

An Enhanced Approach for Detecting Crime Hotspots

Sherif M. Akl, Assem A. Alsawy, Bahaa T. Shabana, and Hesham A. Hefny

Computer & Information Sciences Dept., Institute of Studies and Statistics Research,
Cairo University, Egypt

Shereefakl2@gmail.com, asem_issr@yahoo.com, dr.bahaa.shabana@gmail.com, hehefny@cu.edu.eg

Abstract

This Paper introduces a method for improving the classification process for hotspots and focuses on finding spatial criminal hotspots using the weighted estimates of crime experts also, visualizing the hot locations on map using GIS technology enable the decision makers to build a lot of shapes for criminal analysis and solutions. Also it determines the most gravity crime types and criminal hotspots by ranking the locations of committed crimes. The proposed approach can potentially help people and decision makers knowing accurately hot crimes, hot places and its ranks.

Keywords: *Hotspot, Classification, Weighted estimates, Naïve Bayes, GIS.*

1. Introduction

Crimes are a common social problem that affects the quality of life and the economic growth of a society [1]. Sometimes, they are considered an essential factor that determines whether or not people move to a new city and what places should be avoided when they travel [1]. The solution introduced by this paper can potentially help people stay away from a locations on a certain cities for saving their lives, Cars and their Owns. In addition, having this kind of knowledge would help people improving their living place choices. Moreover, this would be useful for police resources reallocation. It can help in the distribution of police at most likely crime places for any given time, to grant an efficient usage of police resources [1].

Locations that the research refers to have a frequent crimes defines a high crimes occurrence as hotspots locations .Within this definition, the phenomenon of hot spots appears to be widespread in a lot of Egypt cities. Therefore, by having all of these information available and the new approach used, we hope to make our community safer for the people living here and also for others who will come.

1.1. Criminological Theory

The main question in criminological theory is the rise of empirical work on how and where crimes happen. This work has taken various units of analysis. There is increasing evidence showing that some locations have much repeat victimization while other places have no crime events. Therefore, this research focuses on the places of reappearance crimes to discriminate these places which are considered as hotspots from other which are less in crimes rate.

In the last few decades, Crimes were spread via large number of places in Egypt. Some areas have more crime than others, and some others have different kinds of crimes than others. Therefore this paper classifies the crime places using the power of Naïve Bayes classifier to detect hot spots. The determination of hotspots based on the introduced

classification approach, helps decision maker for finding suitable locations to establish new police stations which help for crime reduction.

1.2. Naïve Bayes Classifier

Naïve Bayesian classifier is a supervised learning algorithm, which is effective and widely used. It is a statistical model that predicts class membership probabilities based on Bayes' theorem [1]. It assumes the independent effect between attribute values. While our selected crime features have an independent effect on each other, this classifier was an ideal choice [1]. Naïve Bayes classifier now is used in many different fields and technical arenas to address a variety of questions and problems. The classification based on Naïve Bayes and weighted estimate presents a relatively new approach to determine the gravity of each crime type. Upon this determination we can classify the repeated high gravity crime locations as hotspots and unsafe locations.

1.3. Visualizing Tool

Geographical Information Systems (GIS) are important technology for spatial planning [6]. It helps understanding relationships, patterns, and trends. This paper adopted ARC-MAP visualization tool for representing hot locations and help decision makers follow their goals on a visualized real system.

2. Problem Definition

There are a countless of work done related to crimes. Large datasets have been reviewed, and information such as location and the type of crimes have been extracted to help people follow crime rates. Existing methods have used these databases to identify crime hotspots based on locations. There are several maps applications show the exact crime location along with the crime type for any given city. Even though crime locations have been identified, but there is no information available that includes ranking the crime upon its types and ranking the locations according to its severity. Also there are no approach gives police offices a mapping model for identifying and ranking location according to its seriousness.

3. Research Objectives

Using of weighted estimates given by the experts can give reasonable solutions for different situations in crime world. This research introduces enhanced approach of Naïve Bayes' classifier and weighted estimates to give each location the correct classification based on location seriousness. This would be useful for police resources reallocation. It can help in the distribution of police at most likely crime places for any given time, to grant an efficient usage of police resources.

4. Detecting Hotspots by Classification

The proposed solution used Naïve Bayes classifier. The details of used datasets are shown in TABLE I.

4.1. Using Naïve Bayes

Naïve Bayes is the most widely used classifier; it's a conditional probability model [1]. The process of Naïve Bayes is illustrated as follows:

For a given problem instance to be classified, represented by a vector $\{x\} = \{x_1, x_2, \dots, x_n\}$ is N objects, each object has d feature values as $x_i = \{f_j, \dots, f_d\}$ $i = 1 \dots N$ and $j = 1 \dots d$. representing some n features (independent variables).

5. Experimental Results

5.1. Cairo Crimes Dataset

The dataset represents the real-world crimes occurred in Cairo [9]. It includes criminal offenses and crime incidents in Cairo. 98% of the crimes in the dataset occurred in the year 2016 while the other 2% of the crimes occurring before 2016. The dataset information was obtained from Middle East Monitor website, owned by the Arabia Inform Company which works in collecting Egyptian daily Newspapers [9]. The dataset is composed of 36 Town with 81 of crime instances. The crime category was ranked according to the experts, who ranked crime from the most gravity on public security, Such as Terrorism, Murder and Theft, consecutively.

Table 1. Cairo crimes and its content values

Postal code	Locations	Terrorism	Killing	Drugs	Robbery	Accidental Accident	Human rights	Commercia l crimes
11835	Tagamoaa	6	90	6	60	72	28	16
12616	Bolakeldakror	3	86	165	210	30	12	15
13632	Qaliob	2	24	28	96	20	18	9
11567	Kasrelnil	1	19	16	60	15	10	6
11582	Kasrelainy	1	8	9	15	22	8	3
11619	Rodelfarag	7	32	72	32	42	23	10
12961	Awseem	3	20	37	10	25	38	52
11639	Al-darbelahmar	9	112	91	88	107	66	82
11816	Nasrcity	29	180	124	210	205	140	259
11634	Al-basateen	3	70	112	38	58	49	90
11936	Al-qatamia	1	60	48	85	24	14	26
11632	Masrelqadema	35	170	180	220	130	68	104
11613	Abdeen	6	50	99	58	44	68	45
11653	Al-Sharabia	3	48	57	29	65	36	28
11638	Babelkhalq	2	21	117	81	45	30	49
11571	Elmokattam	7	18	41	49	33	10	46
11611	Bolakaboelela	1	44	102	118	58	35	94
11833	Al-marg	22	109	130	201	105	24	83
11732	Manshyetnasser	3	28	82	59	75	26	42
11659	Al-zawyaelhamra	5	80	100	85	48	49	90
11931	Sakrquorysh	2	40	63	56	70	42	25
11918	Helwan	11	104	109	84	102	60	66
11511	Al-musky	2	68	115	189	49	42	195
11663	Hadyekelqobba	1	24	40	28	24	6	10

Follow Table 1. Cairo crimes and its content values

Postal code	Locations	Terrorism	Killing	Drugs	Robbery	Accidental Accident	Human rights	Commercial crimes
11782	Ainshams	12	95	109	130	69	44	56
11581	Al-Khalifa	1	86	110	71	28	64	38
12952	Elsaf	2	18	62	15	40	18	6
11745	Darelsalam	4	114	120	89	82	58	36
11684	Al-dowyqa	1	43	98	86	28	40	11
11522	Ramses	4	28	91	219	108	89	188
11736	Elnozha	1	28	22	60	12	2	13
13763	Elkhanka	2	15	110	28	15	3	17
13766	Shubra	9	188	210	165	210	96	87
11724	Helmyetelzaytoun	1	59	91	79	41	22	40
11788	Madinetelsalam	5	105	96	210	216	189	205

Step 1. The following formula is used to calculate the probability of occurrence of each crime type by a location. In condition of specific crime location, where T represents the location c and X_i represents the crime type.

$$P(T_c | X_i) = \frac{P(T_c)P(X_i|T_c)}{P(X_i)} \tag{1}$$

T: refers to the location
 X: refers to crime type

Step 2: Calculating the probability of occurrence of a crime at "C" Location from dividing all the crimes in this location by all crime events.

$$P(T_c) = \frac{\sum(X_c)}{\sum(X)} \tag{2}$$

$$P(X_i) = \frac{\sum(X_i)}{\sum(X)} \tag{3}$$

Step 3. Finding the probability of occurrence of a crime "X" at location T from dividing the probability of intersection between them meaning that, all X_i crimes in T locations by all crime in this location T.

$$P(X | T) = \frac{P(T \cap X)}{P(T)} \tag{4}$$

Step 4. Calculating the crisp formula after obtaining its variables.

$$P(T_c | X_i) = \frac{P(T_c) P(X_i | T_c)}{p(x_i)} \tag{5}$$

Step 5. Weighted Estimates

By multiplying the crisp formula which represents the probability of crimes in each location with the weight (W_i) of each corresponding crime type (as shown in table 3) which represents the seriousness of each crime type according to the expert's opinions to obtain the (R_i) which is the final weight of each Location from each crime type.

$$R_i = \frac{P(Tc)P(Xi|Tc)}{P(Xi)} * W_i \tag{6}$$

Table 2. Cairo crimes values after multiplying Weighted values

M1	M2	M3	M4	M5
0.029	0.039	0.0016	0.009	0.0036
0.014	0.037	0.0424	0.0315	0.0016
0.01	0.01	0.0072	0.0145	0.0024
0.005	0.008	0.004	0.009	0.0012
0.005	0.003	0.0024	0.002	0.001
0.034	0.014	0.0184	0.005	0.003
0.014	0.009	0.0096	0.0015	0.0048
0.043	0.049	0.0232	0.013	0.0084
0.139	0.078	0.032	0.0315	0.018
0.014	0.03	0.0288	0.0055	0.0062
0.005	0.026	0.012	0.0125	0.0018
0.168	0.074	0.0464	0.0325	0.0088
0.029	0.022	0.0256	0.0085	0.0088
0.014	0.021	0.0144	0.0045	0.0046
0.01	0.009	0.0304	0.012	0.0038
0.034	0.008	0.0104	0.0075	0.0012
0.005	0.019	0.0264	0.0175	0.0044
0.106	0.047	0.0336	0.03	0.003
0.014	0.012	0.0208	0.009	0.0034
0.024	0.035	0.0256	0.0125	0.0062
0.01	0.017	0.016	0.0085	0.0054
0.053	0.045	0.028	0.0125	0.0078
0.01	0.03	0.0296	0.028	0.0054
0.005	0.01	0.0104	0.004	0.0008
0.058	0.041	0.028	0.0195	0.0056
0.005	0.037	0.028	0.0105	0.0082
0.01	0.008	0.016	0.002	0.0024
0.019	0.05	0.0312	0.013	0.0074
0.005	0.019	0.0248	0.013	0.0052
0.019	0.012	0.0232	0.0325	0.0114
0.005	0.012	0.0056	0.009	0.0002
0.01	0.007	0.028	0.004	0.0004
0.043	0.082	0.0536	0.0245	0.0124
0.005	0.026	0.0232	0.012	0.0028
0.024	0.046	0.0248	0.0315	0.0242

A summation (L) of each location crimes types will give a value for each location which calculated from the formula:

$$L = \sum \frac{P(Tc)P(Xi|Tc)}{p(Xi)} * W_i \tag{7}$$

R_i: Represents the weight of each Location from each crime type.

L : Represents the weight of each Location from all crime types within this location.

The Experts neglected the low crimes estimates such as Human rights, Commercial crimes, Personal and Traffic.

Table 3. A weighted value of crime types upon its seriousness according to expert estimates

Crime Type	Terrorism	Killing	Drugs	Theft	Human Rights	Commercial Crimes	Accidental	Personal	Traffic
Weight	1	1	0.8	0.5	0	0	0.2	0	0

"1" value represents most high estimate while "0" the lowest estimate.

Step 6. Calculating the value of each location crime type after obtaining the weight for every location crime type by multiplying every value in each town by its corresponding crime weight using this formula:

$$R_i = P (T) * W_i \tag{8}$$

Step 7. Making a summation for each row values to obtain the weight of each location.

$$L = \sum R_i \tag{9}$$

Step 8. Identifying results values using the following formula to obtain continuous values. Normalize the result between 0 and 100

$$Deg = \frac{(V-\min(n))*100}{\max(n)-\min(n)} \tag{10}$$

Where V refers to the column value being normalized and n is the column set of values which select min and max from its content.

Step 9. Finally classifying the results by its Ranks from {A, B, C, D, and E}

Table 4. Cairo crimes with ranked values

Locations	Round %	Ranks	Seriousness
Tagamoaa	22	D	M
Bolakeldakror	36	C	H
Qaliob	10	D	M
Kasrelnil	4	E	L
Kasrelainy	0	E	L
Rodelfarag	19	D	M
Awseem	8	E	L
Al-darbelahmar	39	C	H
Nasrcity	90	A	H
Al-basateen	22	D	M
Al-qatamia	14	D	M
Masrelqadema	100	A	H
Abdeen	25	D	M
Al-Sharabia	14	D	M
Babelkhalq	16	D	M
Elmokattam	15	D	M
Bolakaboelela	19	D	M
Al-marg	65	B	H
Manshyetnasser	14	D	M
Al-zawyaelhamra	28	D	M
Sakrquorysh	14	D	M
Helwan	42	C	H
Al-musky	28	D	M
Hadyekelqobba	5	E	L
Ainshams	44	C	H
Al-Khalifa	24	D	M
Elsaf	8	E	L
Darelsalam	34	C	H
Al-dowyqa	17	D	M
Ramses	27	D	M
Elnozha	6	E	L
Elkhanka	11	D	M
Shubra	64	B	H
Helmyetelzaytoun	18	D	M
Madinetelsalam	43	C	H
Al-jabalelahmar	7	E	L

Thus, the locations were ranked into 5 levels of Crime's seriousness as shown in the following table. Also each rank was classified into the degree indicate to the level of seriousness on the national security.

Very High, High, Medium, Low and Very Low Crimes.

Table 5. Ranked color according to its gravity

S	Rank	Linguistic	Color
1	A	Cities which has very high gravity crimes	
2	B	Cities which has high gravity crimes	
3	C	Cities which has medium gravity crimes	
4	D	Cities which has low gravity crimes	
5	E	Cities which has very low gravity crimes	

5.2. The result in Arc Maps

As shown at the map, high and over medium gravity crimes' locations are characterized by red color to show hot spot for these locations such as (Nasr City, Shubra, AL musky, etc.)

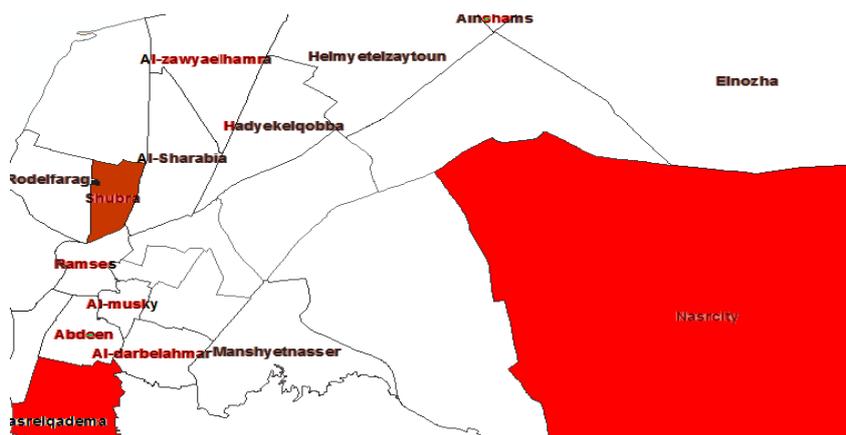


Fig.1 High Hotspot locations by red color

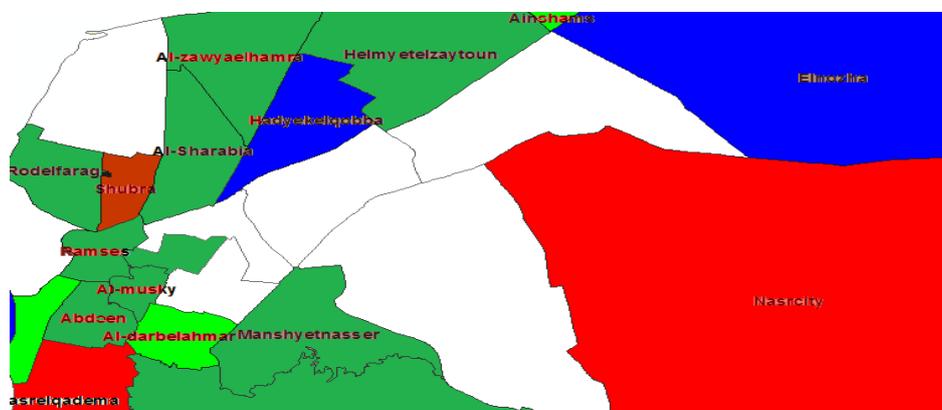


Fig.2 Hotspot locations upon its gravity by many colors

6. Conclusion

A new method of a classification using weighted estimate for detecting hotspots of crime has introduced. Crime mapping and spatial analysis are significant tools for analysis and visualizing crime data. Naïve Bayes classifier with weighted estimate has produced a proper method for ranking crime hotspots according to its severity. Using GIS gives a good visualization for crimes locations. It facilitates spotting high severity locations that needs more attention from those who are interested in fighting against crime. Ranking the crime hotspots indicates the seriousness of these locations. The experts and decision makers in a crimes field can potentially use the proposed approach, and modify the weights according to their needs.

References

- [1]. T. Almanie, R. Mirza, E. Lor, “Crime prediction based on crime types and using spatial and temporal criminal hotspots”, *International Journal of Data Mining & Knowledge Management Process (IJDKP)* Vol.5, No.4, July 2015.
- [2]. R. Kaur, S.S Sehra."Analyzing and displaying of crime hotspots using fuzzy mapping method." *International Journal of Computer Applications* 103.1 (2014).
- [3]. M. Ahmed and R. S. Salihu, “ Spatiotemporal Pattern of Crime Using Geographic Information System (GIS) Approach in Dala L.G.A of Kano State, Nigeria”, *American Journal of Engineering Research (AJER)*, pp. 51–58, 2013.
- [4]. E. POLAT, “Spatio-temporal crime prediction model based on analysis of crime clusters”, Middle East Technical University, graduate school of natural and applied sciences, September, 2007.
- [5]. N. Levine, “Crime Mapping and the Crimestat Program”, *Geographical Analysis* 38.1, the Ohio State University, pp. 41–56, 2006.
- [6]. V. Badenko, D. Kurtener "Fuzzy Modeling in GIS Environment to Support Sustainable Land Use Planning", 7th AGILE conference on geographic information science. Vol. 29, 2004.
- [7]. T. H Grubestic, T. Murray, “Detecting Hot Spots Using Cluster Analysis and GIS”, *Proceedings from the fifth annual international crime mapping research conference*. Vol. 26. 2001.
- [8]. S. Chakravorty, “Identifying crime clusters: the spatial principles”, *Middle States Geographer* 28 (1995): 53-58.
- [9]. Memonitor.com. [Online]. Available: [https:// www.memonitor.com](https://www.memonitor.com). [Last visited: 2-January-2017].
- [10]. L.W.Sherman, "Hotspots of crime and criminal careers of places", *crime and place* 4(1995): 35-52.