IJCSE International Journal of Computer Sciences and Engineering Open Access
Research Paper Vol.-6, Issue-8, Aug 2018 E-ISSN: 2347-2693

Finding Patterns in Crime Against Women Using a Fuzzy Clustering Technique

S. Das[1*], A. Das[2], A.U. Islam[3]

[1,2,3]CSIT, Cotton University, Guwahati, India

[*]*Corresponding Author: ssaimm@rediffmail*.com , *+91-9435314744*

Available online at: www.ijcseonline.org

Accepted: 16/Aug/2018, Published: 31/Aug/2018

Abstract— In most of the recent works pertaining to crime analysis traditional hard clustering techniques are seen to be applied for obtaining the intensity of crime in a particular region. Such clustering techniques which are based on crisp set theory are unable to deal with partial belongingness and as a result it is not possible to find regions partially belonging to multiple clusters with different crime intensities. Keeping this limitation of hard clustering techniques in view we will apply a fuzzy clustering technique which can deal the situations pertaining to partial belongingness, on a dataset of crime against women to reveal some important patterns in it.

Keywords— fuzzy clustering, crime against women, YFCM, FCM, patterns.

I. INTRODUCTION

The incidences of crime are being proliferated day by day all over the world and therefore it is very much essential to find adequate ways to minimize such crimes. Different data mining techniques are being applied to find meaningful information from crime data. It is seen in most of the recent literature pertaining to crime analysis, K-Means clustering algorithm has been used for analyzing and finding patterns of crimes in different locations.

Hard clustering techniques, K-Means for example, are based on crisp set theory and as a result are unable to deal with partial belongingness of a region in multiple clusters with different crime intensities. Therefore, with such clustering techniques, it is not possible to predict the possibility of a region moving towards a cluster with higher crime rate. Keeping this limitation of a hard clustering technique in view we will apply a fuzzy clustering technique in our work. As the Yuan Fuzzy C-Means (YFCM) algorithm of Das and Baruah is already claimed to have enhanced performance than the Fuzzy C-Means (FCM) clustering algorithm of Bezdek, therefore, in this present work YFCM is considered to be more appropriate for finding patterns of crimes on a dataset consisting of data of crime against women where there is every possibility of a region belonging partially to more than one clusters with different crime rates [1], [2].

Section-II includes the related work. Section-III consists of the methodology, algorithms involved in the present work. The application of YFCM in a dataset of crime against women has been explained in this section. Section-IV comprises of the results and discussions. Finally, Section-V consists of the conclusions and future scope.

II. RELATED WORK

Chau et. al. have applied an entity extraction technique to automatically identify person, address, vehicles and personal properties from police narrative reports [3]. De Bruin et. al. introduced a framework for crime trends using a new distance measure for comparing all individuals based on their profiles and then clustering them accordingly [4]. Tayal, Jain and Arora proposed a data mining approach for the design and implementation of crime detection and criminal identification in different cities of India. In their approach they used K-Means clustering algorithm for analyzing crime detection [5]. Alkhaibari and Chung applied different partitioning and agglomerative clustering algorithms in the Stop, Question and Risk Report database of New York Police Department (NYPD) for analyzing the locations of crime and reducing city crime rates [6]. Their analytic and visual results claimed and justified K-Means clustering algorithm to be the best among all the clustering algorithms used for their work. Chauhan and Sehgal reviewed thoroughly different data mining techniques and algorithms used for analyzing crimes[7]. Thota et. al. used K-Means clustering algorithm on a dataset consisting of criminal's record in India with an intention to find crime trends in different zones of states in India[8]. Aljrees et. al. applied a modified K-Means clustering algorithm in recognizing patterns of crimes[9]. They also found that the performance of their modified K-Means clustering algorithm was better than that of the normal K-Means clustering algorithm. With the advent of the concept of fuzzy set theory (FST) developed by Zadeh which particularly deals with the situations pertaining to non-probabilistic uncertainty, the traditional hard clustering technique, K-Means for example, has unlocked a new way of clustering known as fuzzy clustering, where a single object may belong exactly to one cluster or partially to more than one cluster depending on the membership value of that object [10] . The applications of FST in dealing with ambiguous problems where uncertainty prevails have been reflected in the works of Dewit [11], Lemaire [12], Ostaszewski [13]. Pardeshi and Patil have presented a thorough survey on fuzzy logic and rule mining [14]. Pattern recognition is a field whose objective is to assign an object or event to one of a number of categories, based on features derived to emphasize commonalities. Zheng and He reviewed the general processing steps of pattern recognition where they have discussed several methods used for the steps of pattern recognition such as Principal Component Analysis (PCA) in feature extraction, Support Vector Machines (SVM) in classification and so forth [15]. SenthilSelvi and Parimala designed a suitable model where they proposed to use PCA and Kernel Principal Component Analysis (KPCA) to improve the clustering efficiency and quality [16]. Derring and Ostaszewski have explained in their research work a method of pattern recognition for risk and claim classification [17]. They have also made similar application to classify claims with regard to their suspected fraud content.Das has tried the FCM algorithm of Bezdek with three different distances namely Euclidean distance, Canberra distance and Hamming distance which revealed that out of the three distances, the algorithm produces the result fastest as well as the most expected when Euclidean distance is considered and the slowest as well as the least expected when Canberra distance is considered [18]. Das and Baruah have shown the application of the FCM algorithm of Bezdek in vehicular pollution where it is justified that adopting a Fuzzy clustering approach on a dataset consisting of data of vehicular pollution is more beneficial over that of the traditional approach[19]. Das and Baruah have tried to make a comparison between the FCM clustering algorithm of Bezdek and the Gustafson-Kessel (GK) clustering algorithm of Gustafson and Kessel based on three (03) validity measures and found that the overall performance of FCM is better than that of GK on the dataset used for analysis [20], [21]. Although in most of the situations it is evident that the FCM clustering algorithm performs better than other fuzzy clustering algorithms, due to the random initialization of the membership values of the feature vectors the performance of the FCM clustering algorithm of Bezdek varies significantly in its different executions. Yager and Filev proposed a simple and effective method, called the Mountain Method, for estimating the number and initial location of cluster centers [22]. Although this method, unlike the FCM clustering algorithm of Bezdek, did not depend on any randomly initialized membership value to estimate its initial location of cluster centers, the problem with this method, Mountain Method, was that its computation grew exponentially with the dimension of the problem. Chiu developed a new method called Subtractive Clustering (SC) with which he could solve this problem by using data points as the candidates for cluster centres, instead of grid points as in the Mountain Clustering [23]. Yuan et al. proposed a systematic method for finding the initial centroids where there is no scope of randomness and therefore the centroids obtained by this method are found to be consistent [24]. To remove the effect of randomness from the FCM clustering algorithm Das and Baruah proposed a new algorithm, Subtractive Fuzzy C-Means (SUBFCM)[25]. In this algorithm the SC algorithm of Chiu has been used as a pre-processor to the FCM clustering algorithm of Bezdek where it has been tried to overcome two key limitations of it - the inconsistency due to randomness and the inability to handle the situations where the number of clusters is not predetermined. Although SUBFCM could overcome two key limitations of the FCM clustering algorithm of Bezdek, its performance has not been found better than that of FCM in some executions. Keeping that in view Das and Baruah proposed another new algorithm, YFCM, where the algorithm of Yuan et al. has been applied as a pre-processor to the FCM clustering algorithm of Bezdek to remove the effect of randomness as well as to achieve better performance.

III. METHODOLOGY

In this sections we will include the algorithms involved in this present work. We have also provided an application of YFCM algorithm in a dataset of crime against women.

A. The FCM algorithm of Bezdek

Step 1: Choose the number of clusters, $c, 2 \leq c < n$, where n is the total number of feature vectors. Choose $m, 1 \leq m < \alpha$. Define the vector norm $\| \;\; \|$ (generally defined by the Euclidean distance, the mathematical formula of which has also been provided in Equation-1

$$\| x_k - v_i \| = \sqrt{\sum_{j=1}^{p} (x_{kj} - v_{ij})^2} \tag{1}$$

where x_{kj} is the jth feature of the kth feature vector, for $k = 1,2, \dots \dots, n$; $j = 1,2, \dots ., p$ and v_{ij}, j-dimensional centre of the ith cluster, for $i = 1,2, \dots \dots, c$; $j = 1,2, \dots ., p$; n, p and c denote the total number of feature vector , features in each feature vector and total number of clusters respectively.

Choose the initial fuzzy partition

$$U^{(0)} = [\mu_{s_i}{}^{(0)}(x_k)]_{1\le i\le c, 1\le k\le n} \quad (2)$$

Choose a parameter $\in > 0$ (this will tell us when to stop the iteration). Set the iteration counting parameter l equal to 0.

Step 2: Calculate the fuzzy cluster centers $\{v_i^{(l)}\}_{i=1,2,\ldots,c}$ given by the following formula

$$v_i{}^{(l)} = \frac{\sum_{k=1}^{n} (\mu_{s_i}{}^{(l)}(x_k))^m x_k}{\sum_{k=1}^{n} (\mu_{s_i}{}^{(l)}(x_k))^m} \quad (3)$$

for i = 1, 2 , ….. c; k = 1, 2, , ….. n.

Step 3: Calculate the new partition matrix (i.e. membership matrix)

$$U^{(l+1)} = [\mu_{s_i}{}^{(l+1)}(x_k)]_{1\le i\le c, 1\le k\le n},$$

where $$\mu_{s_i}{}^{(l+1)}(x_k) = \frac{1}{\sum_{j=1}^{c} \left(\frac{\| x_k - v_i{}^{(l)} \|}{\| x_k - v_j{}^{(l)} \|} \right)^{\frac{2}{m-1}}} \quad (4)$$

for i = 1,2, … … . . , c and k = 1,2, … … . . , n.

If $x_k = v_i^{(l)}$, formula (2.4) cannot be used. In this case the membership function is

$$\mu_{s_i}{}^{(l+1)}(x_k) = \begin{cases} 1 & if k=i \\ 0 & if k\ne i, i=1,2,\ldots,c \end{cases} \quad (5)$$

Step 4: Calculate $\Delta = \| U^{(l+1)} - U^{(l)} \|$ (6)

If $\Delta > \in$, repeat steps 2, 3 and 4. Otherwise, stop at some iteration count l^*. To make the result operational the fifth step had been introduced by Derring and Ostaszewski.

Step 5: The final fuzzy matrix U^{l^*} is structured for operational use by means of the normalized α -cut, for some $0 < \alpha < 1$. All membership values less than α are replaced with zero and the function is renormalized (sums to one) to preserve partition the condition.

B. The Algorithm of Yuan et al.

The following steps illustrate the algorithm of Yuan et al.

Step 1: Set m =1;

Step 2: Compute the distance between each data point and all other data points in the dataset X;

Step 3:Find the closet pair of data points from the dataset X and form a data point set A_m ($1 \le m \le c$, c is the number of clusters) which contains these two data points, delete these two data points from the dataset X;

Step 4: Find the data point in X that is the closet to the data point set A_m, add it to A_m and delete it from X;

Step 5: Repeat Step-4 until the number of data points in A_m reaches 0.75*(n/c); (where .75 is a multiplication factor (MF))

Step 6: If m<c, then m = m+1, find another pair of data points from X between which the distance is shortest, form another data point set A_m and delete them from X, go to Step-4;

Step 7: For each data point set A_m ($1 \le m \le c$) find the arithmetic mean of the vectors of data points in A_m, these means will be the initial centroids.

C. The YFCM Algorithm of Das and Baruah

Step 1: Normalize the data points so that these are bounded by a hypercube.

Step 2: Find the initial centroids by the algorithm of Yuan et al. (see Section-B)

Step 3: Take these centroids obtained in Step-2 as inputs to FCM.

Step 4: Calculate the membership values of the objects by FCM (See Equation-4)

Step 5: Calculate the value of Δ. (See Equation-6)

if ($\Delta > \in$) then repeat Steps-6, 7 and 5

else stop iteration.

Step 6: Update the centroids by FCM. (See Equation-3)

Step 7: Update the membership values by FCM. (See Equation-4)

The pictorial representation of the above algorithm has been provided in Figure 1.

C. An application of YFCM algorithm for finding patterns in crime against women.

In the present scenario, providing security to the society has become a matter of high concern all over the world especially to those people who are involved in reducing crime incidences. It is essential to analyze crime data so that important patterns of crime can be revealed and adequate measures can be adopted in advance to minimize such crimes. Broadly, crimes can be classified into certain categories namely property crime, violent crime, crime against women and children, traffic violation, cyber crime and others. The dataset used in this section consists of data of crime against women of fifty (50) different cities (feature vectors) each of which is comprising of three (03) attributes (features) namely Dowry Death (DD), Rape (RP) and Cruelty by Husband and his relatives (CH) (see Table-1). Three (03) predefined clusters namely Low Crime Rate (C1), Average Crime Rate (C2) and High Crime Rate (C3) have been considered (see Table-2) for identifying crime intensities in each of these fifty(50) different cities. We have executed both FCM algorithm of Bezdek and YFCM algorithm of Das and Baruah on the dataset of crime data and made a comparison of these two algorithms based on three

(03) validity measures PC, CE and PI (see Equations- 7,8 and 9). The number of iterations to reach the final clusters has also been considered for the comparison. We have applied the traditional K-Means algorithm on the dataset of our present work (see Table-1) and recorded the result revealed by it in Table-4.

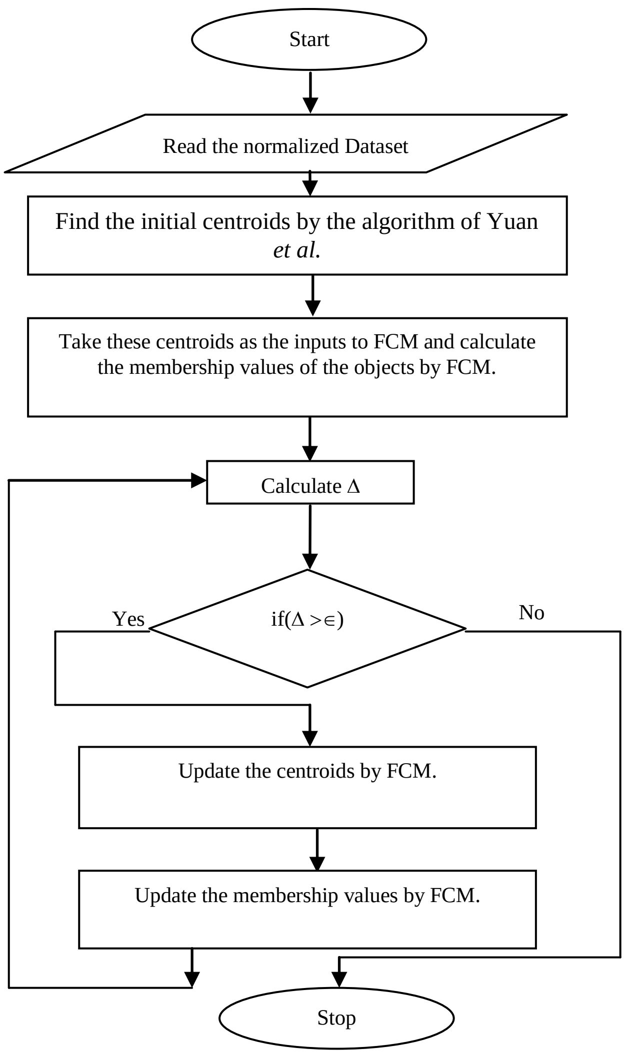

Figure 1. Flowchart of YFCM algorithm

The mathematical expressions of the three validity measures used in the present work have been given in the following.

Partition Coefficient (PC): measures the overlapping between clusters.

$$PC(c) = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}(\mu_{ij})^2 \quad (7)$$

Clustering Entropy (CE): measures the fuzziness of the cluster partition

$$CE(c) = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}\mu_{ij}\log(\mu_{ij}) \quad (8)$$

Partition Index (PI): is the ratio of the sum of the compactness and separation of the clusters.

$$PI(c) = \sum_{i=1}^{c}\frac{\sum_{j=1}^{n}(\mu_{ij})^m \| x_j - v_i \|^2}{n_i \sum_{k=1}^{c} \| v_k - v_i \|^2} \quad (9)$$

Table 1. The dataset of crime against women

City_ID	DD	RP	CH	City_ID	DD	RP	CH
1	0.5	1	4	**26**	0.3	0.7	3
2	0.4	0.9	3	**27**	0.2	0.7	3
3	0.5	1	5	**28**	0.1	0.6	1
4	0.3	0.9	2	**29**	0.8	1.5	5.9
5	0.6	1	5.8	**30**	0.5	1	6
6	0.2	0.7	2	**31**	0.6	2	5
7	0.1	0.7	2	**32**	0.7	2	5.9
8	0.3	0.8	1	**33**	0.9	3	7
9	0.4	0.9	2	**34**	2	5	10
10	0.5	0.9	4	**35**	0.2	0.7	2
11	0.7	2	5.8	**36**	0.7	1.5	5
12	0.9	2.5	8	**37**	0.5	1	5
13	0.8	2	6	**38**	1	2.5	8
14	0.1	0.5	1	**39**	0.4	0.8	4
15	0.2	0.4	1	**40**	0.2	0.3	2
16	0.4	0.7	1	**41**	0.1	0.2	2
17	0.6	2	5	**42**	0.6	2	6
18	1	2.5	8	**43**	0.9	3	7
19	0.9	2	7	**44**	0.7	2	5.9
20	0.8	1.5	9	**45**	0.2	1	1
21	0.6	1	6	**46**	0.5	0.8	4
22	0.7	1.5	5.9	**47**	0.4	0.6	4
23	0.5	1	5	**48**	0.3	0.7	3
24	0.4	0.8	4	**49**	0.2	0.5	1
25	0.2	0.6	1	**50**	0.1	0.2	1
DD: Dowry Death RP: Rape CH: Cruelty byHusband and his relatives							

Table 2. Predefined clusters of crime against women
Source: National Crime Records Bureau (NCRP)

Cluster Name	DD (per 100,000)	RP (per 100,000)	CH (per 100,000)
Low Crime Rate (C1)	Below 0.5	Below 1	Below 5
Average Crime Rate (C2)	0.5-0.9	1-3	5-6
High Crime Rate (C3)	Above .9	Above 3	Above 6
DD: Dowry Death RP: Rape CH: Cruelty by Husband and his relatives			

IV. RESULTS AND DISCUSSION

In this section we have recorded our numerical findings. We have also tried to provide an analysis of the results revealed by our experiments. The values of different validity measures of FCM and YFCM algorithms have been recorded in Table-3 and the graphical representations of the same have been provided in Figures-2 and 3.

Table 3.The values of different validity measures of FCM (the best out of ten different executions) and YFCM algorithms (considering MF=.65) on the dataset of crime against women.

Algorithm	N	C	Itn	PC	CE	PI
FCM	50	3	9	0.74763	0.52529	0.03842
YFCM	50	3	6	0.75798	0.51483	0.02726

N: the size of the dataset, C: the no. of predefined clusters, Itn: the no. of iterations to reach the final clusters.

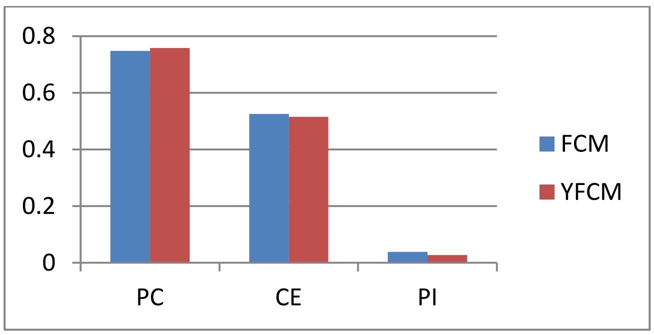

Figure 2.Comparisons of the values of different validity measures of FCM and YFCM algorithms when applied on the dataset of crime against women.

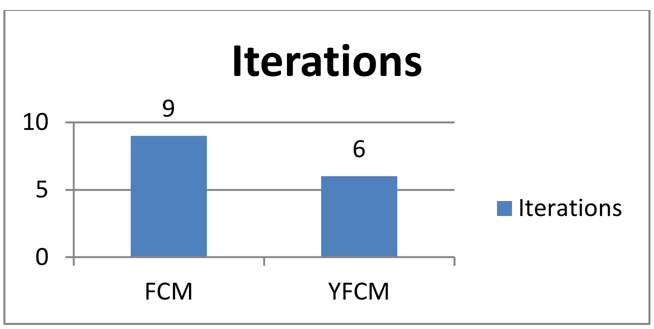

Figure 3. Comparisons of the no. of iterations of FCM and YFCM algorithms when applied on the dataset of crime against women.

From Figure-2 & 3 it is clear that the performance of YFCM algorithm is better than that of FCM on the dataset used in this present work. Therefore, we have considered YFCM algorithm to be more appropriate as compared to FCM to obtain the membership values (full and partial) of different cities in the predefined clusters (see Table-5).

Table 4. Cities belonging to three different clusters revealed by K-Means algorithm

Cluster Name	City_IDs with full membership values
Low Crime Rate (C1)	2,4,6,7,8,9,10,14,15,16,23, 24,25,26,27,28,35,39, 40,41,45,56,47,48,49,50
Average Crime Rate (C2)	1,3,5,11,12,13,17,19,22,29,30,31,32,36,37,44
High Crime Rate (C3)	18,20,33,34,38,42,43

In Table-4 we see that each of the fifty (50) cities belongs exactly any one of the three (03) predefined clusters of crimes. Here, we do not find a single city which belongs partially to more than one cluster with different crime rates. But such outcomes are quite irrational, because due to the availability of three different features of crimes in a city its possibility of belonging to exactly any one of the three predefined clusters of crimes is very less. Rather, there is every possibility of a city belonging partially to more than one cluster of crimes. In the results revealed by the K-Means clustering algorithm (see Table-4) this situation is not seen to be dealt with and therefore, is a major limitation of a clustering algorithm, k-Means for example, which is based on the crisp set theory.

Table 5. Membership values of the cities revealed by YFCM in different predefined clusters of crime against women.

Cluster Name	City_IDs with full membership values	City_IDs with partial membership values
Low Crime Rate (C1)	2,4,6,7,8,9,14,15,16, 24,25,26,27,28,35,39, 40,41,47,48,49,50	1(.2589), 10(.7812), 23(.6239), 30(.2109), 45(.7923), 46(.7895)
Average Crime Rate (C2)	5,11,17,22,29,31,32, 36,37,44	1(.7411), 3(.7312), 10(.2188), 12(.6818), 13(.6912), 19(.6572), 20(.2747), 21(.5032), 23(.3761), 30(.7891), 45(.2077),

		46(.2105)
High Crime Rate (C3)	18,33,34,38,42,43	3(.2688), 12(.3182), 13(.3088), 19(.3428), 20(.7253), 21(.4968)

On the other hand, as YFCM clustering algorithm is a fuzzy one, we can see some cities (see column-3 of Table-5) which partially belong to more than one predefined clusters of crimes. The key limitation seen in Table-4 is that there is no scope of generating prior alerts to those cities which have a tendency to move towards a cluster with higher crime rate. This limitation of K-Means can be dealt with YFCM (see column-3 of Table-5) as it is possible to identify those cities for which there is a possibility of generating prior alerts for taking adequate measures in advance. In other words, we can say that the advantage of partial belongingness is that we can obtain the possibility of shifting a city with partial membership value to a cluster with less crime rate. In Table-6, we have shown the city_IDs with respective possibility of shifting to a cluster with less crime rate. A graphical representation of the same has been shown in Figure-4. This partial belongingness of the cities to the clusters of crimes indicates the percentage of care has to be taken to a particular city for reducing its crime rate. For example, in Table-5 the cities with city_ID s 1, 2 and 10 belong to the cluster of Low Crime Rate (C1) with membership values .2589, 1 and .7812 respectively. This membership values indicate that more care has to be taken for city with city_ID 2, followed by cities with city_IDs 10 and 1 respectively for making these cities free from crimes. In the absence of partial belongingness setting priority of care for each city is not possible. Similarly in Table-5 it is seen that the city with city_ID 3 is partially belonging to cluster of Average Crime Rate (73%) and that of High Crime Rate (27%). This is an indication that this city has a tendency (27%) to move towards a cluster of higher crime rate. Therefore, if adequate measure is taken in advance, we can stop its further deterioration. That means there is a possibility (73%) of shifting this city from High Crime Rate to Average Crime Rate (see Table-6). Similar arguments are applicable for the other cities in Table-6.

Table 6. City_IDs with respective possibility of shifting to a cluster with less crime rate.

	cluster names where the cities are partially present		
CITY_ID	cluster with less crime rate	cluster with more crime rate	possibility of shifting to a cluster with less crime rate
1	Low Crime Rate	Average Crime Rate	26%
3	Average Crime Rate	High crime Rate	73%
10	Low Crime Rate	Average Crime Rate	78%
12	Average Crime Rate	High crime Rate	68%
13	Average Crime Rate	High crime Rate	69%
19	Average Crime Rate	High crime Rate	66%
20	Average Crime Rate	High crime Rate	27%
21	Average Crime Rate	High crime Rate	50%
23	Low Crime Rate	Average Crime Rate	62%
30	Low Crime Rate	Average Crime Rate	21%
45	Low Crime Rate	Average Crime Rate	79%
46	Low Crime Rate	Average Crime Rate	79%

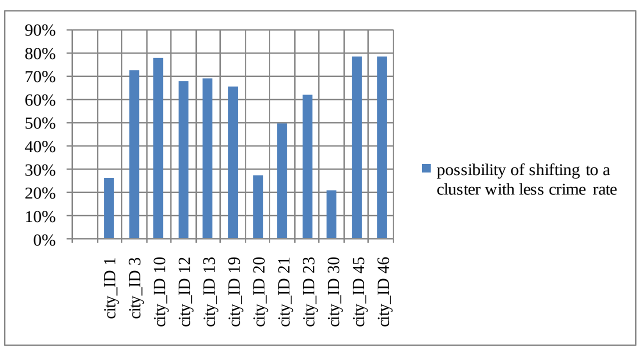

Figure 4. City_ID s with respective possibility of shifting to a cluster with less crime rate.

V. CONCLUSION AND FUTURE SCOPE

In this work we have executed both YFCM and FCM algorithms on the same dataset and found that YFCM algorithm performs better than FCM. we have justified that the application of a fuzzy clustering technique is more beneficial over a hard clustering technique on a dataset of crimes against women where there are objects with partial belongingness. We have shown the key limitations of K-Means algorithm in dealing situations pertaining to partial belongingness and therefore, it is not suitable for our present dataset where due to the availability of three different features there is every possibility of a city belonging partially to multiple clusters with different crime rates. We have shown that by applying a fuzzy clustering algorithm, YFCM algorithm in this case, it is possible to identify those cities which have a tendency to move towards a cluster of higher crime rate. With such identifications it is also possible to obtain the possibility of shifting a city from higher crime rate to lesser crime rate. This possibility will certainly be immensely helpful for adopting adequate measures in advance in reducing crimes.

It is hoped that such a fuzzy clustering technique, YFCM in this case, which has high as well as consistent performance will certainly encourage the researchers associated in the fields of data mining and pattern recognition to carry out their research work for revealing important patterns from data where partial belongingness prevails.

REFERENCES

[1] S. Das and H. K. Baruah., "A New Method to Remove Dependence of Fuzzy C-Means Clustering Technique on Random Initialization", International Journal of Research in Advent Technology, Vol. 2, Issue.1, pp.322-330, 2014.

[2] J. C. Bezdek, , "Pattern Recognition with Fuzzy Objective Function Algorithms", Plenum Press, New York, 1981.

[3] M. Chau, J. Xu, and H. Chen,"Extracting meaningful entities from police narrative reports", Proceedings of the National Conference for Digital Government Research, Los Angeles, California, USA, pp.1-5, 2002.

[4] J.S. De Bruin, T.K Cocx., W.A Kosters., J Laros. and J.N Kok., "Data mining approaches to criminal carrer analysis", Proceedings of the Sixth International Conference on Data Mining, pp.171-177, 2006.

[5] D. K. Tayal, A Jain. and S. Arora, "Crime detection and criminal identification in India using data mining techniques", AI & SOCIET, Vol.30, Issue. 1, pp.117–127, 2015.

[6] A. Alkhaibari and P. T. Chung, "Cluster analysis for reducing city crime rates", Systems, Applications and Technology Conference (LISAT), Long Island, NY, USA, 2017.

[7] C. Chauhan and S Sehgal., "A review: Crime analysis using data mining techniques and algorithms", International Conference on Computing, Communication and Automation (ICCCA), Uttar Pradesh,India,2017.

[8] L. S Thota., M. Alalyan, A. A Khalid., F. Fathima, S. B. Changalasetty. and M, Shiblee. "Cluster based zoning of crime info", 2nd International Conference on Anti-Cyber Crimes (ICACC), 2017.

[9] T. Aljrees, D. Shi, D. Windridge and W. Wong, "Criminal pattern identification based on modified K-means clustering", IEEE International Conference on Machine Learning and Cybernetics (ICMLC),Jeju, South Korea, 2017.

[10] L. A., Zadeh "Fuzzy Sets", Information and Control, Vol. 8, Issue. 3, pp.338-353, 1965.

[11] G. W. Dewit, "Underwriting and Uncertainty", Insurance: Mathematics and Economics, Vol. 1, Issue. 4, pp. 277-285, 1982.

[12] J. Lemiare, "Fuzzy Insurance", Astin Bulletin, Vol. 20, Issue. 1, pp. 33-55, 1990.

[13] K. Ostaszewski, "An Investigation into Possible Applications of Fuzzy Sets Methods in Actuarial Science", Society of Actuaries, Schaumburg, Illinois, 1993.

[14] P. Pardeshi and U. Patil, " Fuzzy Association Rule Mining-A Survey", International Journal of Scientific Research in Computer Science and Engineering, Vol.5, Issue.6,pp.13-18, 2017.

[15] L. Zheng and X. He, "Classification Techniques in Pattern Recognition", Conference Proceedings of 13th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision, ISBN 80-903100-8-7 WSCG, Science Press ,Australia, pp. 77-88, 2005.

[16] T. SenthilSelvi and R. Parimala, "Improving Clustering Accuracy using Feature Extraction Method", International Journal of Scientific Research in Computer Science and Engineering, Vol.6, Issue.2, pp.15-19, 2018.

[17] R. A. Derrig and K. M Ostaszewski., "Fuzzy Techniques of Pattern Recognition in Risk and Claim Classification", Journal of Risk and Insurance, Vol. 62, Issue. 3, pp.447-482, 1995.

[18] S. Das, "Pattern Recognition using the Fuzzy c-means Technique", International Journal of Energy, Information and Communications, Vol. 4, Issue 1, pp.1-14,2013.

[19] S. Das and H. K. Baruah, "Application of Fuzzy C-Means Clustering Technique in Vehicular Pollution", Journal of Process Management – New Technologies, Vol. 1, Issue. 3, pp.96-107, 2013.

[20] D. E Gustafson. and W. C. Kessel, "Fuzzy clustering with a fuzzy covariance matrix", Proc. IEEE CDC, San Diego, CA, USA, pp.761- 766, 1979.

[21] S. Das and H. K. Baruah, "A Comparison of Two Fuzzy Clustering Techniques", Journal of Process Management – New Technologies, Vol. 1, Issue. 4, pp.1-15,2013.

[22] R.R. Yager and D.P. Filev "Approximate Clustering Via the Mountain Method", Tech. Report #MII-1305, Machine Intelligence Institute, Iona College, New Rochelle, NY, 1992.

[23] S.L Chiu., "Fuzzy Model Identification Based on Cluster Estimation", Journal of Intelligent and Fuzzy Systems, Vol.2, pp.267-278,1994.

[24] F. Yuan, Z.H Meng., H.X. Zhang and C.R Dong., "A New Algorithm to Get the Initial Centroids", Proc. of the 3rd International Conference on Machine Learning and Cybernetics, pp.26-29, 2004.

[25] S. Das and H. K. Baruah, "An Approach to Remove the Effect of Random Initialization from Fuzzy C-Means Clustering Technique", Journal of Process Management – New Technologies, Vol. 2, Issue. 1, pp.23-30,2014.

Authors Profile

S Das completed his Masters in Mathematics from Gauhati University of Assam, India. He has obtained his Ph.D. in Computer Science and Engineering in 2017 from Gauhati University and currently working as Assistant Professor in Department of CSIT, Cotton University of Assam, India since 2008. He has published more than 10 research papers in reputed international journals. He has 43 citations of his research work. His main research work focuses on Pattern Recognition, Fuzzy Logic, Fuzzy Clustering, Data Mining, and Computational Mathematics based education. He has 15 years of teaching experience and 8 years of Research Experience.

A Das completed his Masters in Mathematics and Computer Science from Gauhati University of Assam, India. He has obtained his Ph.D. in Computer Science and Engineering in 2015 from Gauhati University and currently working as Assistant Professor in Department of CSIT, Cotton University of Assam, India since 2007. He is a member of Assam Academy of Mathematics since 2005, International Association of Engineers (IAENG) since 2013 and CSI since February, 2018. He has published more than 10 research papers in reputed international journals and conferences including Springer . His main research work focuses on Network Security, Cloud Security and Privacy, Data Mining, IoT and Computational Intelligence based education. He has 15 years of teaching experience and 8 years of Research Experience.

A.U. Islam completed his Masters in Computer Application from Dibrugarh University of Assam, India. He has obtained his Ph.D. in Computer Science and Engineering in 2018 from Gauhati University and currently working as Assistant Professor in Department of CSIT, Cotton University of Assam, India since 2008.

He has published multiple research papers in reputed international journals including Scopus and ACM. His main research work focuses on Graph Theory and Data Mining based education. He has 10 years of teaching experience and 7 years of Research Experience.