Yadanabon University Research Journal, 2019, Vol-10, No.1

Discovering of Criminal Patterns in Myanmar Using Fuzzy Clustering

Min Thein*, Khin Nandar Win**

Abstract

In these days, there are a lot of crimes are occurring around the world in different variety. Discovering and predicting crime patterns are crucial issues and become one of the main community problems for the safety of our environment almost in every country. In this paper, we have discovered some criminal patterns according to their target types and the crimes types of Myanmar using fuzzy clustering and CAC algorithm in accordance with their time and location from spatiotemporal dataset. Moreover, we also presented the usefulness of the fuzzy clustering algorithm in the real-world application. And at last, we discussed frequently occurring crimes by the place where it has been targeted and the crimes types in accordance with the relevant years.

Keywords : discovering crime patterns, fuzzy, clustering, spatiotemporal, crime activity clustering, Myanmar, fuzzy Cmeans

I. Introduction

In every country, there are a lot of and different categories of crimes are taking place day by day and year by year. Therefore, it is an essential issue to discover and prevent from happening a crime in order to create a peaceful environment and an enjoyable society. It can be discovered by emphasizing on each of the crime types, and by choosing important place for establishing the harmony of the community and avoiding the serious problem.

There are abundance and different ways to discover the crimes according to the requirements of crime types. For example, fingerprint can be used to detect the terrorist, the murder or the sexual abuse, etc. Additionally, spatiotemporal data can be used to discover the time and the place of the accident can occur and can be prevented.

Many researchers have done for the criminal analysis based on the spatiotemporal data using different Machine Learning algorithms. For the discovering of the crimes, many researches have proposed, such as criminal behavioral discovering [6,7,8], Residential Burglary Patterns analysis based on Space-Time[9], Criminal Mapping in the examination of criminal activity[10], criminal classification, Criminal hotspot proposing for robbery and violent crimes in terms of space - time and time - space [11], web-based crime geointelligence platform for Maxico City [12] non-hierarchical clustering for assessing patterns of crime[8], Spatio-temporal crime hotspots and the ambient population[13], On The Application of Fuzzy Clustering for Crime Hot Spot Detection[14], Marked point process hotspot maps for homicide and gun crime prediction in Chicago[7], crime patterns mining using spatiotemporal Data [15] crime, criminal forecasting [16,17,18,14,19], algorithmic criminology [20], graph mining for criminal group investigation [21].

Our work focus on the discovering of the crime patterns in Myanmar based on the Spatiotemporal data according to the target types and the attack types using the CAC algorithm [3]. The reminder of the paper is organized as follows: we discussed fuzzy clustering algorithm, its real-world applications and its usage in criminal investigation in Section II. And, we extensively discovered the crime patterns of Myanmar by explaining fuzzy C means algorithm, the data description that is used in the experiment and the methodology in Section III. Finally, we conclude the paper in Section IV and acknowledgement in Section V.

* Lecturer, Department of Computer Science, Yadanabon University

** Student, Department of Computer Science and Engineering, Hunan University, Changsha, China

II. Related Works

A. Fuzzy Clustering Algorithm

Fuzzy Clustering is one of the clustering algorithms in unsupervised learning of Machine learning algorithm. In another words, it is called soft clustering. The main difference between soft clustering and hard clustering is that hard clustering has only two values 0 and 1 (in another words "on" and "off"). But for the name of soft clustering, the data may belong to many clusters with different belongings in percentage. In accordance with the name of fuzzy, fuzzy is presenting the results or something with their percentage and possibilities. It is the main difference with others normal hard algorithm. And fuzzy clustering means possessing the possibility of one data to all assigned clusters. The data will belong to all cluster according to their belonging degree. As an example, in one cluster, the data may belong to 50 % while the rest of the clusters may be included 10 %, 25%, and 25 %, respectively. It is called membership values in the fuzzy clustering. The detailed steps of fuzzy algorithm function is presented in section III.

B. Fuzzy Clustering in Real World Applications

Fuzzy clustering is an emerging and ongoing research topic in the area of engineering, mathematics and computer science. Fuzzy logic has been used in numerous applications such as facial pattern recognition, air conditioners, washing machines, vacuum cleaners, antiskid braking systems, transmission systems, control of subway systems and unmanned helicopters, knowledge-based systems for multiobjective optimization of power systems, weather forecasting systems, models for new product pricing or project risk assessment, medical diagnosis and treatment plans, stock trading, control systems engineering, image processing, power engineering, industrial automation, robotics, consumer electronics, and optimization [1]

C. Fuzzy Algorithm in Criminal Investigation

Crime Activity Clustering (CAC) Algorithm has been proposed by Khin et al for discovering the crime patterns based on fuzzy clustering algorithm using apache spark cloud computing platform and they additionally proposed parallel process algorithm of the CAC algorithm [2]. Furthermore, Khin et al. proposed PCPD (Parallel Crime Pattern Discovering) system based on the fuzzy clustering for large scale spatiotemporal data [3]. In that system, they proposed four algorithms for the crime activity clustering. These four algorithms are CAC membership matrix initialization algorithm, Crime-Activity-Cluster center calculation algorithm, CAC membership matrix updating algorithm, and parallel process of the Parallel Crime Pattern Discovery (PCPD) algorithm [3].

III. Criminal Pattern Discovering In Myanmar

In this section, we presented the discovering of the highest targeted place (mostly, very often targeted place) according to the crime type in one year, in the group of the years and in the period of the years based on CAC algorithm[2] which has been proposed depending on Fuzzy Clustering. Firstly, we show the nature of Fuzzy C means algorithm. And then we described the attack types and target types from the large-scale spatiotemporal database [4]. After that, we have discovered the crime patterns of Myanmar relying on these target and attack type information by the years of 1994, 1995, 1996, 1997, and 1999 as one group, from the year of 1988 to 1992 as one group and then for the year of 1988, we have implemented as another one experiment. And at last, we discussed frequently occurring types of the crimes and the place which is usually experiencing by observing the results of the implementation based on CAC algorithm.

1) **Fuzzy C Means Algorithm:** The FCM algorithm is the most widely used and one of the most famous fuzzy clustering algorithms. It allows one data to belong to two or more clusters. And it was developed by *Dunn* in 1973 and improved by *Bezdek* in 1981. And it is

often used in pattern recognition. And it depends on the minimization of the objective function [5]. Sometimes, objective function is also called target function.

$$J_m = \sum_{i=1}^{n}\sum_{j=1}^{c} w_{ij}^{m} \,||\, x_i - c_j||^2 \,, \qquad 1 \le m < \propto$$

Where m is any real number greater than 1, u_{ij} is the degree of membership of x_i in the cluster j, x_i is the i^{th} of d-dimensional measured data, c_j is the d-dimension center of the cluster, and ||*|| is any norm expressing the similarity between any measured data and the center.

Fuzzy partitioning is carried out through an iterative optimization of the objective function shown above, with the update of membership u_{ij} and the cluster centers c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{||x_i - c_j||}{||x_i - c_k||}\right)^{\frac{2}{m-1}}}, \qquad c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \,.\, x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

This iteration will stop when $max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right\} < \delta$ where δ is a termination criterion between 0 and 1, whereas k are the iteration steps. This procedure converges to a local minimum or a saddle point of J_m.

The algorithm is composed of the following steps:

1. Initialize $U = [u_{ij}]$ matrix, $U^{(0)}$
2. At k-step: calculate the centers vectors $C^{(k)} = [c_j]$ with $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \,.\, x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

3. Update $U^{(k)}, U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{c}\left(\frac{||x_i - c_j||}{||x_i - c_k||}\right)^{\frac{2}{m-1}}}$$

4. *if* $\left\|U^{(k+1)} - U^{(k)}\right\| < \delta$ *then STOP*; otherwise return to step 2.

2) **Dataset Description**: We have taken Myanmar data from the large scale spatio-temporal dataset for our implementation [4]. There are 22 places that can be purpose as the crime destination and and 9 types of crime that can be occurred. We have described them as follows:

Description of the Nine Attack Types (Crime types)		
No	Attack Type ID	Attack Type Description
1	1	Assassination
2	2	Armed Assault
3	3	Bombing / Explosion
4	4	Hijacking
5	5	Hostage Taking (Barricade Incident)
6	6	Hostage Taking / Kidnapping
7	7	Facility/ Infrastructure Attack
8	8	Unarmed Assault
9	9	Unknown

Table. 1. Description of the Attack Type (Crime Types)

Description of the Target Types		
No.	Target type ID	Target Type Description
1.	1	Business
2.	2	Government (General)
3.	3	Police
4.	4	Military
5	5	Abortion Related
6	6	Airports & Aircraft
7	7	Government (Diplomatic)
8	8	Educational Institution
9	9	Food or Water Supply
10	10	Journal lists & Media
11	11	Maritime
12	12	NGO (Non-Government Organization)
13	13	Other
14	14	Private Citizens & Property
15	15	Religious Figures/ Institutions
16	16	Telecommunication
17	17	Terrorists/ Non-State Militia
18	18	Tourists
19	19	Transportation
20	20	Unknown
21	21	Utilities
22	22	Violent Political Party

Table. 2. Description of the Target Place

iyr	coun	country_txt	city	attacktyp	attacktype1_txt	targtyp	targtype1_txt
1988	138	Myanmar	Rangoon	3	Bombing/Explosion	19	Transportation
1988	138	Myanmar	Maesan Village	2	Armed Assault	14	Private Citizens & Property
1988	138	Myanmar	Poinsa	2	Armed Assault	14	Private Citizens & Property
1988	138	Myanmar	Rangoon	3	Bombing/Explosion	8	Educational Institution
1988	138	Myanmar	Namsam	6	Hostage Taking (Kidnapping)	21	Utilities
1988	138	Myanmar	Nawngmu village	2	Armed Assault	14	Private Citizens & Property
1988	138	Myanmar	Rangoon	3	Bombing/Explosion	7	Government (Diplomatic)
1988	138	Myanmar	Shan State	3	Bombing/Explosion	14	Private Citizens & Property
1988	138	Myanmar	Bhamo	3	Bombing/Explosion	11	Maritime
1988	138	Myanmar	Kyaikto Township	2	Armed Assault	15	Religious Figures/Institutions
1988	138	Myanmar	Kyaikto Township	3	Bombing/Explosion	15	Religious Figures/Institutions
1988	138	Myanmar	Kyaikto Township	3	Bombing/Explosion	15	Religious Figures/Institutions
1988	138	Myanmar	Kyaikto Township	3	Bombing/Explosion	15	Religious Figures/Institutions
1988	138	Myanmar	Kyaikto Township	3	Bombing/Explosion	15	Religious Figures/Institutions
1988	138	Myanmar	Kawkareik	2	Armed Assault	14	Private Citizens & Property
1988	138	Myanmar	Moegphan	2	Armed Assault	14	Private Citizens & Property
1988	138	Myanmar	Mandalay	3	Bombing/Explosion	19	Transportation
1988	138	Myanmar	Kayah State	3	Bombing/Explosion	21	Utilities
1988	138	Myanmar	Myawadi	2	Armed Assault	19	Transportation
1988	138	Myanmar	Lelitho	3	Bombing/Explosion	21	Utilities
1988	138	Myanmar	Kyaikkow	3	Bombing/Explosion	19	Transportation
1988	138	Myanmar	Maukpiin	2	Armed Assault	19	Transportation
1988	138	Myanmar	Sinbo	2	Armed Assault	4	Military

Table. 3. Description of the Crimes in the Global Terrorism Dataset that happened in Myanmar for the year of 1988.

3) Discovering Crimes in Myanmar by Crime Type and Target Place

In this subsection, we presented our observation of the crime patters in Myanmar in accordance with the crime types shown in Table (1) and the place that has been targeted shown in Table (2) .Our experiment relied on CAC algorithm [3] and the discovered crime patterns are presented in Fig.[2,4,6.8.10.12,14,16] according to the target types and the crime types. We have labeled the graphs according to the target type ID and crime type ID.

For one year discovering Myanmar crime pattern, we chose 1988 for demonstration. In Fig.4, 1988, for crime types, there are many data points in 3 and 2 for X axis. Therefore, we can see that Bombing/ Explosion and Armed Assault were frequently occurred in Myanmar for the year of 1988. And they were mostly targeted to the Private Citizen and Property (14), Transportation (19), Government (Diplomatic) (7), Educational Institution (8) and Religious figures/ Institutions (15) according to the cluster data points and the cluster centers as shown in Fig.4. Additionally, we presented widely for the year of 1988 by clustering it as two cluster in Fig. 2, three cluster in Fig. 4, four cluster in Fig. 6, five cluster in Fig. 8 and six cluster in Fig.10. And, Figures (1,3, 5, 7, 9) are presented for its cluster relevant target function and in which iteration, it is closed to the target respectively. We have shown the1988 Myanmar crime data in Table. 3 which are taken from the Global Terrorism Database [4].

For the year of 1994, 1995, 1996, 1997 and 1999, the crime patterns and the target are presented in Fig. 13 and Fig. 14. In the dataset, the year of 1998 is missing as in exception. We calculated with three clusters.

Moreover, for the period of 1988 to 1992, we have discovered the crime patterns like in the figures 11 and 12. And in Fig. 15 and Fig. 16, we have discovered crime patterns for the period of 2001 to 2010.

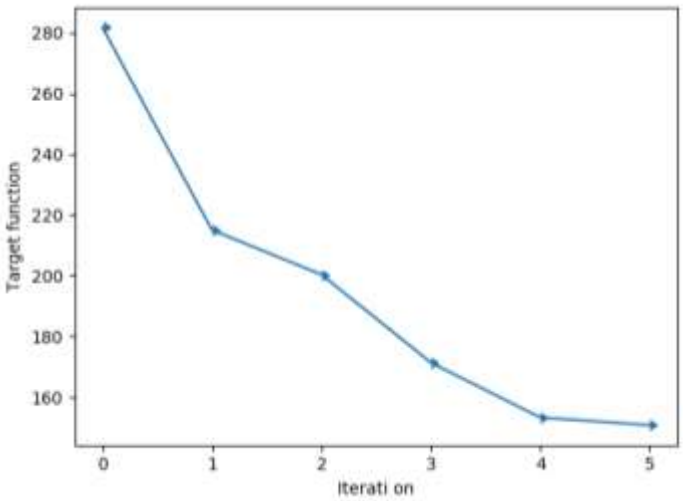

Fig.1. Two Cluster target function for the year of 1988 in Myanmar

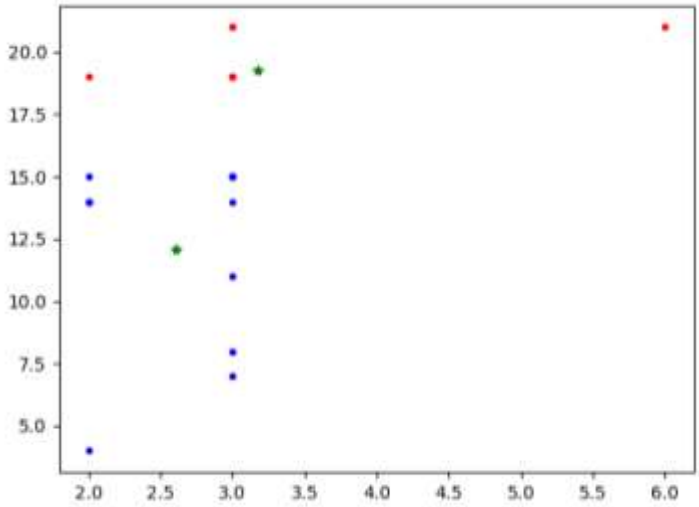

Fig.2. Two Clusters of Crime in Myanmar for the year of 1988

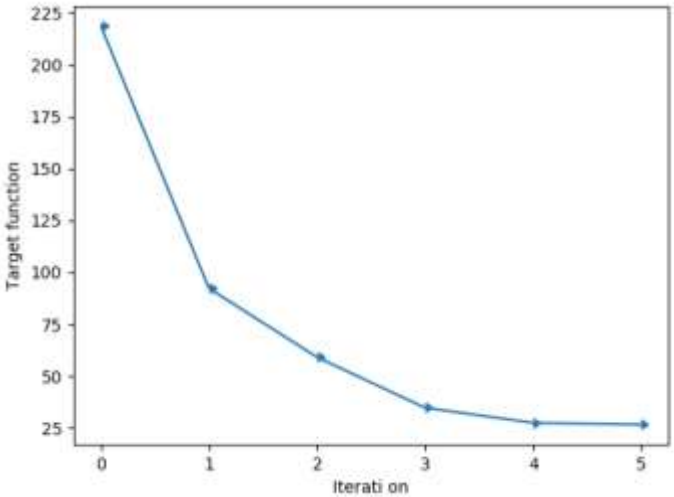

Fig.3. Three Cluster target function for the year of 1988 in Myanmar

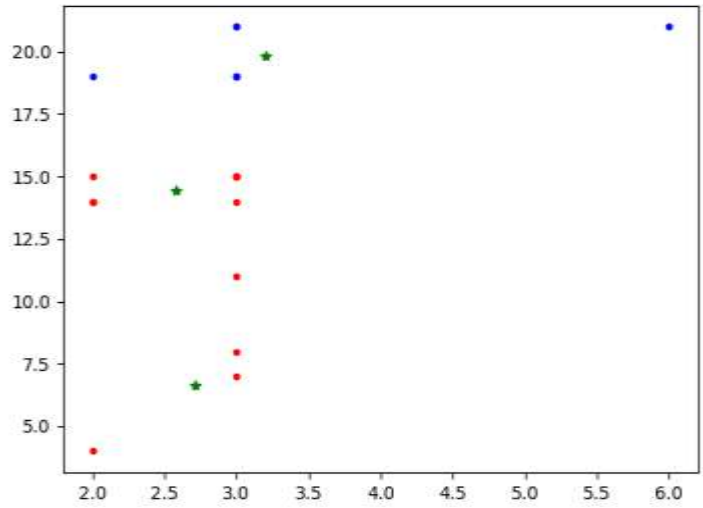

Fig.4. Three Clusters of Crime in Myanmar for the year of 1988

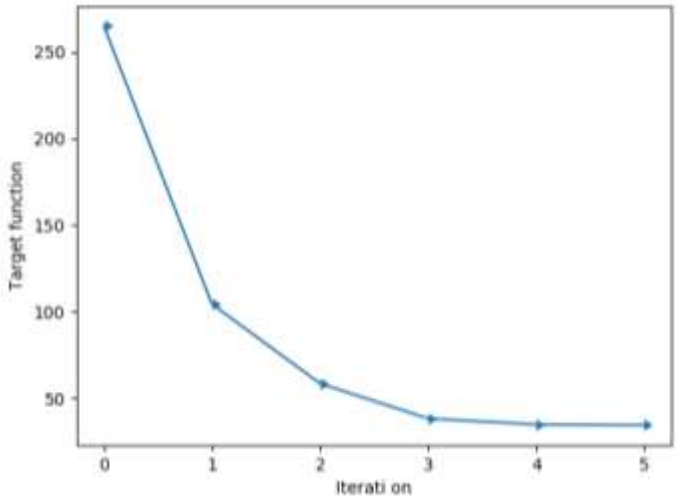

Fig.5. Four Cluster target function for the year of 1988 in Myanmar

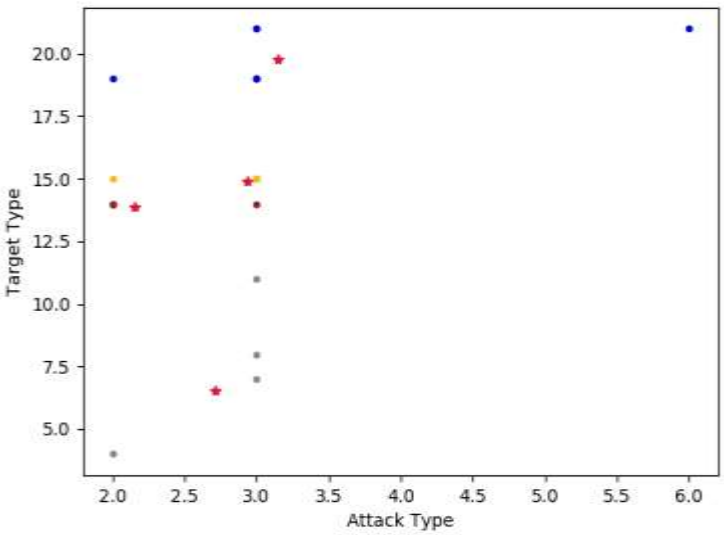

Fig.6. Four Clusters of Crime in Myanmar for the year of 1988

Yadanabon University Research Journal, 2019, Vol-10, No.1

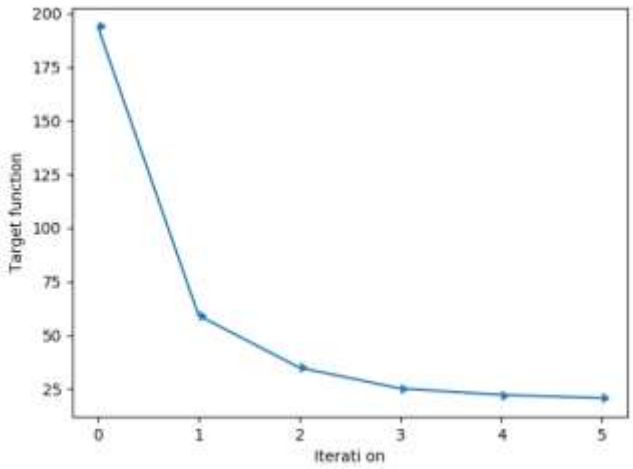

Fig.7. Five Cluster target function for the year of 1988 in Myanmar

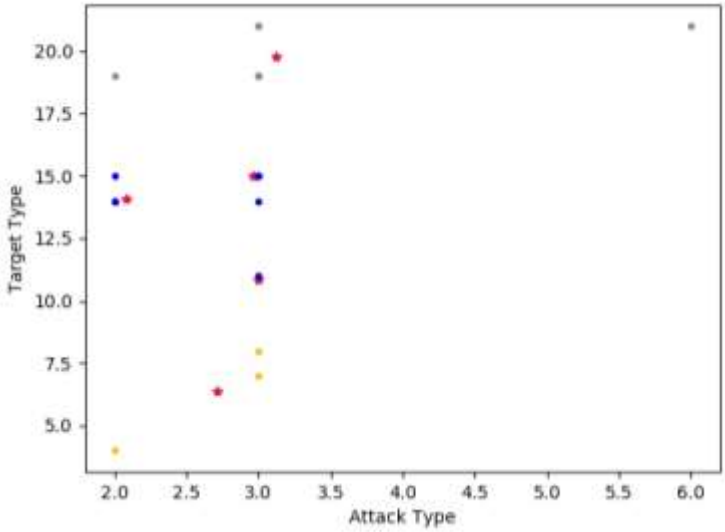

Fig.8. Five Clusters of Crime in Myanmar for the year of 1988

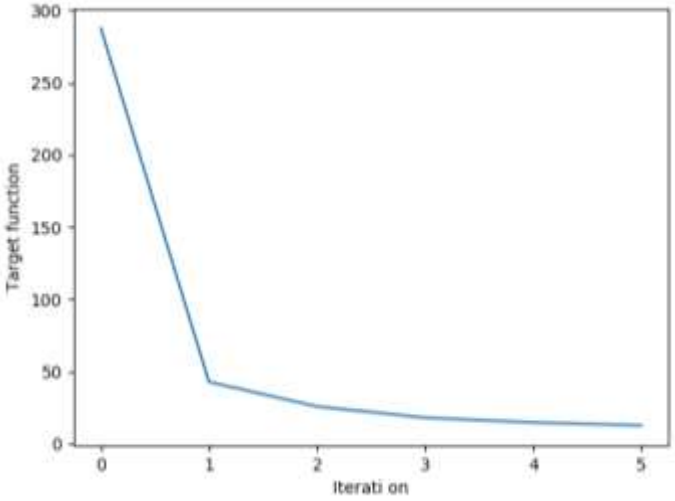

Fig.9. Six Cluster target function for the year of 1988 in Myanmar

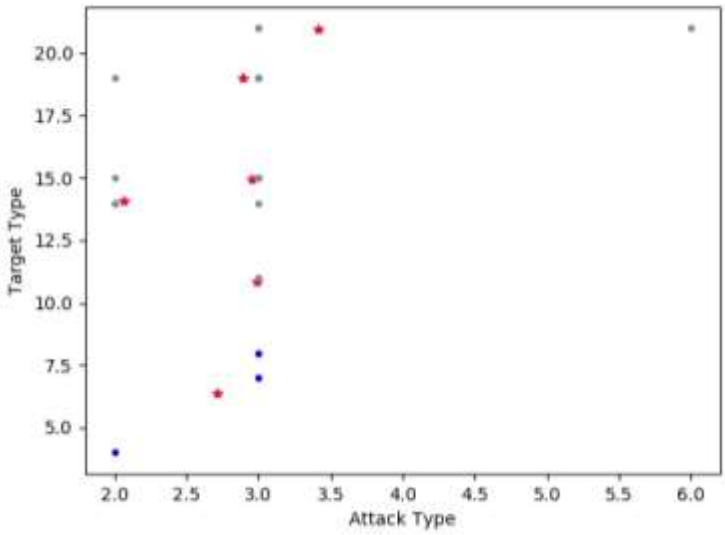

Fig.10. Six Clusters of Crime in Myanmar for the year of 1988

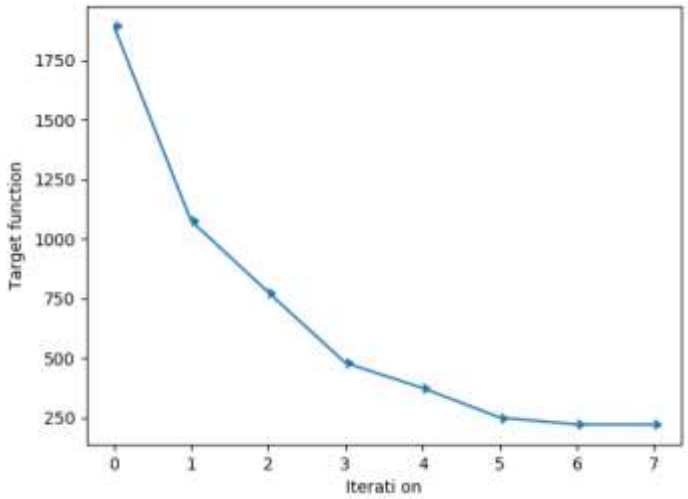

Fig. 11. Target Function of Three Crime Clusters happened in Myanmar for the periods from 1988 to 1992.

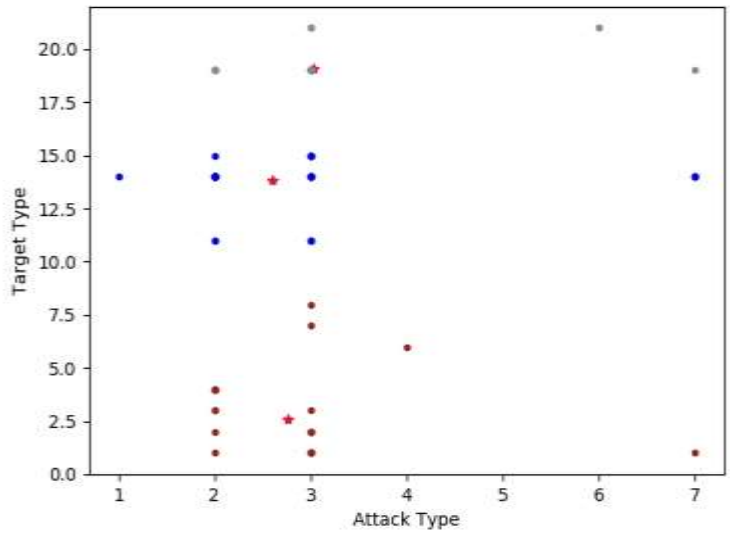

Fig. 12. Three Crime Clusters happened in Myanmar for the period from 1988 to 1992

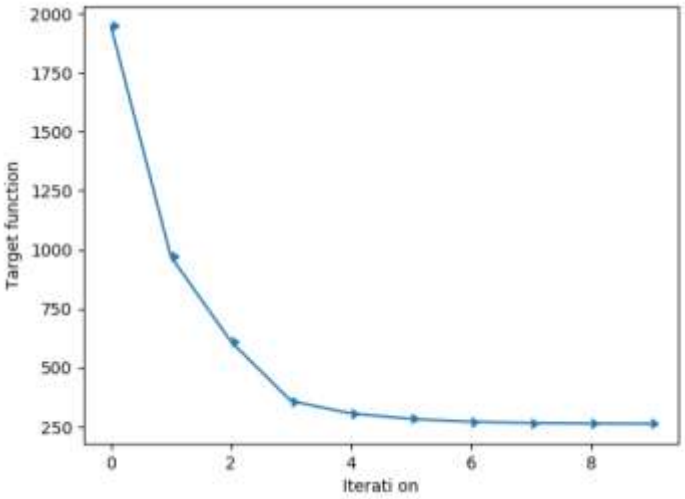

Fig. 13. Target Function of Three Crime Clusters happened in Myanmar for the Years of 1994, 1995, 1996, 1997, 1999.

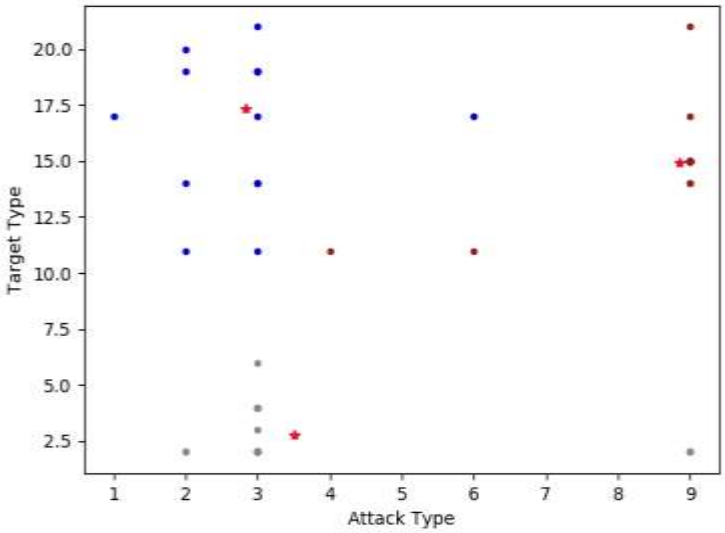

Fig. 14. Three Crime Clusters happened in Myanmar for the Years of 1994, 1995, 1996, 1997, 1999.

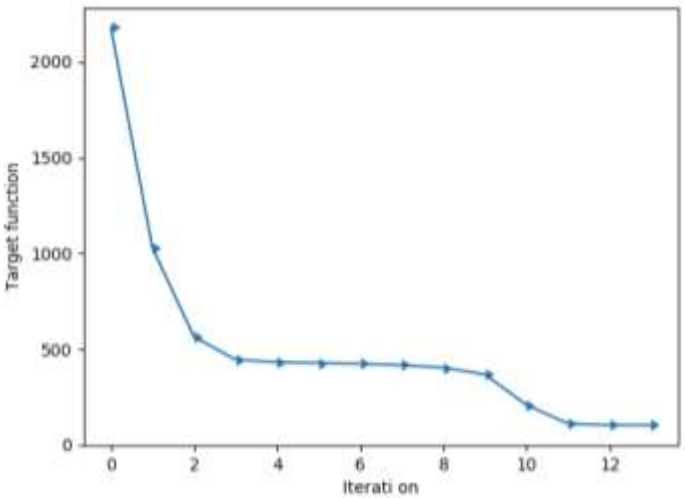

Fig. 15. Target Function of Three Crime Clusters happened in Myanmar for the period of 2001 to 2010.

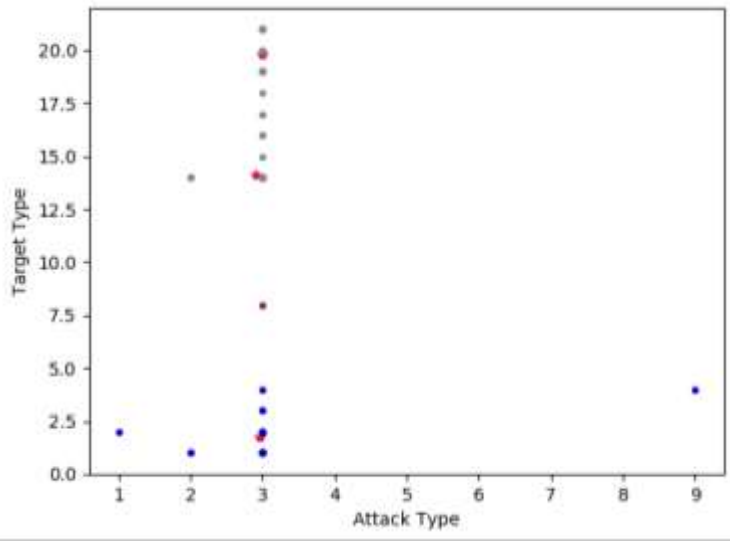

Fig. 16. Three Crime Clusters happened in Myanmar for the period of 2001 to 2010.

4) Methodology

We did our experiment using open source Python 3.6.8 and Anaconda1.9.7 distribution. As the challenges of our experiments, there are a lot of noisy data in the dataset. Therefore, firstly, we cleaned the data in order to be ready for our next step implementation.

IV. Conclusion

In this paper, we have discovered the criminal patterns of Myanmar according to one year, the group of the years and the period of the years in accordance with the place where the crimes usually happened and the types of the crimes. Furthermore, we explained Fuzzy clustering algorithm and their differences with hard clustering algorithm. In addition, we described its essentialities and the applications of fuzzy clustering algorithm in these days. And, at last, we discussed the crime patterns which have been discovered in Myanmar from our experiments.

V. Acknowledgement

First and foremost, we would like to thanks to our teachers who have supported us to become today. Secondly, we are especially grateful to our universities for supporting the facilities that is required for the implementation. Thirdly, we want to express our gratitude to our parents and families who had encouraged us with their full supports and inspirations. Finally, we wish to extend our thanks to all the associated person for their guidance to be successful in this work.

References

[1] H. Singh, M. M. Gupta, T. Meitzler, Z.-G. Hou, K. K. Garg, A. M. Solo, and L. A. Zadeh, “Real-life applications of fuzzy logic,” Advances in Fuzzy
Systems, vol. 2013, 2013.

[2] K. N. Win, J. Chen, G. Xiao, Y. Chen, and P. Fournier-Viger, “A parallel crime activity clustering algorithm based on apache spark cloud computing
platform.”

[3] K. N. Win, J. Chen, Y. Chen, and P. Fournier-Viger, “Pcpd: A parallel crime pattern discovery system for large-scale spatiotemporal data based on fuzzy
clustering,” International Journal of Fuzzy Systems, pp. 1–14, 2019.

[4] U. of Maryland, “Global terrorism database (gtd),” http://www.start.umd.edu/gtd.

[5] POLITECNICO, "Fuzzy Clustering Tutorial" https://home.deib.polimi.it/matteucc/Clustering/tutorial_html/cmeans.html.
[6] G. Mohler and M. Short, GEOGRAPHIC PROFILING FROM KINETIC MODELS OF CRIMINAL BEHAVIOR, SIAM J. APPL. MATH, vol. 72, no. 1, pp. 163180, 2012.
[7] G. Mohler, Marked point process hotspot maps for homicide and gun crime prediction in Chicago, Int. J. Forecast., vol. 30, no. 3, pp. 491497, 2014.
[8] A. Murray and T. Grubesic, Exploring Spatial Patterns of Crime Using Non-hierarchical Cluster Analysis, in Crime Modeling and Mapping Using Geospatial Technologies SE - 5, vol. 8, M. Leitner, Ed. Springer Netherlands, 2013, pp. 105124.
[9] S. J. Rey, E. a. Mack, and J. Koschinsky, Exploratory Space-Time Analysis of Burglary Patterns, J. Quant. Criminol., vol. 28, no. 3, pp. 509531, 2012.
[10] J. H. Ratcliffe, CrimeMapping: Spatial and Temporal Challenges, in Handbook of quantitative criminology, Springer Sciences, 2010.
[11] Christopher R. Herrmann;The dynamics of robbery and violence hot spots, Springer-December 2015
[12] 2012–Construction of a Web-Based Crime Geointelligence Platform for Mexico City Public Safety.
[13] N. Malleson and M. a Andresen, Spatio-temporal crime hotspots and the ambient population, Crime Sci., vol. 4, no. 1,p. 10, 2015.
[14] T. H. Grubesic, On The Application of Fuzzy Clustering for Crime Hot Spot Detection, J. Quant. Criminol., vol. 22, no.1, pp. 77 105, 2006.
[15] Diego Mayorga, Miguel A. Melgarejo, Nelson Obregon,A Fuzzy Clustering Based Method For the Spatiotemporal Analysis of Criminal Patterns, 2016, IEEE
[16] LianDuan, Xinyue Ye, Tao Hu and XinyanZhu,Prediction of Suspect Location Based on Spatiotemporal Semantics, International Journal of Geo-Information, June 2017
[17] Hvistendahl, M. Crime forecasters.Science 2016, 353, 14841487. [CrossRef] [PubMed]
[18] Fuzzy C-Means (FCM) Clustering Algorithm: A Decade Review from 2000 to 2014.
[19] 2011–ICDM–Crime Forecasting Using Data Mining Techniques
[20] Berk, R. Algorithmic criminology.Secur.Inf. 2013, 2, 114. [CrossRef]
[21] 2015–TIFS– CISRI: A Crime Investigation System Using the Relative Importance of Information Spreaders in Networks Depicting Criminals Communications