Criminal Pattern Identification using Genetic Algorithm
Vibhuti H. Jani1
1SOCET

Abstract—Data mining is the process of identifying interesting patterns in data. These patterns must be meaningful in that they lead to some advantage. These patterns allow us to make prediction on new data. Clustering is a technique that group data items into classes with similar characteristics to maximize or minimize interclass similarity. For example, in the same way or differently to identify the suspects who carried out the crimes associated with gangs are different between the groups. Classification is a technique that finds common properties among different crime entities and organizes them into predefined classes. Many learning techniques can be quite complicated and typically are expressed as a set of rules or decision trees. For example, in the same way or differently to identify the suspects who carried out the crimes associated with gangs are different between the groups.

I. INTRODUCTION
Nowadays, a lot of crimes in all countries of the world are expanding. The cost of living is rising, while the main reason unemployment is growing. Each one of us to reach our goals and our lives should be employed to accomplish. Population explosion had difficulties to find jobs for everyone. Jobless person both ends meet and can find hard to develop his frustration. They are easy to get money where antisocial or criminal activities, up to the last end. The other main reason is the broken families or single parenthood. Here kids will not get any attention from family or society; that alone makes them engage in crimes. For example, children will rebel to society, it is unpleasant past experiences. Like Internet Medias, porn films and detective novels make them worse play a tangible role. The Committee is a crime for people to make their living, but all this time I have a good reason to believe that. After studying research papers, I found interest in criminal pattern identification. So, I am motivated to carry the work in this area as a dissertation. My efforts in this area will surely give important results to reduce the crime by identifying crime patterns efficiently.

II. CRIME DATA MINING TECHNIQUES
The main techniques of the crime data mining are clustering, association rule mining, classification and sequential pattern mining [10]. Although all of these efforts, the crime mining still is a highly complex task.

<table>
<thead>
<tr>
<th>Crime Type</th>
<th>Description</th>
</tr>
</thead>
</table>

A. Clustering Techniques
The methods of assigning items [10] does not set a predetermined classes. Some researchers, such as automatically Crime Records individuals, organizations, and vehicles of different objects to associate the concept of space uses an algorithm based on statistics. There are many techniques like link analysis techniques used to identify similar transactions, the Financial Crimes Enforcement Network AI system, money laundering and other financial crimes detection and analysis of information to support the absorption of bank secrecy act. Create a clustering analysis of crime a major part of the crime can be automated, but is typically limited by the required high computational intensity.

B. Association Rule Mining
Researcher also potential future network attacks [10] to help find intruders' profiles network can apply this technique. Association rule mining, sequential pattern mining often at different times for the same things on the set of transactions that have been occurring sequences are seeking. Network intrusion detecting, this approach is time-stamped data can identify patterns among the intrusion. Hidden patterns show the benefits of crime analysis to achieve meaningful results, but wealthy and highly organized information is needed.

C. Deviation Detection
This technique differs significantly from the rest of the information that data to study the use of special measures. Also called outlier detection, investigators fraud detection, network intrusion detection, and other crime can apply this technique [10] analyzes. However, such activities sometimes makes it difficult to identify outliers, can appear to be normal.

D. Classification
This technique is often used for crime prediction, crime classification that can reduce the time required to identify. However, this classification technique requires a predefined classification scheme. A high degree of accuracy to predict the missing data will be limited because the classification requires a reasonably full training and test data.

E. String Comparator Techniques
Researcher often use intensive computation [10] that requires a string comparators can be used to evaluate the data. It is interesting to compare a string of computer scientists that string matching or string distance whether measures. So the "edit distance", the minimum number, deleted, inserted into a receptacle as a general measure of similarity between two strings to define, and substitutions required to transform one string to another.
Criminal Pattern Identification using Genetic Algorithm
(IJSRD/Vol. 3/Issue 04/2015/315)

<table>
<thead>
<tr>
<th>Crime Type</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Traffic Violations</td>
<td>Driving under the influence of alcohol, fatal/personal injury/property damage traffic accident, road rage</td>
</tr>
<tr>
<td>Sex crime</td>
<td>Sexual offences</td>
</tr>
<tr>
<td>Arson</td>
<td>Arson on buildings</td>
</tr>
<tr>
<td>Gang / drug offences</td>
<td>Narcotic drug offences (sales or possession)</td>
</tr>
<tr>
<td>Violent crime</td>
<td>Criminal Homicide, armed robbery, aggravated assault, other assaults</td>
</tr>
<tr>
<td>Cyber crime</td>
<td>Internet frauds, illegal trading, network intrusion/hacking, virus spreading, hate crimes, cyber piracy, cyber pornography, cyber-terrorism, theft of confidential information.</td>
</tr>
</tbody>
</table>

Table 1: The Classification of Crime [7]

IV. PROPOSED ALGORITHM

Currently the most useful method for detecting crime hotspots is clustering. The nearest neighbor hierarchical clustering & K-means clustering are typical methods in spatial ellipse techniques. These ellipses represent some hotspots. But crime hot spots are not natural spatial ellipses forms [1]. So these methods do not represent the actual spatial distribution of crime and often mislead the researchers into focusing on areas of low crime importance within an ellipse.

A genetic algorithm (or GA) is a search technique used in computing to find true or approximate solutions to optimization and search problems. Genetic algorithms are categorized as global search heuristics. Genetic algorithms are a particular class of evolutionary algorithms that use techniques inspired by evolutionary biology such as inheritance, mutation, selection, and crossover (also called recombination).

Algorithm works like this:

1) [Start] a population is created with a group of individuals created randomly. The individuals in the population are then evaluated.

2) [Caused Factors] The evaluation function gives the individuals a score based on how well they perform at the given task.

3) [New Population] Create a new population by repeating following steps until the new population is complete

4) [Selection] Two individuals are then selected based on their fitness, the higher the fitness, the higher the chance of being selected.

5) [Crossover and Mutation] These individuals then "reproduce" to create one or more offspring, after which the offspring are mutated randomly.

6) [Accepting] Place new offspring in a new population

7) [Replace] Use new generated population for a further run of algorithm

8) [Test] If the end condition is satisfied, stop, and return the best solution in current population

9) [Loop] Go to step 2

V. IMPLEMENTATION

Eclipse LUNA will been used to develop the proposed algorithm as it is java based. Developing language is advanced java.

Fig. 1: Dataset
VI. CONCLUSION

It is the application of data mining techniques can produce important results is one of the most important fields. Crime Analysis, is a part of Criminology, explore and discover crimes and criminals, including their relationships with that task.

The proposed algorithm will help to identify effective criminal and crime patterns. And it will also help to predict crime. In proposed algorithm genetic parameters and the experiment has been conducted to find a better outcome. The genetic algorithm operation has been examined to give better results. This results could potentially be used to prevent crime for the forth coming years. Crime data mining has a promising future for increasing the effectiveness and efficiency of criminal and intelligence analysis.

REFERENCES


