Volume 6, Issue 4, April 2016 ISSN: 2277 128X
International Journal of Advanced Research in Computer Science and Software Engineering
Research Paper
Available online at: www.ijarcsse.com

A Review of Different Clustering Techniques in Criminal Profiling

Adeyiga J.A
Computer Science Department &
Bells University of Technology,
Ota, Ogun State, Nigeria

Bello A.O
Information Technology Department &
Bells University of Technology,
Ota, Ogun State, Nigeria

Abstract*— *Finding criminals and solving crime is no easy task and have been the prerogative of the law enforcement agencies. Several profiling systems have been developed to assist the law enforcement agencies in solving crimes but most of them lacked essential components which would have made the quality of their output more in tune with reality. This research work therefore aims to review several clustering techniques that has been developed for Criminal Profiling. The aim is to propose a better clustering technique (fuzzy clustering) for criminal profiling in order to provide investigators with rich sources of intelligence and information which can be used to predict and prevent criminal activity.

Keywords*— *Profiling, Criminal, Fuzzy Clustering, Crime.

I. INTRODUCTION

Crime has become very sophisticated in Nigeria and fighting crime with theold traditional methods do not seem to yield productive desired results. It is therefore imperative to apply the modern means of solving and investigating crime such as; the use of clustering algorithm in analyzing crime data.Crime analysis; a part of criminology, is a task that includes exploring and detecting crimes and their relationships with criminals [28]. It is defined as a set of systematic, analytical processes directed at providing timely and pertinent information relative to crime patterns and trend correlations to assist operational and administrative personnel in planning the deployment of resources for the prevention and suppression of criminal activities [38]. Studying crime data includes all the intricate details of a particular crime. More specifically, crime analysis is the breaking up of acts committed in violation of laws into their parts to find out their nature and reporting statements of these findings [38]. But due to increased crime rate which leads to an increase in both crime and criminal datasets there is a need to find better and more efficient ways to analyze these growing crime volumes.

A major challenge facing all law-enforcement and intelligence-gathering organizations is accurately and efficiently analyzing the growing volumes of crime data [8]. Another major challenge faced by the Nigerian law enforcement agencies is the lack of a central repository where all data collections concerning crimes and criminals are stored which possess a bigger problem, as there are many cases of data repetition, and as such it is difficult for law enforcers to see patterns in crimes during analysis. According to [22], detecting crime from data analysis can be difficult because the daily activities of criminals generate large amounts of data and stem from various formats. Again, the quality of data analysis depends greatly on the background knowledge of the analyst. However, intelligence and law enforcement agencies often are faced with the dilemma of having too much data, which in effect translates into too little value [18], notwithstanding the issues that results from the large volume of inconsistent, incomplete and incorrect data.

Furthermore, complex relationships between various crime attributes can be overlooked by human analysts. The development of a crime analysis tool has four steps; namely, data cleaning/ entity extraction, fuzzy clustering, classification and Knowledge representation. Various crime analysis tools using clustering mechanism have been developed. For example,[19] used a new distance measure and also used crime severity to cluster crimes (i.e. crimes were grouped according to major, minor and intermediate). Initially this approach tackled most of the problems at that time, but over time it gave rise to more complicated issues. For example, the distance measure between an innocent person and a onetime offender was relatively lower than the distance between a one-time offender and a minor criminal; this was influenced by the weighted value. One of the works [22] improved on was [19] by proposing several guidelines.

In many clustering algorithms, the spatial distance (usually Euclidean distance) among different instances is used to measure the similarity. This has proven to be successful in many applications [37]. It is intuitive that similar objects are closer together while objects from different groups are far from each other. This intuition, however, is not always true in more complicated applications, especially when the shape of data clusters is not limited to be spherical. The step from the well-known C-means clustering algorithm to fuzzy c-means algorithm and its vast number of sophisticated extensions and generalizations involves an additional clustering parameter, the so called fuzzifier [10].

Fuzzy clustering accepts the fact that the clusters or classes in the data are usually not completely well separated and thus assigns a membership degree between 0 and 1 for each cluster to every datum. Although the extension from deterministic (hard) to fuzzy clustering seems to be an obvious concept, it turns out that to actually obtain membership degrees between zero and one, it is necessary to introduce a so called fuzzifier in fuzzy clustering. The main purpose of

the fuzzifier is to control how much clusters are allowed to overlap. The concept of fuzzy connectedness which was first introduced by Rosenfeld in 1979 has shown that it allows the discovery of clusters of any shape and can detect not only the known types, but also their variants. It has been successfully applied to image processing, [15] generalized the concept of fuzzy connectedness. Since crime investigations deal with uncertainties (probability), the focus of this research work is to review different clustering algorithm techniques that have been used in profiling criminal.

II. PROFILING AND CRIMINALITY

Profiling is used to link a specific person or persons to a specific crime that have already occurred. Criminal profiling is the process of investigating and examining the criminal behavior in order to identify the type of person responsible [36]. A profile based on characteristics patterns or factors of uniqueness that distinguishes certain individuals from the general population. [32] explained that profiling criminals is important because not only does it help narrow down the pool of suspects in an investigation, it also helps to link related crimes, identify the potential for escalation of nuisance criminal behavior to more serious crimes, to focus on the investigation, etc.

Identification of patterns may help in profiling. For example, an attacker may use a specific mode of operation, and target specific victims. Profiling can help identify the victim types, and assist in the identification and installation of sufficient deterrents. In order to develop useful profiles of different offender categories, a large amount of data is required and in order to improve the reporting of crime, there is a need to increase the trust between the public and private sectors, which will result in reporting of crimes when they occur. This will allow researchers to more precisely identify whether or not any unique patterns and characteristics actually exist [3].

Profiling methods can harness data from both crime and business processes. Both also have a large information system component in data and digital image storage for later retrieval and analysis through information sharing, data matching and data mining techniques. Sometimes data is collected without permission or knowledge of the user e.g., CCTV, cookies, etc. Profiling may be scientific or non-scientific [16], singular or aggregative [29], and proactive or reactive [12]. Crime profiling of the offender/perpetrator can also follow the methodology of organized (above average traits e.g., intelligent quotient (IQ), competent) or disorganized i.e., below average IQ, inadequate [33]. Profiling techniques, aided by machine learning programs, can be classified as supervised or unsupervised. An example, of unsupervised learning includes profiling of superimposed frauds in the telecommunications and mobile phone sectors.

A study by [23] puts forward an initial pre-conception model of identity fraud profiling. In their model they highlight five sequential stages from perpetrator, mode of attack, target system, target entity, through to the victim and suggest that "as a profile carries limited value independently it is important to analyse the interactions of profiles, which may carry more value for identity fraud detection and prevention strategies" [23]. Building a comprehensive understanding of the way personal documentation and information is obtained by perpetrators is one of the best mitigation strategies for identity fraud – profiling does this.

A. Criminal Profiling

A profile is a psychological sketch of an offender. While criminalprofiling is the act of developing a psychological profile of an offender based on the state of the crime scene. It is most often done by a forensic psychologist, which is someone who has studied the criminal mind. This profile can then be used by law enforcement agencies to assist in apprehending the criminal [17]. Criminalprofiling is used not only to find potential offenders, but also to narrow down a list of offenders that has already been compiled by the police. Criminalprofiling has helped investigators to apprehend hundreds of criminals [7]. By assessing the patterns and motives of previous criminals, profiling allows investigators to fairly accurately predict the characteristics of current and future offenders, allowing killers and other perpetrators to be caught before they can continue on to other crimes. Criminal profiling works on the principle that each and every criminal regardless of the level or severity of their crime will work to a certain set of values.

III. THE CLUSTERING METHODS

Clustering is the technique that is used to group objects (crimes and criminals) without having predefined specifications for their attributes [38]. It ensures that the data objects are similar to one another within the same cluster and dissimilar to other objects in a different cluster. Clustering groups objects (crime and criminals) based on the information found in the data describing the objects or their relationships. The goal of clustering is that the objects in a group will be similar to one another and different from the objects in other groups. Clustering partitions a data set into several groups such that the similarity within a group is larger than that among different groups, the idea of clustering is similar to the way humans think. But we usually tend to have high volumes of data which is too cumbersome for the human mind to summarize into smaller groups. There is no golden standard or benchmark for the type of clustering technique to use, it depends on the goal of the researcher [22].

Cluster Analysis is also perceived as unsupervised learning because there is no pre-defined pattern in which objects should be grouped i.e. it learns by observation rather than examples; it deals with finding a *structure* in a collection of unlabeled data. A loose definition of clustering could be "the process of organizing objects into groups whose members are similar in some way" [14]. It is therefore a collection of objects, which are "similar" between them and are "dissimilar" to the objects belonging to other clusters. Any cluster should exhibit two main properties: low inter-class similarity and high intra-class similarity. Cluster analysis is used in a number of applications such as data analysis, image processing, fraud prediction, crime analysis etc. Since clusters can formally be seen as subsets of the data set, one can possible classify clustering methods according to whether the subsets are fuzzy or crisp (hard).

A. Hard clustering

These methods of clustering are based on the classical set theory, and require that an object either does or does not belong to a cluster [11]. It means partitioning the data into specified number of mutually exclusive subsets. That is, each document belongs to exactly one cluster, it only assigns a value of 1 or 0. Such as,

***1) Partitioning Algorithm*:** Construct various partitions and then evaluate them by some criterion. Partition techniques create a flat partitioning of the data points. If *K* is the desired number of clusters, then partitioning approaches typically find all *K* clusters at once. Given *k*, the *k-means* algorithm is implemented in four steps:

1. Partition objects into *k* nonempty subsets
2. Compute seed points as the centroids of the clusters of the current partition (the centroid is the center, i.e., *mean point*, of the cluster)
3. Assign each object to the cluster with the nearest seed point
4. Go back to Step 2, stop when no more new assignment[2].

The weakness of k-means is that it is applicable only when mean is defined, then what about categorical data?, the need to specify k;the number of clusters, in advance, also not suitable to discover clusters with non-convex shapes, the k-means algorithm is sensitive to outliers. A major example is K-Means clustering which aims to partition a dataset into *k* clusters in which each observation belongs to the cluster with the nearest mean.

***2) Density-Based*:** This is based on connectivity and density functions. In density-based clustering, clusters are defined as areas of higher density than the remainder of the data set. Objects in these sparse areas that are required to separate clusters - are usually considered to be noise and border points. The most popular density based clustering method is DBSCAN. In contrast to many newer methods, it features a well-defined cluster model called "density reach ability". Similar to linkage based clustering; it is based on connecting points within certain distance thresholds. However, it only connects points that satisfy a density criterion, in the original variant defined as a minimum number of other objects within this radius [20].

3) Hierarchical Algorithm:Hierarchical clustering builds a cluster hierarchy (a tree of clusters). Create a hierarchical decomposition of the set of data (or objects) based on a criteria and uses distance matrix as clustering criteria. This method does not require the number of clusters ***k*** as an input, but needs a termination condition.Connectivity based clustering, also known as hierarchical clustering, is based on the core idea of objects being more related to nearby objects than to objects farther away. As such, these algorithms connect "objects" to form "clusters" based on their distance. A cluster can be described largely by the maximum distance needed to connect parts of the cluster. At different distances, different clusters will be form, which can be represented using a dendrogram, which explains where the common name "hierarchical clustering" comes from: these algorithms do not provide a single partitioning of the data set, but instead provides an extensive hierarchy of clusters that merge with each other at certain distances[20]. Types include Agglomerative (bottom-up) which simply means starting with one point cluster and recursively merging two or more most similar cluster to one parent cluster until the termination and Divisive (Top-down) which starts with one cluster of all objects and recursively splitting each cluster until the termination criterion is reached.

B. Fuzzy Clustering

Fuzzy clustering also called soft clustering. Allow the objects to belong to several clusters simultaneously, with different degrees of membership. It is more natural than hard clustering because objects on the boundaries between several classes are not forced to fully belong to one of the classes, but rather are assigned membership degrees between 0 and 1 indicating their partial membership [11]. For example, the clusters or groups that are identified will be overlapping, meaning that one instance may fall into several clusters.

In fuzzy clustering we make a fuzzy partition of the data set. But to accommodate the introduction of fuzzy partitioning, the membership matrix (U) is randomly initialized according to equation 2.1. Each point has a degree of belonging to clusters, as in fuzzy logic, rather than belonging completely to just one cluster. Fuzzy clustering uses membership function in partition data set.

$$\sum_{i=1}^{c} \quad \mathrm{U}ij \quad = 1, \ \forall \quad j = 1 \ldots \mathrm{n} \qquad 2.1$$

This function is called membership function and its value is between 0 and 1 [11].

This can be applied to crime analysis because police officers commonly receive descriptions of suspects that are fuzzy in nature; the concepts of fuzzy variables will be introduced into criminal analysis to describe a fuzzy logic-based mathematical procedure that is capable of handling such variables. Fuzzy logic stipulates that an element can be a member of a given set in an uncertain manner. Unlike classical mathematical set theory in which an element can only be in two situations, a member or not a member of some set; fuzzy logic generalizes the possibilities and introduces the concept of shades of membership patterns. The concept of fuzzy logic was first put forth by [39]. Its incorporates the condition that an element is either a member or not a member of some set but also extends the condition by introducing the possibility of membership falling in mixed modes [13].

1) Fuzzy C-Means: Fuzzy C-Means Clustering (FCM) also known as Fuzzy ISODATA, is a clustering technique which is separated from hard k-means that employs hard partitioning. The FCM employs fuzzy partitioning such that a data point can belong to all groups with different membership grades between 0 and 1. FCM is an iterative algorithm. The aim of FCM is to find cluster centers (centroid) that minimize a dissimilarity function. This algorithm works by assigning membership to each data point corresponding to each cluster center on the basis of distance between the cluster center and the data point.

This algorithm works on following these four steps.

Step 1.In these step we find the membership matrix (U) initialize randomly

Step 2. Calculate the centroids (ci)

Steps 3. Using dissimilarity function to calculate the dissimilarities between centroid and data points. That is, the sum of square error which measures how well a clustering fits the data.

Step 4. Check threshold value then stop if we find the correct threshold value [11].

2) Expectation Maximization Algorithm: Expectation Maximization (EM)is a model based approach to solve clustering problems. It is an iterative algorithm that is used in problems where data is incomplete or considered incomplete. EM clusters data, in a manner different from K-means. It is known to be an appropriate optimization algorithm for constructing proper statistical models of the data [5]. EM aims at finding clusters such that maximum likelihood of each clusters parameters is obtained. EM starts with an initial estimate for the missing variables and iterates to find the maximum likelihood (ML) for these variables [31].

The EM algorithm extends the basic approach to clustering in two important ways. Instead of assigning cases or observations to clusters to maximize the differences in means for continuous variables, the EM assigns a probability distribution to each instance which indicates the probability of it belonging to each of the clusters. The goal of the clustering algorithm then is to maximize the overall probability or likelihood of the data, given the final clusters [31]. It's a framework to approach maximum likelihood of parameters in statistical models by using an iterative refinement technique

The EM algorithm

1. Select an initial set of model parameters.
2. Repeat
3. E-Step: for each object, calculate the probability that each object belongs to each distribution, i.e., calculate $P(\Theta_j|o_i, \Theta)$.
4. M-Step: given the probabilities from E-step, find the new estimates of the parameters that maximize the expected likelihood.
5. Until the parameters do not change.
6. Or, stop if the change in the parameters is below a specified threshold [40].

Naser [31] proposed the hybridization of two algorithms k-means and Expectation maximization (EM) in order to improve on the drawback of k-means that does not always guarantee convergence. But with EM, it searches for a local maxima and converges very well. They therefore proposed a fuzzy K-means Expectation Maximization Algorithm (FKEM), the approach used to obtain initial cluster in their algorithm was by computing the weighted fuzzy averages instead of using a simple mean which was also similar to the approach by [24]. The algorithm was tested using IRIS data set and the results compared with fuzzy K-means.

It was concluded that Fuzzy K-means was not able to cluster data correctly when the initial value of K was small. Hence, if good seed centers are not chosen, Fuzzy K-means will not perform very well. If the correct random seeds are not chosen, EM can still perform to make the K-means work correctly, since EM will iterate to find best centers for the given data. It was suggested that Fuzzy K-means along with EM gives a better clustering, than Fuzzy K-means. While[11] proposed a fuzzy clustering algorithm, Expectation maximization fuzzy c-means clustering (EMFCM) which is a hybridization of FCM and Expectation maximization algorithm. It was expected that their proposed algorithm will improve over fuzzy c-means clustering by avoiding the looping problems. But the performance of the algorithm was not ascertained and it was therefore proposed as an extension, so that the algorithm can be implemented and its performance evaluated using the mouse data set in MATLAB tool. Their algorithm uses fuzzy partition which as the limitation of detecting outliers and assigning them to extra clusters.

IV. REVIEWED WORK

There has been many research works in data mining especially in the area of security that has to do with criminal data analysis and profiling. For example, in order to address data mining in the area of profiling,[18] provided a data mining perspective for criminal network analysis involving investigations of organized crimes (e.g. terrorism, narcotics, trafficking fraud, etc.), because it involved multiple and collaborating offenders. Furthermore, the categories of existing criminal network approaches and tools were split into three generations. The first generation (manual approach) was helpful in crime investigation, but when data sets are very large it becomes an extremely ineffective and inefficient method. The second generation tools automatically produced graphical representation of criminal networks hence the name graphic-based approach. Common examples are Analyst's Notebook, hyperbolic tree, etc. although they were also described as capable of using various methods to visualize criminal networks. The downside was it relied mainly on the knowledge of the analysts to study the graphs as it produce only graphical representations of criminal networks without much analytical functionality, they also foretold the properties of the third generation approach.

Techniques also mentioned in their works include Social network analysis that were designed to discover patterns of interaction between social actors in social networks, and is also capable of detecting patterns, subgroups and uncovering a network organization and structure. These subgroups can be detected by the use of cluster analysis such as hierarchical clustering method that can detect underlying groups that were not apparent. Also patterns of interaction between subgroups can be discovered using a black modeling approach; this combination can help show the hidden structure of a criminal network and the knowledge gained may aid law enforcement agencies to better combat crime. They concluded that reliable data and sophisticated analytical techniques are critical for law enforcement agencies. Using automated social network analysis and visualization techniques in more advanced and improved methods would help to bring better

crime analysis tools in view. Also, [1] proposed a method to employ computer log files as history data to search some relationships by using the frequency occurrence of incidents. Then, they analyzed the result to produce profiles, which can be used to perceive the behavior of criminal. Because criminal profiling helps in identifying crime characteristics, which is the first step in developing crime analysis. [19] also introduced a framework for crime trends using a new distance measure for comparing all individuals based on their profiles and then clustering them accordingly.

While [19] and [35] both used data mining approaches in criminal career analysis. Four important factors play a role in the analysis of criminal career (crime nature, frequency, duration and severity). The tool that was developed in both research, extracts these important characteristics of a criminal from the database and creates a digital profile for all offenders, compared all these individuals by a new distance measure and clusters them accordingly. But in [35] there was significant improvement over the research made by[19],by making four major key enhancements mainly to improve the semantics and the efficiency, thereby improving the existing methods of automated criminal career analysis. A new distance measure was introduced that more closely resembles the reality of policing. Instead of the previous, more rigid, comparison of career changes over time,[19]proposed the collection of crimes in a single year as a multiset, which then describes severity, nature and frequency inherently. A new distance measure called Jaccard distance was employed for calculating the difference between two crime-multisets. Instead of the former method of a strict number wise comparison between years (comparing the first year of criminal A with the first year of criminal B, the second year of A with the second year of B, etc.), with the possibility of stretching or shrinking careers, they proposed a novel alignment of the mentioned multisets.

In conclusion, the enormous cloud of one-time offenders gave an unclear sketch of their distance space and the runtime of the chosen approach was not yet optimal, as the clustering method (push and pull clustering) used in the former research was computationally intensive thereby causing performance delays, which is highly inefficient in the real world where the volumes of data increase daily. This was addressed with the use of torus clustering. This was able to solve the problem of time complexity and some major problem encountered in the former method. Finally, it was suggested by [35] to equip the tool with a sub-cluster detection algorithm to provide better insights into the comparability of criminal careers. And also to set fuzzy borders between the different years' crimes within months ending or beginning such a time unit can be (partly) assigned to the next or previous year respectively as well, thus eliminating the problems arising with strict coherence to the change of calendar year. , they also suggested that other clustering techniques be adopted, if they adhere to the demand of incremental addition of single items.

According to [30], solving crimes is a complex task that requires human intelligence and experience, and data mining is a technique that can assist them with crime detection problems. While [9] was able to do a survey that categorized, compared, and summarised from almost all published technical and review articles in automated fraud detection within the past 10 years. It defined the professional fraudster, formalised the main types and subtypes of known fraud, and presents the nature of data evidence collected within affected industries. It was affirmed from their work that this field can benefit from other related fields. Specifically, unsupervised approaches from counterterrorism work, actual monitoring systems and text mining from law enforcement, and semi-supervised and game-theoretic approaches from intrusion and spam detection communities can contribute to future fraud detection research. Based on these,[4] was able to profile and analyse a criminal with the presence of evidence from the crime scene as oppose to the works of [19]; [35] who did their own analysis without evidence from the crime scene by constructing a Bayesian networks for criminal profiling from limited data with the presence of evidence from the crime scene. The method adopted in their work was that they developed a Bayesian network (BN) model of offender behaviour from a database of cleared homicides.

The BN was able to infer the characteristics of an unknown offender from the crime scene evidence, and help narrow the list of suspects in an unsolved homicide. Their research shows that 80% of offender characteristics are predicted correctly on average in new single-victim homicides, and when confidence levels are taken into account this accuracy increases to 95.6%. Their model is shown in Figure 2.1.

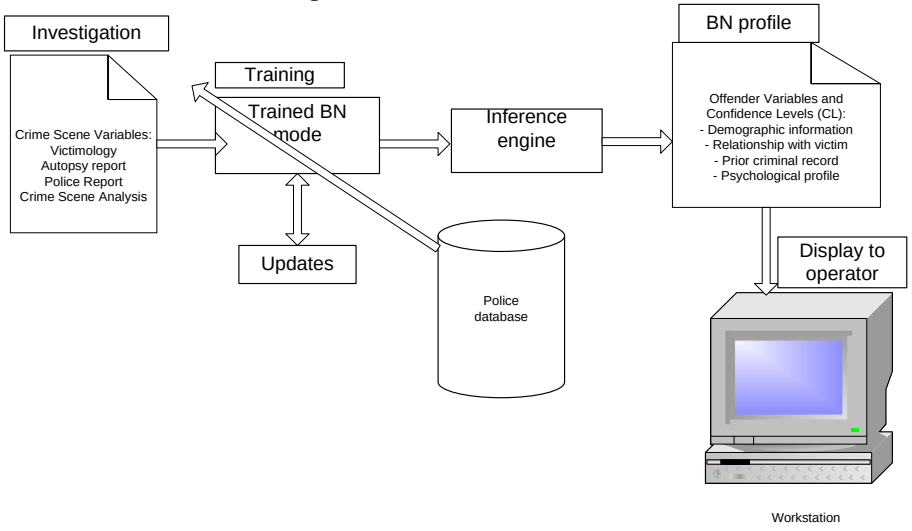

Figure 2.1 A BN model as decision-support tool for police investigation.

Source:[4]

The knowledge that is gained from data mining approaches is a very useful tool which can help and support police forces [21]. These influenced the study of[6]who applied fuzzy association rule mining for community crime pattern discovery. They were able to develop an approach that relieves the need of law enforcement personnel to go through uninteresting, obvious rules in order to find interesting and meaningful crime patterns of importance to their community. Rules discovered in their study offers utility for use from the national level down to the state and community level.

Also, [38] was able to use decision tree and simple K-means algorithm to cluster crimes according to their attributes using online police data from Iraq. The model of their work is shown in Figure 2.2. While [27] used a clustering/classify based model to anticipate crime trends which could be used to lessen and even prevent crime for the coming years. The first task they did was the prediction of the size of the population of a city and the basic approach used to do this was to cluster population sizes, create classes from the clusters, and then classify records with unknown population sizes, the reason why they used clustering to create classes was because classes from clusters are more likely to represent the actual population size of the cities. WEKA mining software package was used and adopted for the clustering. The next task was to predict future crime trends by tracking crime rate changes from one year to the next, and using data mining to project those changes into the future. The basic method they used was to cluster the cities having the same crime trend, and then using ”next year” cluster information to classify records.

They concluded that from the encouraging results, it is believed that crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis, and that there was still lots of options to explore. Visual and intuitive criminal and intelligence investigation techniques can be developed for crime pattern. While [34] used various data mining techniques to combat crime and terrorism in Nigeria, they analyzed how data mining techniques can be adopted by law enforcement agencies in tracking the activities of terrorist and their criminal activities, and also examined the limitation of data mining in fighting crime in Nigeria.

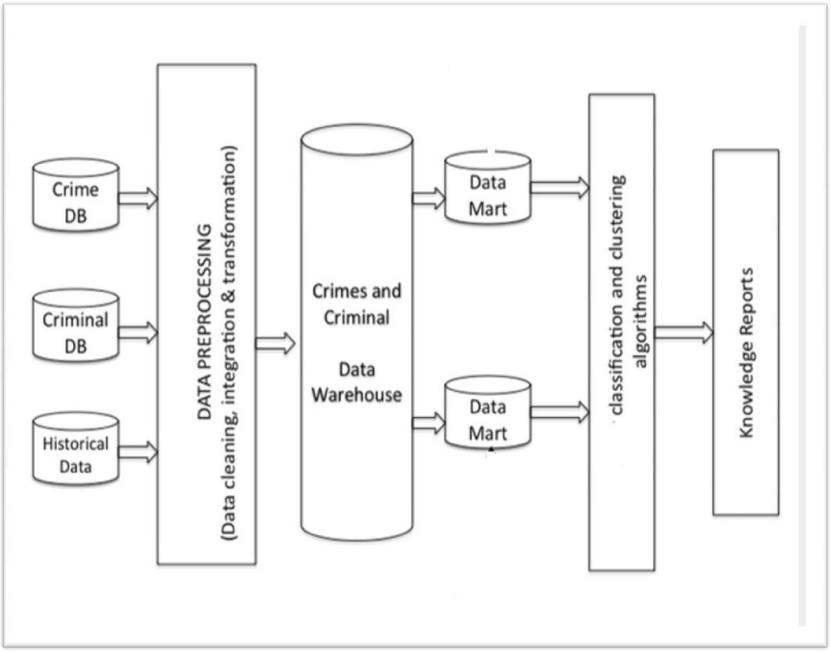

Figure 2.2: Framework Architecture for Crime Analysis

Source: [38].

Malathi [26]looked at use of missing value and clustering algorithm for crime data using data mining. They also looked at MV algorithm and Apriori algorithm with some enhancements to aid in the process of filling the missing value and identification of crime patterns. They also used semi-supervised learning technique for knowledge discovery from the crime records and to help increase the predictive accuracy. As always data pre-processing was discussed stating the various methods of treating missing values, because it is of grave importance to any data mining technique implementation. Two clustering techniques, K-means and DBScan (Density-Based Spatial Clustering Application with Noise) were used to test the data. Thus the two clustering techniques were analyzed in their efficiency in forming accurate clusters, speed of creating clusters, efficiency in identifying crime trend, identifying crime zones, crime density of a state and efficiency of a state in controlling crime rate. Experimental results showed that HYB algorithm show improved results when compared with k-means algorithm and therefore was used in further investigations.

In conclusion the research focused on developing a crime analysis tool for Indian scenario using different data mining techniques that can help law enforcement department to efficiently handle crime investigation. The tool was proven to be effective in terms of analysis speed.

Mande [25] did an intelligent analysis of crime data using data mining methods (clustering and classification) and auto correlation models which aims to identify a criminal based on the witness/clue at the crime spot. The clustering of data was based on criminal/ crime and thereby minimizing the search space, based on the clusters the classification algorithm is then applied to classify the criminal then the auto correlation model authenticates the criminal.

V. CONCLUSION

Hard Clustering was majorly used as the clustering techniques in criminal profiling and theis has made the quality of their output lack essential components which would have made it more in tune with reality. Fuzzy clustering is therefore proposed to address this issuein which a criminal can belong to more than one cluster at a time in order to establish the fact that soft clustering gives rooms for overlapping depending on the degree of membership or probability going by the nature of criminals who can belong to as many group as possible.

Based on the issues identified above, the proposed Fuzzy Clustering will examine how real life data from the law Enforcement Agencies can be integrated and analyzed to produce "profiles" of activity and behavior of criminals. The aim is to provide investigators with rich sources of intelligent information which can be used to predict and prevent criminal activity.

REFERENCES

[1] Abraham, T. and de Vel, O. (2006) Investigative profiling with computer forensic log data and association rules, in Proceedings of the IEEE International Conference on Data Mining (ICDM' 02),Pp. 11 – 18.

[2] Abbas Osama Abu (2008),"Comparisons between data clustering algorithm" International Arab journal of information technology, vol 5 No 3 July 2008.

[3] Askeriya I, H. Jahankhani, S. Wee Lee & Ameer Al-Nemrat, (2010): "Education, Trainingand Awareness (ETA) Four Dimensional Cybercrime prevention model"

[4] Baumgartner, K, Ferrari S and Palermo G (2008): "Constructing Bayesian Networks for Criminal profiling from limited data" Elsevier 21 March 2008.

[5] Bradley P, Fayyad U and Reina C (1998)"Scaling Expectation Maximization clusteringto large databases, Microsoft Research Report, Aug 1998

[6] Buczak, L. Anna, and Christopher M, Gifford (2010)"Fuzzy Association Rule Mining for Community Crime Pattern Discovery" Published in Proceeding ISI-KDD' 10 ACMSIGKDD Workshop on Intelligence and Security Informatics, Article No, 2.

[7] Claridge, Jack (2012) "Criminal Profiling an its use in Crime Solving" Dec 2012 retrievedfrom www.exploreforensics.co.uk

[8] Chen, H. Atabakhsh. H. Petersen, T. Schroeder J. Buetow, T.. Chaboya, L. & Huang, Z(2003, may).COPLINK: Visualization for Crime Analysis In proceedings of the 2003 annual national conference on Digital government research (pp, 1-6). Digital Government Society of North America

[9] Clifton Phua, Vincent lee. Kate smith & ross Gayler (2007) "A Comprehensive survey ofData Mining based Fraud Detection Research"

[10] Dellaert Frank (2002): "The Expectation Maximization algorithm", Technical report GeorgiaInstitute of Technology 2002.

[11] Esh Narayan, Yogesh Birla and Gaurav Tax(2012) "Enhancement of fuzzy C-means clustering using Expectation Maximization algorithm". International journal of Computer Applications Vol 43-no 13 April 2012.

[12] Fredrickson, D. and Siljander, R. (2002). Racial Profiling: Eliminating the confusion betweenRacial and Criminal Profiling. Springfield, IL: Charles C. Thomas.

[13] Gomez, J, and Dasgupta, D. (2002). "Evolving Fuzzy Classifiers for Intrusion Detection". InProceedings of 2002 IEEE workshop on Information Assurance, 68 – 75 West Point, NY : IEEE.

[14] Hoppner F, Klawonn F Fruse R and Runkler(1999),"fuzzy Cluster Analysis"J. Wiley and Sons, Chichester, England 1999.

[15] Herman G and B. Carvalho (2001). "Multiseeded Segmentation Using Fuzzy Connectedness". IEEE Pattern Analysis and Machine Intelligence 23(5): 460-474, 2001,

[16] Hick, S, and Sales, B, (2006). "Criminal Profiling: Developing an Effective Science and Practice". American Psychological Association, Washington, DC.

[17] Huet William (2012) "Uncovering the truth: the reality of Criminal profiling ", oct 2012 retrieved from www.coloradotech.edu

[18] Jennifer Xu, Hsinchun Chen (2005) "Criminal Network Analysis and Visualization: A Data Mining Perspective" communications of the ACM, Vol, 48 No. 6, pp, 100-107

[19] Jeroen S, De Bruin, Tim K. Cocx, Walter A, Kosters, Jeroen F, J. Laros and Joost N, Kook (2006) "Data Mining Approaches to Criminal Career Analysis," in Proceedings of theSith International Conference on Data Mining (ICDM' 06), pp, 171 – 177

[20] Kumar Anoop Jain and Satyam Maheswan (2012) "Survey of Recent Clustering Techniques in Data Mining" International Archive of Applied Sciences and Technology Volume 3 [2] June 2012: 68 – 75.

[21] Keyvanpour, M.R, Javideh, M. and Ebrahimi, M.R. (2010) Detecting and investigating crimeby means of data mining : a general crime matching framework, Procedia Computer Science, World Conference on Information Technology, Elsvier B,V. Vol. 3, pp. 872 - 830

[22] Krishnamurthy Revatthy (2012) "Survey of Data Mining Techniques on Crime Data",International Journal of Data Mining Techniques and Applications vol. 01, pp. 48 –55.

[23] Le Lievre, E. & Jamieson, R, (2005). ''Investigation of Identity Fraud in Australian Organisations''. Collaborative Electronic Commerce Technology and Research (CollECTeR), 1 10.

[24] Looney C. G (2002) "Interactive clustering and merging with a new fuzzy expected values. Pattern Recognition Lett. Vol 35, 2002 187-197.

[25] Mande Uttam, Y. Srinivas, J.V.R.Murthy (2012) An Intelligent analysis of Crime Data using data mining and auto correlation models. "International Journal of Engineering Research and Applications" (IJERA) ISSN: 2248-9622 www.ijera.com Vol. 2, Issue 4,July – August 2012. Pp. 149 – 153.

[26] Malathi and Santhosh, (2011) ''an enhanced algorithm to predict a future crime using datamining'' International journal of computer applications (0975 - 8887) volume 21 – no. 1, May 2011.

[27] Malathi, Dr .S. Santhosh Baboo (2011) ''Algorithm crime prediction Model based on theanalysis of crime clusters''Global journal of computer science and technology volume 11 Issue 11 version 1.0 july 2011 Type:pp 46-51

[28] Malathi.A, Dr. S.Santhosh Baboo and anabras (2011) ''an intelligent analysis of city crimedata using data mining'' international conference on information and electronics engineering IPCSIT vol6 (2011).

[29] Marx, G. and Reichman. N (1984), Routinising the Discovery of Secrets Computer asInfromants. American Behavioural Scientist, (27:4), March/April, 423-452.

[30] Nath,S.(2007) crime data mining,advances and innovations in systems, K. Elleithy (ed.),computing sciences and software Engineering,Pp.405-409.

[31] Naser Sara, Rawn Alkhaldi and Gregory Vert (2005) "A modified fuzzy k-means clustering using Expectation maximization algorithm.

[32] Nykodym, N, Taylor, R, Vilela, j,(2005): ''Criminal profiling and insider cybercrime''computer law and security report 21(5),pp.408-414.

[33] Petherick, W.(2006). Serial crime: theoretical and pratical issues in behavioural profiling,Elsevier Inc, London, Great Britain

[34] Raphael O. O & Francis O.E (2011): "Combating Crime And Terrorism Using Data MiningTechniques", 10th International Conference On Information Technology for People Centred Development Nigeria Computer Science Conference Proceedings, Vol 22, pp80 – 89 2011.

[35] Tim k. Cocx, Jeroen F. J Laros, Walter A. Kosters (2008) "Enhancing the Automated analysis of criminal careers".

[36] Turvey, B. (2002): "Criminal Profiling, An Introduction To Behavioural Evidence". UK,Elsevier.

[37] Wang Qiang & Vasileios M, (2004): "A Clustering Algorithm for Intrusion Detection" DataEngineering Laboratory,

[38] Swadi B, Kadhim. Al- Janabi (2011) "A Proposed Framework for Analyzing Crime Data Set Using Decision Tree and Simple K – Means Mining Algorithms" Journal of kufa forMathematics and Computer Vol. 1. No 3, may 2011, pp 8 – 24.

[39] Zadeh L.A,(1975)"Fuzzy logic and approximate reasoning, synthese ,vol, 30 pg 407-428

[40] Han, Kamber and Pei,(2012) "Knowledge discovery and data mining"

