

Prediction of Cybercrime by using Euclidean Voronoi on Persons Arrested under Cybercrime

R. Deepalakshmi



Abstract: Cybercrime has many dimensions. Mails sent from one person to another may have much illegal content. Cybercrime has gained recent significance in current situation. Mathematical models have proved to give more accurate results in Cybercrime analysis and prediction. Euclidean distance with Voronoi diagrams is a recent mathematical diagram which is divided into regions based on points in a plane. Thiessen polygons gain popularity in science and technology. This paper finds the solution to analyse data based on persons arrested under cybercrime.

Keywords : Cybercrime, Euclidean distance, Voronoi diagrams, Thiessen polygons

I. INTRODUCTION

Cyber crime is the great threat among today's world. In circumstances, when there are no comprehensive trust between sender and receiver, some method is needed to prevent cybercrime. Cyber crime is the crimes in which any ICT devices has been used knowingly or unknowingly [1]. Cyber attacks done purposefully are termed as cybercrimes. These have many side effects both politically and economically. Cyber fraud is a challenge in many fields. Euclidean Voronoi is more useful in finding different types of cyber crimes in various fields. Voronoi diagram, provides a solution to most complicated problems. Cyber crime problem which is included includes calculation of Euclidean distance and nearest neighbouring algorithm. Data mining also provides a vital role in statistical implementation of data in crimes. The patterns that are discovered from spatial data or crime data is of much use to investigation team. Since cybercrime is much sensitive area, we can encompass the mathematical model to predict the result and for future proceedings [2]. The proposed method using Euclidean distance and Voronoi diagram not only identifies the crime but also useful for future predictions. It is used for identifying that the results produced by this type prove to be important [3].

Manuscript published on 30 September 2019

* Correspondence Author

Dr. R. Deepalakshmi*, Computer Applications, Tamilnadu Dr. Ambedkar Law University, Chennai, India. Email: profdeepalakshmi@gmail.com

© The Authors. Published by Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP). This is an [open access](https://creativecommons.org/licenses/by-nc-nd/4.0/) article under the CC-BY-NC-ND license <http://creativecommons.org/licenses/by-nc-nd/4.0/>

II. LITERATURE SURVEY

This research paper has various data mining methods for cybercrime detection in vast area. Among those some are proven to be successfully implemented, while others are in the process of advancement. Everyone will be affected from the succession of fraud irrespective of the management of that particular company. Every organization must be protected with firewall to minimize the effect of cybercrime. Cyber-attacks were also examined in banks where brute force attack and ATM frauds are detected [4]. Networking algorithms also help in detecting cyber attacks.

III. METHODS AND DISCUSSIONS

Web mining plays a vital role in development of research area. Web structure, content and usage mining is widely used in cyber crime investigation. Web mining acquires best knowledge and has proven effective in the field of cyber crime investigation. As Internet is ruling over the world, most of cyber crime is knowingly or unknowingly committed by a normal user who is new to computers.

A. Web Mining and its usage in detecting Cyber crime

Web content mining needs innovative applications of text mining and its unique approaches. In recent years we can see the sudden raise of web mining. Digital evidence is collected from textual data by means of blogs, e-mails, SMS, MMS. Many tools are used to extract information.

B. Tools for extracting information

1. Scrapy: is a good web mining tool. It helps in extracting data from the websites. It is a genuine web scraping tool.
2. BIXO: web mining tool available in open source.
3. Oracle Data Mining: is intended by Oracle. It is excellent data mining software, encompassing many data mining algorithms mainly used in predicting data.

C. How tools are used in cybercrime

Investigation of cyber data is essential in web mining and data mining techniques is used for that. Web mining and criminal mining are used in Data Mining Techniques. If a particular webpage is taken for review, using data mining methods, succession methods of crimes involved can be detected.

IV. IV DATASET

The data set for analysis is taken from the Government website data.gov.in under cyber crime website.



Prediction of Cybercrime by using Euclidean Voronoi on Persons Arrested under Cybercrime

Dataset refers to the person arrested by various age categories under different section of IT Act and IPC sections detected under cybercrime during the year 2018. This crime statistics is provided by the Ministry of Home Affairs and National Crime Records Bureau (NCRB).

Table- I: Partial Table showing cyber crime details of the year 2018

State/UTs	Crime Heads	Below 18 Years	Between 18-30 Years	Between 30-45 Years	Between 45-60 Years	Above 60 Years	Total
ANDHRA PRADESH	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	4	3	1	0	8
ARUNACHAL PRADESH	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	0	0	0	0
ASSAM	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	0	0	0	0
BIHAR	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	0	0	0	0
CHHATTISGARH	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	0	0	0	0
GOA	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	0	0	0	0
GUJARAT	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	4	3	0	7
HARYANA	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	5	0	0	0	5
HIMACHAL PRADESH	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	0	0	0	0
JAMMU & KASHMIR	TAMPERING COMPUTER SOURCE DOCUMENTS (SECTION 65 IT ACT)	0	0	0	0	0	0

A. Euclidean distance to find the crime

Euclidean distance is a technique used to find the distance/dissimilarity among objects. Euclidean distance ($\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$) = 0

$$\sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2} = 0$$

This file contains the Euclidean distance of the data after the min-max, decimal scaling, and Z-Score normalization.

Euclidean distance Calculations in excel after normalization - min max, z score and decimal scaling

B. Proposed formula for predicting cybercrime

Use Euclidean distance measure $Dis\sqrt{(\sum (\beta-N))/n}$.

This formula can be used on the cyber data set to find the similarities. Data is divided into uniform grid cells. Each point must process to the prescribed locations [5]. A window which is on mid of each cell is passed through study area. Number of crimes is calculated based on the formula $\lambda(s) = \sum(1/\lambda^2)R(S-Si/\lambda)$ [6].

Here

k refers to kernel function

s refers to grid reference point

sirefers to point event being measured for intensity

trefers to bandwidth

R – study region (not part of formula)

Every point that lays within the area of the kernel function is being progressed to reference point. Each point that falls within the bandwidth of the kernel function as it progressively moves from one grid reference point to the next contributes to an estimate

of the intensity of the point process [7]. When the calculation is completed for all the surface points, hotspots are generated.

C. Image for Euclidean voranooi

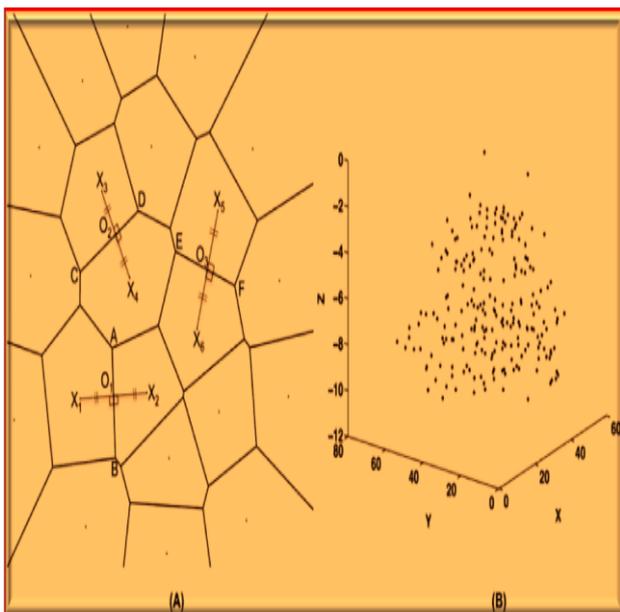


Fig.1 Crime analysis using Voronoi diagram

(A) Voronoi diagram and calculation of Euclidean distance

(B) Distribution of Crime Data

$$Vor(s) = \{p: distance(s,p) \leq distance(s^1,p), s^1 \in S^1 \}$$

Where Vor(s) is the Voronoi region for a point s.

D.Working of Diagram

O is a single point where sametypes of crimes are met. (X_i, X_n) is the set of two points which is considered in a single plane. Here X_i coordinates refer to the age and type of crimes. This is used in drawing the perpendicular bisector with X coordinates. The polygons are formed for each tuple in the data set. This Voronoi diagram is drawn only for partial dataset in cybercrime. The point in diagram B represents the collision of crimes among various ages. Pearson method also can be used to find the similarity in cyber crime among different age groups.

V CONCLUSION

This paper aims in detecting Cyber crime by using trending Mathematical concept Euclidean Voranoi. Many data mining methods are used to calculate and predict Cybercrime in different fields. Mathematical model gives an accurate result for predicting persons involved in cybercrime. Euclidean Voranoi has gained significance in predicting huge number of data. The graph obtained by Euclidean Voranoi is of much use to the cybercrime department for disrupting future crimes. The data set is selected from Tamilnadu Government Website. The analysis is performed in partial dataset where Euclidean distance is used to measure the cyber crimes based on age. Various Mathematical formulas are used to find the similarities grouped on age. Voranoi models are more are less similar to networking model for detecting cybercrimes but this proves to be more accurate compared to networking model. Detection of Cybercrimes on Persons arrested under Cybercrime is of great use to crime departments. These types of mathematical models are more accurate and these models are used for predicting the occurrence of cybercrime. It is clear in near future problems of cybercrimes can be easily solved by implementing mathematical models.

REFERENCES

1. Deepalakshmi R “Usage of social networks sites and level of awareness in cyber security”, Zenith International journal of Multidisciplinary research, volume 9, issue 2, 2019
2. Neill, D.B., A.W. Moore, F. Pereira and T. Mitchell, “Detecting significant multidimensional spatial clusters.” L.K. Saul, et al., eds. Adv. Neur. Info. Proc. Sys., 17: 969-976, 2005.
3. VijayaKumar M, Dr. Abdel-Badeeh M. Salem, Dr. S. Karthick and Dr.Hanumanthappa, 2013. “Mining Based Crime Classification Algorithms for Detecting Crime Hot Spot and Cold Spot”, Egyptian Computer Science Journal (ECSJ), 37(4): 1-13M, 2013
4. AtulBamara,”Cyber attacks and defense strategies in India”. 49-61,SSRN electronic journal, 2013
5. 5.Bowers et al, Stability of space time clusters of burglary Bowers , “British journal of Criminology”, volume 44, issue 1, 2004.
6. Gatrell, A.C., T.C. Bailey, P.J. Diggle and B.S. Rowlingson, 1996
7. 7.KavehA Survey On Cyber-Crime Prediction Techniques. GRD Journals- Global Research and Development Journal for Engineering, Volume 2, Issue 5, April 2017 ISSN: 2455-5703, 2017.

AUTHORS PROFILE



Dr. R.Deepalakshmi, received her MCA degree from The Ethiraj College for women, Chennai, Tamilnadu, with distinction .Recieved her M.Phil degree from MKU. Received her Doctorate degree from Madras University, India. She worked as Head of the Department in Sir Theagaraya College, in the department of Computer Science and Applications. Currently she is working as an Assistant professor in Tamilnadu Dr. Ambedkar Law University. She has more than 15 years of teaching experience and more than 8 years of research experience in the field of Data mining, Algorithm Analysis, Geographical Data mining, Bioinformatics, Image Mining and Cybercrime. Author has published many articles in the National and International Journals with high impact factors and has presented papers in many conferences. She has received Best teacher award 2017 and received many appreciations from many universities. Author has published many books in her subject discipline. She is a coordinator for many international conferences, seminars and workshops. Author has guided many M.Phil research scholars and has produced the best candidate in her area of specialization. She is a woman of passion who is interested to be updated and face challenges.

