



Weight based Approach using Rossmo's Equation for Criminal Policing

Simran Syal
Research Scholar
Computer Science and Engineering
CGC Group of Colleges
Mohali, Punjab, India

Mandeep Singh Tinna, PhD
Associate Professor
Computer Science and Engineering
Chandigarh University
Mohali, Punjab, India

ABSTRACT

This paper discusses the approach of criminal policing by using Rossmo's equation and proposing weight based approach wherein nine types of crimes are put into four clusters and using raw data collected from online sources, weight is calculated based on the number of previous crimes committed. The result is put to use for analytics using rapid miner tool and graph is generated as the final result. Yorkshire ripper case is used as the experimental result to prove the convict taking some of his previous committed crime as data input along with other criminals in the area committing different crimes. The end result generated proves the algorithm discussed in the paper.

General Terms

Predictive Analysis, Data Mining

Keywords

Rossmo's equation, Cluster, Rapid miner tool, Probabilty, Policing.

1. INTRODUCTION

A criminal has a definite approach towards committing crime and it can be generally predicted by using some mathematical equations and analyzing the results. Criminal policing is inhibiting crime activity by actively resolving ongoing crime committed in an area and try to prevent any further crime to take place by applying real-time analysis of situation of area.

However, it is never 100% true that all the time crime can be cracked but if sufficient data is present which in most cases is valid, policing can be more and more simplified by making use of predictive analysis and mathematical tools.

Rossmo's equation was developed and patented by criminologist Kim Rossmo and it widely used in various researches and software tools to do geographic profiling based on intensity of crimes and various other crime related studies.

The proposed weight based approach using Rossmo's equation explained in this paper uses large datasets, however due to space constraint small datasets are used. There has been study which proves that based on previous crime trends of a region, we can predict future crime trend and hence this theory has been incorporated in the approach and verified through both abstract data and experimental data.

In this paper, weight based approach is defined and verified on the result drawn after applying Rossmo's [1] equation. This paper talks about nine types of crimes which are classified into four different clusters based on the nature of

crime which can however be changed as accordingly required. Previous datasets on crime has been taken and applied to predict the criminal based on the crime committed in the area.

2. APPROACH ARCHITECTURE

Past records of criminals are taken from database and a mean crime probability is calculated using the Rossmo's equation. Dataset 1 => Area where past crime is committed by the criminal(s). Areas are divided into sectors for a particular city.

Dataset 2 => Types of crime committed by the criminal. This paper focusses on nine types of crime; Murder, Rape, Assault, Kidnapping, Burglary, Theft, Extortion, Fraud and Bribery.

Each dataset is maintained with respect to each criminal's name and father's name.

A mean sector (area) is being calculated using the mean mathematical formula and noted against each criminal.

$$\text{Mean Area (x)} = \sum (\text{Criminal Crime Area}) / \text{Total Crimes} \quad (1)$$

Mean Area is calculated for each criminal. Rossmo's equation is used to calculate the probability of each criminal doing crime in particular sector.

2.1 Rossmo's Equation

$D = \sum (\text{Absolute (Crime Area} - \text{Mean Area)})$, where area is calculated in x, y co-ordinates and D is mean distance to find criminal.

$B = \text{Buffer zone} = \text{Absolute (Farthest Area} - \text{Mean Area)}$ in x, y co-ordinates.

Buffer zone is area in which Rossmo's equation holds true for probability generated for a particular criminal.

Probability = $1 / (D^f)$ if $D > B$,

$$(B^g) / 2 \times ((B-D)^{(f-g)}) \text{ if } D \leq B \quad (2)$$

Where $f = 0.8$, $g = 0.5$ which is constant and taken from experimental result calculated in "A mathematical modeling approach for geographical profiling and crime prediction" journal [2]

2.2 Weight based Approach

Past crime type data over a period of time (say, 3 years) is taken from the criminal database and data is used to calculate the future crime trend.

The nine types of crimes are studied and put in the paper into 4 categories based on its severity (however the categorization can be changed which is not a part of study in the paper):

- C1: Murder, Rape
- C2: Assault, Kidnapping
- C3: Burglary, Theft, Extortion
- C4: Fraud, Bribery

The individual cluster [3] weight is calculated by dividing the total with the individual cluster.

General equation used to calculate the weight,

Individual Cluster weight = Number of crimes committed falling in one cluster / Total number of crimes over a year.

Example: C1 weight = C1 / (C1 + C2 + C3 + C4) (3)

The individual cluster weight is taken year by year and its mean value is used to predict the ongoing year crime trend.

C1 weight (2015) = (C1 weight (2012) + C1 weight (2013) + C1 weight (2014)) / 3 (4)

The data is taken from Integrated Network for Societal Conflict Research (INSCR) which were prepared by researchers associated with the Center for Systemic Peace and compiled using open source information.

Now after calculating the results gathered by applying Rossmo’s equation and cluster data, the trend is plotted by using the final result which is generated by using both data.

2.3 Final result

The crime type of each criminal is taken and a value is generated by adding up the weights of cluster in which his or her previous crimes fall.

This data is plotted against the data generated by Rossmo’s equation for each criminal using cluster weight as signifying factor against criminal’s activity in a particular area or sector. This can be visualized by a sample trend below using a sample data.

TABLE 1: Data Result of Sample Crime Database

Criminal name	Mean Crime Sector	Probability of Criminal by Rossmo’s Equation	Crime weight of criminal (cluster weight)
Sandeep	3	0.943696	0.354839
Kewal Raj Kapoor	6	0.950845	0.354839
Hafiz	9	0.943981	0.33871
Jha	12	0.935329	0.483871
Monu	13	0.94947	0.354839

Munawwar	13	0.94947	0.33871
Ravi	14	0.954418	0.516129
Francis Dsouza	16	0.931699	0.354839
Jacob	18	0.974988	0.693548
JITENDRA	22	0.925899	0.693548
Ranchandra Kalsangra	23	0.925899	0.709677
Sonu	25	0.9519	1.03226
Dilip	29	1.06561	0.354839
Megha	29	1.06561	1.0

A chart was plotted using rapid miner tool taking the following data results as input to study the characteristics of the result. The data presented above is for experimental study of the method proposed and is not factual.

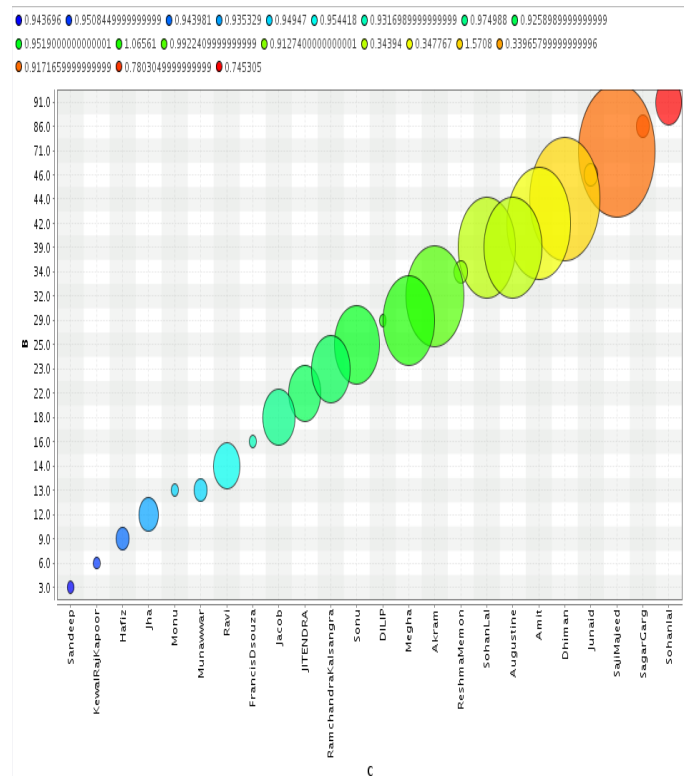


Fig. 1: Trend Plot for data result using rapid miner tool

X-axis: Criminal name

Y-axis: Mean Crime Sector, calculated using equation 1

Bubble size: Represents the cluster weight of each criminal, calculated using equation 3 and 4

Color: Represents the probability of each criminal committing crime, calculated using equation 2 and value of each color is written above the plot generated by the tool

By making use of rapid miner tool, one can easily predict the criminal who committed in a particular area, here marked as sector and hence if any crime is reported in any sector, the

cops can run the analytics using the algorithm defined in this paper and pin point to 2 to 5 criminals among whom one has possibly have committed the crime and thus a lot of policing time and record lookup can be saved as the algorithm would generate the best possible data in minimum time.

3. Experimental result (Proving Yorkshire Ripper Case Convict)

The data below is taken from crimestoppers.uk.org and compiled in the database. A small dataset is used to represent the result.

TABLE 2: Input Data From internet sources (U.K.)

Criminal Name	Crime Type	Crime Area
Leonard Mucaj	Assault	Sheffield
Dorian Greci	Assault	Sheffield
Monty Morris	Assault	Sheffield
Lucasz Zaleski	Assault	Sheffield
Koser Sadiq	Burglary	Sheffield
Freeman David Ibekwe	Fraud	Sheffield
Irij Hossini	Fraud	Sheffield
Rimon Usef	Rape	Bradford, Sheffield
Brandon Pyatt	Fraud, Theft	Sheffield, Manchester
Kym Svendson	Burglary	Doncaster
Zahoor Ahmed	Rape	Dewsbury
Peter Sutcliffe	Assault, Rape, Murder	Bradford, Leeds, Halifax, Sheffield
Troy Martin Howe	Assault	Leeds
Thomas Tooth	Assault	Manchester
Mark Edward Fitton	Burglary	Cheshire, Lancashire, Oldham, Manchester
Arslam osman Ali	Rape	Grimsby

The crime area is categorized with numerical value taking one of the area as a base point. The numerical value is assigned based on the distance from the base point or city. Here, Sheffield has been taken as the base point or base city marking it as 1. Now the distance is taken of other cities in miles and particular value is used to calculate mean area and used in Rossmo's equation.

The graphical representation of area used in shown in the map below.

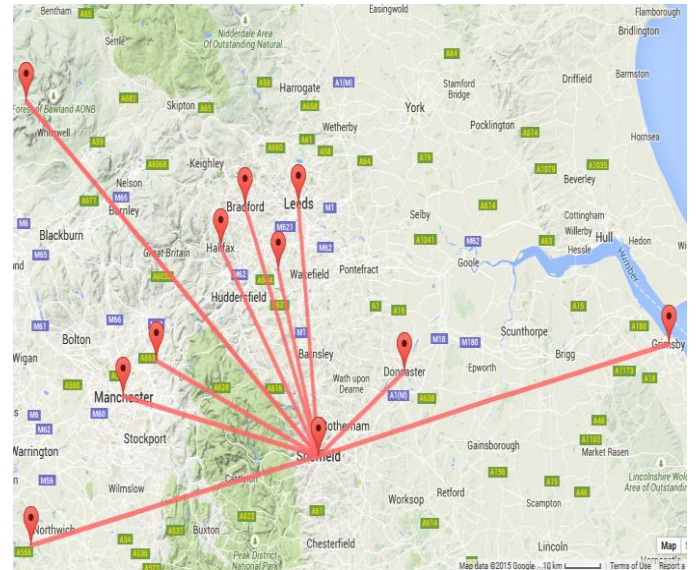


Fig. 2: Map showing cities distance mentioned in TABLE 3 from Sheffield, United Kingdom

TABLE 3: Cities with distance in miles from Sheffield (Sheffield is taken with base value 1 miles)

City	Miles from Sheffield
Leeds	29
Manchester	33
Sheffield	1
Bradford	31
Halifax	27
Grimsby	59
Cheshire	48
Lancashire	62
Oldham	28
Doncaster	17
Dewsbury	22

Buffer zone (B) is taken to be the difference value of distance of city farthest from Sheffield and mean value of distance for a particular criminal. The average constitutes the buffer zone.

In this example, famous Yorkshire Ripper Case is taken in which crime is committed in Oldham, United Kingdom.

Based on the distance value mentioned in TABLE 3, the probability of individual criminal is calculated using Rossmo's equation and taken the result as input in the rapid miner tool.

The rapid miner tool plots the trend as a graph and the graph can be put to analytics to identify the criminal. The result is justified based on the crime that is already known and was committed by Peter Sutcliffe [4] and probable result that the graph shows.

However, the data used to cluster the crime types and put weightage on crime type was done from the same database mentioned in TABLE 2.

TABLE 4: Data result after applying input from TABLE 2

Criminal Name	Mean Crime Sector	Mean Cluster Weight	Probability Value By Rossmo's Equation
Leonard Mucaj	1	0.368421	0.910303
Dorian Greci	1	0.368421	0.910303
Monty Morris Foote	1	0.368421	0.910303
Lucasz Zaleski	1	0.368421	0.910303
Koser Sadiq	1	0.157895	0.910303
Freeman David Ibekwe	1	0.157895	0.910303
Irij Hossini	1	0.157895	0.910303
Rimon Usef	1	0.315789	0.910303
Jahangir Alam	16	0.315789	1.00796
Kym Svendson	17	0.157895	0.887297
Brandon Pyatt	17	0.31579	1.03127
Zahoor Ahmed	22	0.315789	0.880251
Peter Sutcliffe	22	0.68421	0.957468
Troy Martin Howe	29	0.368421	0.863812
Thomas Tooth	33	0.368421	0.854022
Mark Edward Fitton	43	0.157895	0.411973
Arslam osman Ali	59	0.315789	0.674741

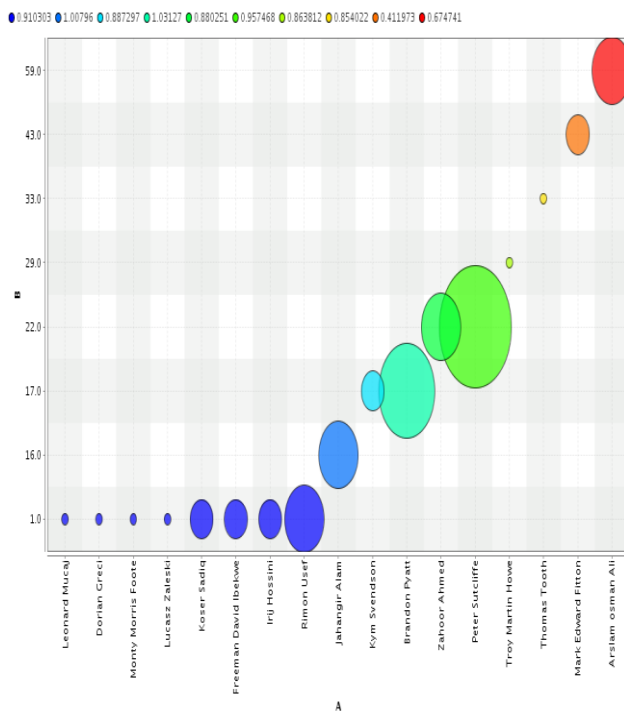


Fig. 3: Trend plot for data result using rapid miner tool

X-axis: Criminal name

Y-axis: Mean Crime Sector, calculated using equation 1

Bubble size: Represents the cluster weight of each criminal, calculated using equation 3 and 4

Color: Represents the probability of each criminal committing crime, calculated using equation 2 and value of each color is written above the plot generated by the tool

Since, the input of crime location was Oldham, United Kingdom by Peter Sutcliffe and Oldham has value defined to be 28 in TABLE 3. Hence, if a line is drawn parallel to x-axis from y intercept value 28, it will cut the bubble somewhere which is light green in color. Now a straight line can be drawn parallel to y-axis from mid of this bubble, it can be seen that the x value comes out to be Peter Sutcliffe who was later proved to be the real culprit in Oldham case. Hence, it can said that by applying weight based approach on the Rossmo's equation, the criminal can be fairly predicted based on previous crime records or it can be better put together by saying, analyzing the crime pattern of criminals using appropriate algorithms and tools.

4. CONCLUSION

Though there cannot be 100% accuracy in the predictive result in all cases as it depends pretty much on previous history of crime and will fail in case of new identity criminal but it will surely help in the primary policing and in certain cases identifying the real criminal.

Since the weight based approach is applied and each iteration of result will ensure trend with the futuristic criminal statistics, hence the analysis is more granular in terms of data feed. By making use of mathematical approach, the algorithm designed proves that crimes can be reduced by effectively learning the behavioral pattern of previous crimes or criminals. Moreover, further studies can be done on the proposed approach and it can be made more accurate by working on the type of crime to be clustered together which falls in the study of a professional criminologist. The idea of this paper was to prove the weight based approach used with Rossmo's equation which came out to be accurate in the case study of Yorkshire Ripper case.

The proposed algorithm can be used in distributed computing and a central database can be maintained wherein policing can be managed across various regions and automated reports can be generated at the end to keep a check on the crime trend and moreover, crime can also be checked more accurately to a greater extent unlike the conventional approach of geographic profiling.

5. REFERENCES

- [1] D. Kim Rossmo, "Place, Space and Police Investigations: Hunting Serial Violent Criminals.
- [2] Xifan Zheng, Yang Cao, Zhiyu Ma, "A Mathematical Modeling Approach for Geographical Profiling and Crime Prediction", 978-1-4244-9698-3/11/\$26.00, 2011 IEEE.
- [3] A. Malathi, Dr. S. Santhosh Baboo, "Algorithmic Crime Prediction Model Based on the Analisis of Crime Clusters", Issue 11., vol. 11. Global Journal Of Computer Science and Technology.



- [4] Naichang Chen, Wei Shi, Dawei Song, “Prediction of Series Criminals: Approach based on Modeling”, IEEE 2010 International Conference on Computational and Information Sciences.
- [5] Devlin, Keith J., Gary Lorden, “The Numbers behind NUMBERS: Solving Crime with Mathematics,” New York, NY: Plume published by Penguin, 2007 print.
- [6] Rossmo, D. Kim. Geographic Profiling. Boca Raton, FL:CRC, 2000. Print. Spatial Aspects of Crime Theory and Practice. Pearson Custom Pub, 2005 Print.
- [7] Putwain, David and Aidan Sammons. Psychology and Crime. East Sussex: Routledge, 2002. Print.
- [8] Mohammad A. Tayebi, Martin Ester, Uwe Glasser, Patricia L. Brantingham, “Crimetracer: Activity Space Based Crime Location Prediction”, 2014 IEEE/ACM International Conference on Advances in Social Networks Analysis and Mining (ASONAM 2014).
- [9] Zhu Shiyao, “ A research on dealing with serial crimes with mathematical method”, IEEE 2011 978-1-14244-9698-2.
- [10] Beth Pearsall, “Predictive Policing: The Future of Law Enforcement?” NIJ Journal, No. 266.
- [11] Chung-Hsien Yu, Max W. Ward, Melissa Morabito, Wei Ding, “Crime Forecasting using Data Mining Techniques”, 11th IEEE International Conference on Data Mining Workshops.