: jan@ukm.edu.my