Crime Forecasting System
(An exploratory web-based approach)

1-Yaseen Ahmed MEENAI
Institute of Business administration (IBA)
Karachi, Sindh, Pakistan

2-Twaha Ahmed MEENAI, 3-Arafat TEHSIN, 4-Muhammad Ali ILYAS
Department of Computer Science and Engineering, Bahria University
Karachi, Sindh, Pakistan

ABSTRACT

With the continuous rise in crimes in some big cities of the world like Karachi and the increasing complexity of these crimes, the difficulties the law enforcing agencies are facing in tracking down and taking out culprits have increased manifold. To help cut back the crime rate, a Crime Forecasting System (CFS) can be used which uses historical information maintained by the local Police to help them predict crime patterns with the support of a huge and self-updating database. This system operates to prevent crime, helps in apprehending criminals, and to reduce disorder. This system is also vital in helping the law enforcers in forming a proactive approach by helping them in identifying early warning signs, take timely and necessary actions, and eventually help stop crime before it actually happens. It will also be beneficial in maintaining an up to date database of criminal suspects includes information on arrest records, communication with police department, associations with other known suspects, and membership in gangs/activist groups. After exploratory analysis of the online data acquired from the victims of these crimes, a broad picture of the scenario can be analyzed. The degree of vulnerability of an area at some particular moment can be highlighted by different colors aided by Google Maps. Some statistical diagrams have also been incorporated. The future of CFS can be seen as an information engine for the analysis, study and prediction of crimes.

Keywords: crime, exploratory analysis, graphs.

1. INTRODUCTION

The aim of this paper is to propose the development of a Crime Forecasting System to support the police operations in Karachi. The proposed system will be able to collect and store crime incidents and analyze it using different tools to process crimes' data. This would help the *Law Enforcement Agencies* in forecasting crime and preventing it before it actually occurs. Thus, leads to a reduction in the crime rate. The system would also benefit *the general public* by providing the victims an unconventional and advanced medium to report the criminal incidents they are subjected to.

The features of the system if highlighted, takes the form as follows. It;

- Comes out as a decision support system that assists the law enforcement agencies in optimal utilization of resources
- Is able to predict crime patterns by analyzing the graphs based on real time data sets.
- Shows the results in multiple views using *Google Maps* with various perspectives.
- Provides ease in use with user friendly interfaces due to which no special training would be required.
- Improves the quality of results/output with grown data sets over the passage of time.
- Highlights Gang based criminal activities located within different parts of the city thus expose criminal social networks.
- Provides information regarding criminal Characteristics such as their facial appearance, dressing style, nature of weapons used for crime.
- Clearly depicts the comprehensive picture of criminal activities in a city

With these features the most valuable prospective outcomes of the system are considerable reduction in the crime rate, maintenance of a standardized and centralized crime database and satisfactory response time by the law enforcement agencies.

2. BACKGROUND

Keeping the focus bounded to a crime prone mega city like Karachi, the study through field survey and the analysis of the current conventional approach/practice for handling and managing crimes reveal a series of flaws. Some of which are stated below.

- **Lack of data sharing**
  Problem-solving clearly depends on the availability of robust data (Read and Tilly, 2000, p. 28) but the manual system currently in use makes the communication channel and information flow very static and rigid.
- **Inefficient Resource Allocation**
  According to Boba (2001) Crime Analysts can assist in effective allocation of resources by determining the times and areas i.e. when and where the offenses are likely to occur. The system under study shows that due to the absence of appropriate data processing mechanism the practice of efficient resource allocation is not in place.
- **Inconsistent standards of data storage**
  The data storage standards are not consistent as different departments use customized formats, schema and parameters.
- **Slow Response Time**
  The delay, due to the absence of a rapid reporting mechanism, non preemptive approach and the fulfillment of departmental protocols at law enforcement agencies produces a mandatory flaw of a slow response.

With the above mentioned flaws of the system the crime rate is likely to increase naturally. Despite several governmental efforts and restructuring of the law enforcement agencies/departments, the history asserts this trend too.

In order to address these problems the need of a crime forecasting system arises. The initial research was based on the consideration of multiple factors specially the Crime Index. This approach employed the use of statistical methods namely Regression, Time Series Analysis, Double Exponential Smoothing. Crime Index, which is the rate of crime over population, was the main input parameter provided to these statistical methods, on the basis of which it forecasted results.

The forecasting process comprised of three steps:
1. Evaluation of Crime Index.
2. Gathering data.
3. Application of the techniques on the gathered data.

However the drawback associated with this approach was that it did not cater for all the effecting factors, thus less effective and the intended scope was not fully covered. For instance, if one considers the murder crime, it can be seen as only murder while it may be a result of an initial crime. i.e. the reasons and causes behind may

possibly be categorized as crimes too, thus affecting the results of forecasting.

The literature survey that was conducted in order to understand and identify the features that we need to incorporate into the system, as well as to learn the various approaches that others have taken to improve this system. The following table- 2.1 contains summary features that depicts how have the different systems been implemented in terms of the features they provide (*details of these systems can be obtained from the first author*).

Features	CrimeView [3]	ATAC[4]	WebCAT[1]	CADmine[5]	Data Detective [2]
Statistical Model	✓	✓	✓	✓	✗
Spatial Analysis	✗	✓	✓	✓	✗
Data Entry	✗	✓	✗	✗	✗
Features	CrimeView [3]	ATAC[4]	WebCAT[1]	CADmine[5]	Data Detective [2]
Calendaring	✗	✓	✗	✗	✗
Reporting	✓	✓	✓	✓	✓
Data Mining	✓	✓	✓	✓	✓
Artificial Neural Networks	✗	✗	✗	✗	✓
Environmental Mapping	✓	✓	✓	✓	✓
Alerts	✓	✓	✗	✓	✗

Table 2.1 (Benchmark)

3. METHODS and MATERIALS

3.1 The Techniques

Exploratory Data Analysis lies at the core of **Crime Forecasting System.** *Exploratory data analysis (EDA)* used to analyze data for the purpose of formulating hypotheses worth testing. It was so named by *John Tukey,* Weiss (2002) . The objectives of EDA are to:

- Suggest hypotheses about the causes of observed phenomena
- Assess assumptions on which statistical inference will be based
- Support the selection of appropriate statistical tools and techniques
- Provide a basis for further data collection through surveys or experiments.

The approach under consideration employs a variety of techniques (mostly graphical) to,
1. Maximize insight into a data set.
2. Uncover underlying structure.
3. Extract important variables.
4. Detect outliers and anomalies.
5. Test underlying assumptions.
6. Develop parsimonious models.
7. Determine optimal factor settings.

The disadvantages of using this technique were,
- It does not provide definitive decisions always.
- Requires a high degree of judgment

Since the system relies heavily on crime data, one of the vital tasks was to collect it. Initially the law enforcement agencies were contacted for this purpose but the effort was not productive. Therefore the data was acquired through **Interviews**, **Questionnaire**, **Observations** and **Research**. These techniques ensure first hand real data with magnified volume and minimum error. Generally most of the cases are not reported due to the problems people have to face in reporting them to the law enforcement agencies. This Web based approach works as a remedy to this problem.

3.2 Architecture

The underlying database design which serves data storage and retrieval. Its design is fully optimized and synchronized with the questionnaire and has the flexibility for future enhancements and adjustments. New factors may be added according to the dynamic crime patterns in order to improve the effectiveness of the system.

The context diagram serves to focus attention on the system boundary and helps in clarifying the precise scope of the system. Figure 3.1 .
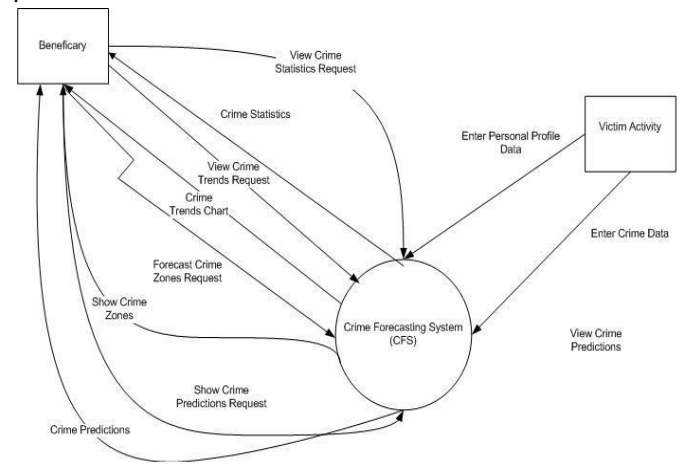

Figure 3.1- the context diagram

3.3 Environmental Mapping

The geographical data assists in mapping with the help of Google APIs which act as a bridge between our system and Google Maps. The process is explained in Figure 3.2
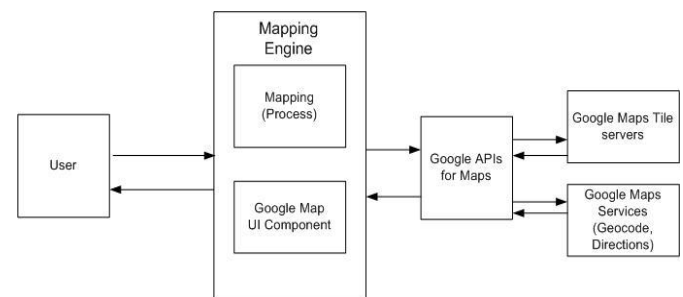

Figure 3.2 – The Mapping Process

3.4 The dependency contours

Figure 3.3 shows the dependencies of different qualitative entities which ensures the relationship among them. This vital result authenticates the effectiveness, productivity and the workability of the system. The contours and the color intensities confirm the relationship between areas and crimes.
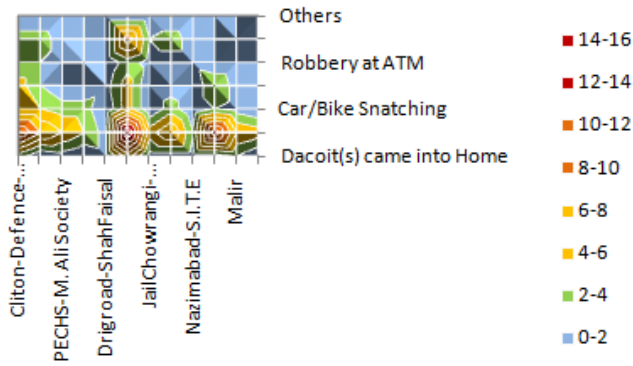

Figure 3.3 - Contours

3.5 Statistics

Figure 3.4 verifies the hypothesis made against the authenticity of data collected through this system as most of the crimes remain unreported due to the procedural hindrances. The analysis conducted on the involvement of number of criminals assists in prediction of the nature of crime expected as per gang based activities.
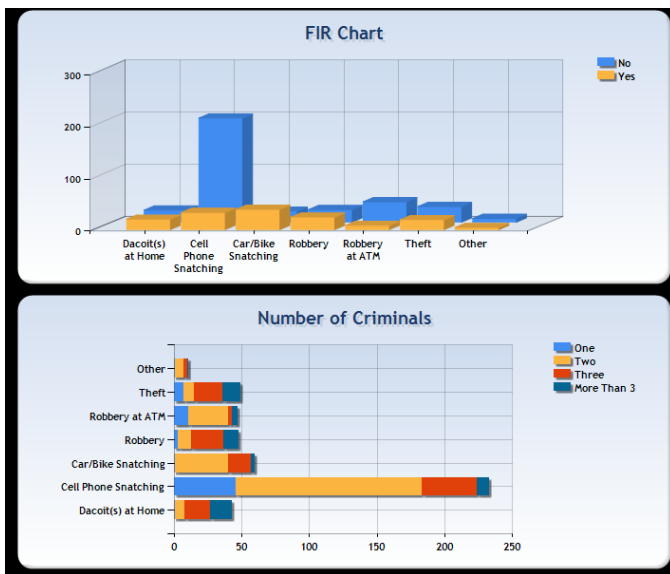
Figure 3.4

4. THE INTERFACES

4.1 Data Representation

The following snapshot in figure shows the overall picture of criminal activities per area. Moreover different colors represent different intensities of crimes. At a glance delivery of information increases the visual value of the system. The interface reflects the changes in real time as per updated data. Thus the most current picture of crimes per area appears with shifts and trends. The interface is a result of high profile integration of the data and Google Maps which enhances the effectiveness and value of the system manifold.
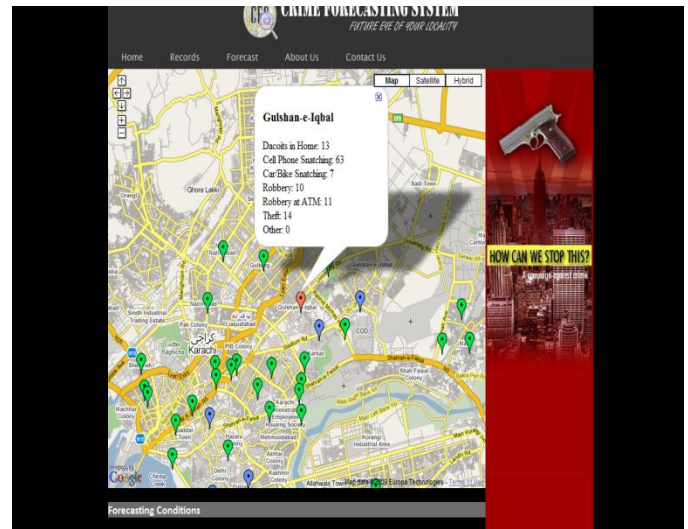

Figure 4.1

4.2 The Forecasting

Once the data is seen there comes the awaited output of the system i.e. forecasting. The forecast can be seen as per area with different factors and aspects. This provides a closely conclusive and demographic picture to the local law enforcement agencies. For instance the parameters of behavior and age tell about what social category of people is to be focused to address the increase in crime in the respective area.
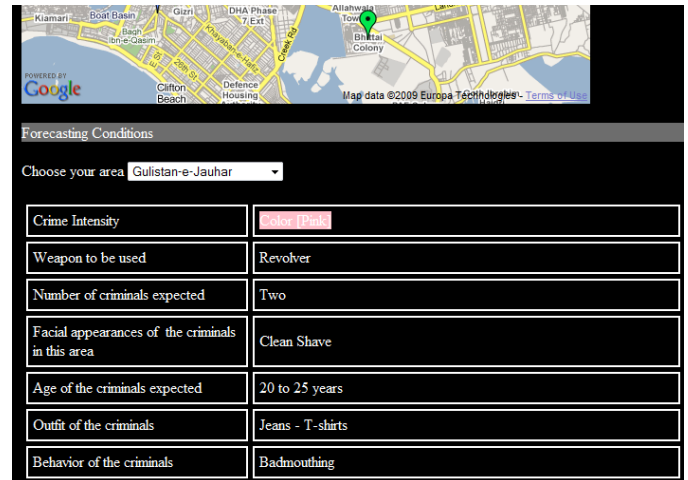
4. CONCLUSION

A thorough study revealed that there are certain areas of criminology which are needed to be addressed using computational capabilities and techniques. Inferential engine and prediction on the basis of the implementation of statistical and AI techniques through computers have been found very beneficial if exploited for the purpose. Following are the results after the implementation of our efforts

- Standardization of input in form of a questionnaire and designing of a centralized database
- Analytical Engines
- Inferential Engines
- Map based visualization of crime affected areas

5. FUTURE WORK

To increase the usefulness of Crime Forecasting system, one of the enhancements that can be made is to develop the system as a mobile application. This would allow the user to report the crime remotely as soon as he observes one and eliminate the need for a desktop system to report a crime. To increase the effectiveness of the Crime

Forecasting System, more techniques could be used for forecasting crime. Techniques belonging to the Data Mining and Artificial Intelligence field may be incorporated into the system to make the forecasting process even more effective and efficient.

6. ACKNOWLEDGEMENTS

7. REFERENCES

Boba, R. (2001), Introductory Guide to Crime Analysis and Mapping, *Report to the Office of Community Oriented Policing Services Cooperative Agreement,* US Department of Justice; Washington DC, (pg. 9)

Read, T., and Tilly, N. (2000*),* Crime Reduction Research Series Paper 6, *Not Rocket Science?* Home Office; London, (pg. 28-35)

Weiss A. Neil (2002), Introductory Statistics, 5th Edition, Addison Wesley. (pg. 193)

ATAC | Automated Tactical Analysis of Crime
http://www.bairsoftware.com/atac.html
CADmine
http://www.coronasolutions.com/products/cadmine.shtml
CrimeView®
http://www.theomegagroup.com/police/crime_mapping_solutions.html
DataDetective
http://www.sentient.nl/?dden
WebCAT™ Crime Analysis System
http://www.daprosystems.com/DPSProducts/crimeanalysis.aspx