CHAPTER 1

INTRODUCTION

1.1 TAMIL NADU POLICE

Police organizations are progressively betting a lot on the Geographical Information System (GIS) for logical model and real-world illustration. This confidence has concerned law enforcement agencies to include completely different knowledge like location, crime and social demographic so as to develop crime analysis.

The GIS component of the system can help Police departments and detective agencies to identify the geographical location of crime events and evaluate the crime spot specifically and perfectly. The GIS provides the information required by police Officers to evaluate and make decisions by where crime is happening and where the density of crime is higher. The ability to contact and process information quickly while displaying it in a crime forecasting model allows police to distribute resources rapidly and more efficiently.

The analysis of the existing crime forecasting model verified that in Tamil Nadu Police strength, the crime forecasting model in position is generally physical and the majority of the information is reserved in filing cupboard. The data collection, crime analysis and convention of information are still poor, incorrect and premature. Data is collected using police records and books which were designed during the grand times. The old model is
slow and flat to errors. Since Police action requires timely and reliable information, this model weakens the Police’s capacity to fulfill its mandate. It is also noted that there is administration due to organizational settings. Authority must be required for a higher officer to get any information which may not necessarily be top secret.

1.2 CRIME ANALYSIS

Crime

In a crime and law enforcement agency, the central focus is on crime, both those reported to the police and those that are not. Thus, the central type of data analyzed is crime and the information surrounding it, such as arrests, offenders, victims, property, evidence, etc.

In addition to information about crime, law enforcement agencies address to several different problems and so collect several different forms of information. Samples of law enforcement knowledge that are usually accessible for crime analysts are service (e.g., burglar alarms, suspicious activity), traffic info (e.g., accidents and citations), citizens’ perceptions (e.g., worry about crime, crime prevention behavior, satisfaction with the police), victimization, probation records, and parole information. Together with this phrase refers to identifying patterns and finding out the relationships of crime and law enforcement knowledge with different forms of information recorded.

Crime Analysis

“The qualitative and quantitative studies of crime and law enforcement information in combination with social-demographic and spatial factors to apprehend criminals, prevent crime, reduce disorder and evaluate organizational procedures”.
The qualitative and quantitative crime analysis uses each qualitative and quantitative information and analysis techniques. Qualitative information and analytical techniques are confronted with non-numerical information also because the examination and interpretation of observations for the aim of discovering underlying meanings and patterns of relationships. This is often the commonest of field analysis, content analysis and historical analysis. Quantitative information is information primarily in numerical or categorical format. Quantitative analysis consists of manipulations of observations for the aim of describing and explaining the phenomena that those observations replicate and is primarily statistical.

Crime analysis employs each kind of information and techniques reckoning on the analytical and sensible want. For instance, crime information is often employed in varied ways that, each quantitatively and qualitatively. The knowledge like date, time, location and sort of crime is quantitative in that statistics are often used to research these variables. On the other hand, narratives of crime reports are thought of qualitative information in that an outsized range of narratives is nearly not possible to analyze statistically and are primarily examined to work out general themes and patterns.

Three basic Fundamentals of crime analysis

In the crime element, Crime data constitutes a kind of criminal offense, crime name, victim of crime and a private accused of obtaining committed the crime. Crime kind is defined as either a felony (serious) or misdemeanor (less serious) whereas crime name is an identity of a criminal offense. An accused person has particulars that embrace identity selection, name, sex, age, race, employment standing and address (both physical and postal).
In the spatial element location and socio demographic information are part of the knowledge needed by law enforcement agencies within the analysis of crime. They are each spatial as they are tied to crime event.

In the location, crime analysis includes a spatial element which could be defined with four attributes: distribution, surroundings, movement and action space. The spatial distribution of a criminal offense describes how the locations of crime events relate to at least one another. This includes purpose distribution patterns (for example, clustered, random or even uniform), densities, progression, axes, and others. The presentation of location on a map or laptop or desktop makes it easy for cops to visualize where crime happens and what forms of crime tend to occur along or near each other. Movement, on the other hand, describes the dynamics of the spatial element of crime. Location is represented at different levels looking forward to the detail of the analysis to be done like, motive analysis that's incredibly detailed and shows the precise location of crime events. Motive analysis normally uses the addressing system where all buildings and houses have an address and thus the address of a building is tied to a criminal offense event. The polygon analysis could be a small quantity detailed as a result of it provides crime location per polygon. A polygon is given constant color assuming the crime happens uniformly in the vicinity. The polygons utilized within the analysis of crime are mainly administrative boundaries as defined by law enforcement agencies or national governments.

Socio-demographic data refers to characteristics of individuals and groups like sex, race, income, age, and education. Social demographic data are additionally represented throughout a map. An area is demarcated into different census areas giving the little print of the population per age cluster and gender. This information helps the police to understand why certain crimes tend to occur where they are doing and such knowledge is analyzed
along with the location of crime events. While using crime analysis, it is conjointly achievable to identify the link between hot spots and the native population, sorts of land use, the designed setting and native crime history.

In the temporal element, time and date of the prevalence of a criminal offense event are vital aspects of crime analysis. They indicate that crimes have a bent to occur throughout which period of the day and those days of the month or perhaps months of the year.

**Crime Analysis Model**

There are five varieties of crime analysis utilized in the past to current analysis of crime. Every type contains some characteristics of crime analysis normally; however every other is restricted within the variety of information and analysis used similarly as in its purpose.

The principle of intelligence analysis is to support sworn personnel within the identification of complicated and apprehension of people to consequently forestall criminal activity. A connected objective is to link information along knowledge, determine crime relationships and determine crime areas for any investigation by inserting the analysis in an exceedingly simple framework that is easy to know. Abundance of the data analyzed within the field of intelligence analysis, not reported to the police by voters however is gathered by law enforcement authorities. Samples of knowledge assortment strategies embody surveillance, informants, and participant observation. Additionally, the sort of knowledge is not restricted to criminal information, however, will embody phone conversations, travel info, financial/tax information, and family and business relationships. Intelligence analysis has historically centered on an additional or less organized criminal activity.
Intelligence crime analysis has conjointly been known as “profiling,” that is the procedure of building a “profile” of an unknown criminal, based mostly on the character of the crime, the necessities of the case and also the uniqueness of the victim. Like intelligence analysis, this kind of research focuses totally on qualitative knowledge, surrounding serious serial crimes like murder and rape. Crime data are collected and analyzed on a private level for those persons primarily or peripherally concerned with the incidents. The spatial nature of the incidents and the connected locations are additionally taken into consideration. The key purpose of criminal investigative analysis is to develop patterns of serial crimes across town, state and even national boundaries, by linking behavior and proof inside and among incidents so as to catch the offender and/or clear cases. This is often an awfully specific kind of crime analysis that is primarily done on the federal law enforcement level since these varieties of crime occur occasionally and cross jurisdictional boundaries.

Tactical crime analysis focuses on crime information, from day-to-day crimes, reported to the police. Tactical crime analysis additionally focuses on specific data concerning every crime like technique of entry, hot purpose of entry, suspect’s actions, form of the offender, furthermore because of the date, time, spot and kind of location. Field information like uncertain activity necessitate service, criminal trespass warnings and persons with scars, marks, or tattoos collected by officers is additionally thought-about within the analysis. It’s largely utilized in specific operations and management. There are three functions of tactical crime analysis: They are 1) linking cases along with identifying the notable crime patterns and trends, 2) identifying potential suspects of a criminal offense or crime pattern and 3) clearing cases. The most important focus of tactical crime analysis is examining crime daily thus on spot patterns, trends and investigative leads for recent criminal and potential
criminal activity. Once a criminal offense pattern, suspect, or investigative lead is identified, the info is compiled and disseminated to patrol officers and detectives.

Strategic crime analysis consists primarily of quantitative analysis of a mixture of daily, monthly and yearly collections of criminal data. That is, general classes like date, time, space and sort of incident are analyzed rather than qualitative information like account descriptions of incidents. Variables together with burglary, class, sex, income, population, location and placement sort are examined in conjunction with law enforcement data within the analysis method. The two primary functions of strategic crime analysis are 1) to assist within the recognition and analysis of long-term issues and 2) to conduct investigative studies.

Administrative crime analysis is totally and completely different from the previous forms of analysis in that it refers to the presentation of the findings rather than to statistical analysis. The choice and the way to gift information are the main focus of administrative crime analysis. Often, the sort of knowledge that is presented represents the “tip of the iceberg” of all the works and analysis that has previously been done.

The Figure 1.1 displays how all of these types of crime analysis relate to a minimum of each other in terms of the quantity of aggregation of the information. That low levels of aggregation focus on individual cases and used qualitative information and analysis techniques and other people with high levels of aggregation focus on a restricted scope of larger amounts of knowledge and data.
At the very best of the Figure 1.1, criminal investigative analysis and intelligence analysis utilize the smallest quantity aggregated and most qualitative information. The info encompasses data regarding informal networks of criminals and their non-criminal acquaintances and relatives still as where the individuals live, work, and play. The most targets here are about the specifics of criminals, the nature of their crimes, their relationships and their lives typically.

**Objectives of Crime Analysis**

The main function of crime analysis is to support law enforcement endeavors. The first objective of law enforcement is that the capturing of criminals; consequently, one in every the first objective of crime analysis is to assist within the apprehension of criminals. For instance, a detective might have a theft incident within which the suspect encompasses a snake tattoo on his left arm. The crime analysts might assist by looking at a database of field incident cards to spot people with such a tattoo. Also, a criminal offense
analyst might conduct a time of day/day-week analysis of burglary incidents that will assist officers in surveillance of a part to catch offenders.

Another primary objective of law enforcement is to prevent crime through ways apart from apprehension. This objective lends itself notably well to help form crime analysis. For instance, members of the police department are conducting themselves against the law prevention campaign concerning residential burglary and would like to focus on their resources within the areas that require it the foremost. Crime analysis will assist in coming up with community education and patrol response tailored to the matter by providing spatial analysis of residential burglary, analysis of how, when and where the burglaries occurred and analysis of what things were stolen. This data might be used to develop crime prevention suggestions like closing and locking a garage door.

Many criminologists contend that social disorder will cause crime; that is, scar and different indicators of social decay left unchecked will attract crime and accelerate additional decay. Thus, reducing disorder may be a law enforcement objective and, by extension, one for crime analysis likewise. Crime analysis will assist with these efforts by providing analysis and analysis of disorder indicators like traffic accidents, noise complaints, or trespass warnings which will assist officers in addressing these problems before they become additional serious issues.

Another objective of crime analysis is supported by the assessment of organizational procedures. Several instances include resource allocation, the evaluation of crime prevention programs, realigning geographic boundaries, forecasting staffing needs, and developing performance actions for the police department.
1.3 CRIME FORECASTING

Forecasting methods were previously used in an extensive range of fields, such as crime forecasting for predicting the crime events and financial forecasting for, among others, forecast stock values. Most forecasting techniques use data from the past records to try and discover patterns, which are then used to build a forecast of what will happen in the future. For crime, related data are available for events that have been reported and from offenders who have been under arrest. However, the police officers are not only interested in the amount of crime that will take place at a particular point of time, they are also interested in its location. This objective is unusual commencing, for example, The police are interested in “crime hotspots” where a high amount of crime is predicted. This information can then be used to take safety measures, such as police patrolling specific areas, or placing resources.

Crime forecasting can be utilized to generate hot spot maps, which give a diagrammatic representation of predicted crime in space. This kind of forecasting is called geographical forecasting. The main purpose of this thesis is to create maps for the Kernel Density Estimation (KDE) deployment of police forces. Because of this, the focus is on the geographical crime forecasts that predict crime hotspots with a reasonable size to patrol rather than specific buildings or places. In some cases, hotspot specific data will be abstracted, but because the focus is on larger hotspots, the generalization does not have a large contact on the usability.

Short-term Crime Forecasting has the necessity to forecast over space and time series information like monthly crime levels across uniform square grid cells at intervals in a town. The grid cells have to be compelled to be as small as achievable, but a mile on a facet, so as to support targeting patrols and alternative police interventions.
During this setting, it is vital to manage the small-area estimation problem; specifically, to search out means that to accurately estimate models primarily based on tiny and thus noisy information aggregates. Information pooling across grid cells, in some kinds, is critical to enhance accuracy. Primarily based on the results of the current analysis is that multivariate models estimated across all grid cells, rather than invert models for every grid cell, maybe the most effective approach.

Multivariate Crime Forecasting - each for the short and long-term attracts on the vast and attractive criminology literature and modeling approaches from the sphere of spatial Econometrics. These literature give appealing theories for controlling fastened effects of place (i.e., crime patterns rely on the character of native populations and land uses, each of that do not amend rapidly over time), for incorporating spatial interactions (e.g., using spatial and time lags to represent crime displacement to nearby areas caused by a crackdown on drug dealing) and for specifying leading indicators to be used in short-term forecasting (a version of the "Broken Windows" theory suggests that "soft crimes" harden over time to become serious crimes).

Geographical Crime Forecasting Techniques

The conventional method of crime forecasting by the police unit was inserting pins on a map on the wall, where each pin stood for a crime. Using this technique, a general idea was created by the crime division for a mark, showing the position of crime hotspots. Compared to the repeat and near repeat victimization theory, it was understood that future crime would happen at those same hot spots. This method, however, is time consuming and not very exact. At the end of the 20th century, the cartography technology was introduced. It had been the mixture of GIS and mapping software and became accessible. In conjunction with the automation
of police records, Geographical Information Systems can be used to ascertain crime patterns digitally.

The pins on a map technique might simply be replicated, ensuring in the point’s map. However, with a large number of crimes it becomes unclear where the precise locations of the hotspots are and what their size is. For instance, if the two crimes happened at a similar place, still just one point is visible. Options are used to counter a number of these problems, like adding up all crimes at the same location and produce a larger point. However, this will not solve all problems. Another methodology created for digitally mapping crime was a boundary (the map divided in totally different areas) or grid- (the map divided into cells of an equivalent size) KDE mapping. Using these strategies, every incident is assigned to a region of the map and every region gets a color representing the quantity of incidents assigned to that area.

The additional incidents besides the strategies that divided the maps into areas, alternative methods were created that visualized distribution of crime during a sleek way, therefore known as continuous KDE maps. With this technique, every crime event is placed on the map like point maps, however currently every point has a certain influence over the area near its location, for instance by using a crime distribution. The influence of an incident spreads around the location where it occurred, with the influence changing into gradually smaller, the farther away from the incident. By stacking the crime distributions of every incident, a density corresponding to the forecasted amount of crime hotspot is created.

1.4 GIS IN CRIME ANALYSIS MAPPING

“Crime mapping” could be a term that has been used for the past few years to sit down with analysis using GIS during a law enforcement
setting. The term crime analysis mapping is employed to explain this method as a result of employing a GIS to investigate a crime is not simply the act of putting incidents on a map, however, conjointly of study. Consequently, “crime analysis mapping” is the method of employing a geographic data system together with crime analysis techniques to concentrate on the spatial context of criminal and alternative law enforcement activity.

**Graphical Information System**

Ever since maps are out that depict the geographic options of communities, like streets and town boundaries, police departments have used such maps to work out patrol areas and emergency routes in addition to assist patrol officers to find specific addresses. Police departments have conjointly mapped crime, a method that, still recently, concerned the manual placement of pins on hand-drawn wall maps. GIS is the emergence of computerized crime mapping as a tool for conducting crime analysis. It begins with an introduction to key terms and then describes basic ideas before presenting a history of crime mapping and knowledge on the field’s current standing and career method.

Definition: A GIS is a set of computer-based tools that allow the user to modify, visualize, query and analyze geographic and tabular data.

A GIS is comparable to a spreadsheet or word processing program in that the software provides a framework and templates for information assortment, collation and analysis and it is up to the user to determine what elements of the system to use and the way to use them. A GIS will moreover enable the user to provide paper maps; it additionally permits him or her to look at the information behind geographic options, combines numerous options, manipulates the information maps and performs statistical functions.
A GIS could be a set of computer-based tools that enable anyone to change, visualize, query and analyze geographic and tabular information.

A GIS could be a powerful software tool that enables the user to do something from an easy purpose map to a three-dimensional visualization of spatial or temporal information. A GIS is totally different from manual pin maps and PC maps in that it permits the analyst to look at the information behind the geographic options, mix varied options, manipulate the information and maps and to perform statistical functions. There are many different kinds of GIS programs that embody desktop packages (e.g., ArcView®, MapInfo®, GeoMedia®, Atlas GIS®, Maptitude®) and also skilled software (e.g., ArcInfo® and Intergraph®).

**GIS Components**

The following could be a description of the five major parts of a GIS that embrace information illustration, information options, visualization, scale, and querying. Real world information is represented by one in every four options in a GIS.

**Features of Geographic Information Systems**

A geographic information system interprets physical components within the universe like roads, buildings, lakes and mountains into forms which will be displayed, manipulated and analyzed alongside police data like crime, arrest, and traffic accident data. There are four categories of geographic features employed in GIS to represent crime objects and spatial locations within the real world; these are named as point, line, polygon and image features.
A point feature could be a discrete location that is typically depicted on a GIS-generated map by an emblem or label. A degree feature is analogous to a pin placed on a paper wall map. A GIS uses totally different symbols to depict the locations of information relevant to the analysis, like crimes, motor car accidents, traffic signs, buildings, police beat stations and cellular phone towers. A line feature is a real-world component that can be symbolized on a chart or map by a line or set of lines.

A polygon feature could be a geographic space represented on a map by a multisided figure with a closed set of lines. Polygons will represent areas as giant as continents or as tiny as buildings; in GIS-generated maps they will be used to depict county boundaries, town boundaries, parks, faculty campuses, or police districts.

An image feature on a GIS-generated map may be a vertical photograph taken from a satellite or an airplane that is digitized and placed among the acceptable coordinates. Such photos, which can seem in black and white or color, show the main points of streets, buildings, parking towns and environmental options (landscaping).

**GIS and Importance of “Hot Spots”**

Geographic or spatial analyses of crime using GIS have established their meaning in each criminological analysis and criminal justice. In recent years, their use in analyzing crime patterns has been viewed as a very important part of the effort in law enforcement agencies toward successful and economical crime management. A theme matter within the geographic analysis of crime that has attracted a lot of attention is that the identification of crime “hot spots,” specifically the locations or small areas within which disproportionately sizable amount of the criminal incidents cluster. Researchers and police use the term in many alternative ways in which, some
talk to hot spot addresses, others talk to hot spot blocks and some others examine clusters of blocks. Compared to researchers, crime analysts discover for concentrations of individual events that may signify a series of connected crimes. They additionally consider small areas that have an excellent deal of crime or disorder, despite the fact that there is also no common offense. Analysts additionally observe neighborhoods and neighborhood clusters with high crime and disorder levels and take a look at them to link these to underlying social conditions.

However, there is no common definition of the phrase hot spot of crime survival, the general understanding is that a hot spot is a neighborhood that features a larger than average selection of crime or disorder events, or a neighborhood where individuals have the next than average risk of victimization. This implies the existence of cool spots or areas but with the typical quantity of crime or disorder. It conjointly suggests that some hot spots could also be hotter than others; that is, they vary in how a higher than average they are.

Identifying crime hot spots incorporates a vital sensible suggestion, as a result, there are accumulating items of proof that ‘the more additional law enforcement efforts are centered on high-crime places or high-crime time, the more additional, successful and economical they might be in controlling crime’.

**Hot Spot Visualization**

The obvious way to visualize a spatial purpose pattern is claimed to plot the information in the shape of a straightforward dot map. This, in flip, can provide an initial sense of perception in the form of the study space and any obvious pattern gift within the distribution of events. In saying this, perceptive concepts regarding what constitutes a ‘random pattern’ is
misleading. Usually, it will be laborious to come back to any conclusions purely on the premise of a feasibility analysis, notably when the info set is large and there are multiple occurrences of events at identical or nearly identical location.

The use of geo referenced data in digital mapping packages and geographical information systems has vastly simplified the method of mapping crime incident data. Systems implementing these concepts are few and far between. The Illinois Department of Justice ‘Spatial and Temporal Analysis of Crime’ (STAC) software is employed extensively in criminal profiling and describing the boundaries of crime clusters as customary deviations ellipses. Two major issues of STAC are an absence of correspondence between the form of the hotspot and therefore the underlying patterns of land use, which STAC is prone to the Modifiable Areal Unit downside.

With recent developments in crime mapping hot spots of any size, from hot spot places to hot regions can be established. Although many of these perspectives on hot spots have something in common which are the concentrations of crime or disorder separated by areas with far less crime or disorder, they differ in the area covered by the hot spots. More importantly, the factors that give rise to hot spot places are different from the factors that give rise to hot spot streets, hot spot neighborhoods, or hot spot cities. Furthermore, the actions one takes to deal with a hot spot place will be different from the actions needed to address a hot spot street, hot spot neighborhood, or hot spot city.

**Hot Spotting Techniques**

Each level has basic units of analysis which is the data being examined. One will think about units as such as the geographic areas being
depicted on maps: points, lines, or polygons. Some criminological theories facilitate making a case for purpose concentrations of crime. Different theories facilitate making a case for linear concentrations of crime or hot spot crime polygons. However, theories of crime are helpful for serving to guide crime and disorder mapping as long as one selects a theory acceptable for the amount of research and action. Crime analysts frequently assume that crime distributions are clustered and whether clusters exist or not, but some are identified from random crime distributions. Testing for clustering is that the initial step in revealing whether or not the information has hot spots of crime. Between these extremes, we have near repeats which take place in an area not more than three or four city blocks in extent.

The kernel density estimation is one of the methods used expansively in analyzing spatial clustering of offenses. It was originally developed in the late 1950s as a method to produce a smooth density distribution from a histogram and the procedure was later extended so that a smooth surface can be fashioned from a point-pattern map. Kernel density estimation has a number of advantages over other methods in searching crime hot spots, as well as its ability of allowing irregular shaped hot spots and its appropriateness for examining changes in those hot spots over time. Hot spots or spaces have been identified using STAC, kernel density analysis or similar techniques. These hot spots commonly cover many city blocks in cooperating more than one census tract. At the other extreme, we have repeated victimization where one address is attacked multiple times to repeat victimization as a “hot place” and combines hot spots and near repeats into “hot space.”

1.5 CRIME MAPPING

Crime mapping is a term used in policing to refer to the process of conducting spatial analysis within crime analysis.
Definition: Crime mapping is the process of using a geographic information system to conduct spatial analysis of crime problems and other police-related issues.

1. It makes possible visual and statistical analyses of the spatial character of crime and other types of events.

2. It permits analysts to connect dissimilar data sources together based on common geographic variables.

3. It provides maps that help to communicate the results of the analysis.

Kinds of Crime Mapping

There are some kinds of crime mapping used routinely in crime analysis. In single-symbol maps, individual, uniform symbols represent options like the locations of stores, roads, or states. A vital factor to stay in mind regarding single-symbol maps is that a GIS places all points on such a map that share a constant address directly on prime of one another, creating it not possible for the map to indicate what percentage points there really are. This drawback of single-symbol mapping is especially relevant for the mapping of crime and different police knowledge, as a result of crime and different police-related incidents that typically occur repeatedly at explicit locations. Thanks to this, crime analysts use single-symbol mapping primarily to show geographic info during which there is no overlap; they use different varieties of maps to convey info regarding multiple incidents at explicit locations. Therefore, analysts use single-symbol maps primarily after they are operating with comparatively tiny amounts of information that do not overlap. Police agencies conjointly typically use single-symbol maps to speak the locations of crimes among patterns to police personnel.
A buffer may be a specified space around a feature on a map. Buffers are set at a less distance, like fifty feet, or larger distances, like five hundred miles, reckoning on the aim and the scale of the map. Buffers facilitate in crime analysis by illustrating the relative distances between options on a map. Buffers may also be used as polygons for knowledge aggregation and comparison.

Crime analysts usually use graduated maps that are, mapping during which completely different sizes or colors of options represent explicit values of variables. During a graduated size map, the sizes of the symbols used for the purpose and line options mirror their worth. As noted on top of, single-symbol maps do not seem to be acceptable for displaying knowledge concerning crimes that occur at constant locations repeatedly. Analysts use graduated size maps for this purpose, as a result of which, these maps will account for multiple incidents at constant locations. However, like single-symbol maps, graduated size maps are subject to overlapping points if too several information are analyzed right away. During a graduate color map, the colors of the symbols mirror their values; this type of mapping is used with points (in a single-symbol map only), lines, and polygons.

Chart mapping permits the crime analyst to show many values among a selected variable at the constant time. There are two kinds of chart mapping: pie and bar. In pie chart mapping, the relative percentages (represented by slices of a pie) of values among a variable are displayed. In bar chart mapping, the relative frequencies (represented by bars) of values among variables are displayed.

In density mapping, analysts use purpose information to shade surfaces that are not restricted to space boundaries (as is the case in graduated color mapping). In their most simple kind, density maps are shaded as per the concentration of incidents above all areas. Such maps are used to match tiny
variations in crime levels from one space to a different space, instead of to match levels of crime among mounted artificial geographic boundaries, as in space maps.

Rather than a kind of mapping, the term interactive crime mapping refers to simplified geographic info systems created for novice users over the web. Several police departments have interactive internet sites where voters and cops will conduct basic crime mapping. These applications generally are not versatile or refined enough to be helpful to crime analysts.

History of Crime Mapping

Mapping itself encompasses a long history. However, crime mapping specifically may be traced back to the first 1800s when social theorists began to form maps parenthetically with their theories and analysis regarding crime. In relation to crime and policing, maps initially were used to look at problems like poverty or demographic characteristics and crime. The first one in all the primary police departments to use mapping was New York town within the 1900s. The maps consisted of straightforward wall maps during which “push pins” were used to point crimes that had occurred. Throughout the 1920s and 1930s, sociologists at the University of Chicago used mapping to look at crime and delinquency, specifically juvenile delinquency and connected social characteristics. Within Nineteen Sixties and Seventies, the primary computer-generated maps of crime were created.

1.6 CRIME CLASSIFICATION

Different crimes receive totally different priorities in terms of justice system response. These priorities are the results of classifying crimes. Several ways of classifying crimes exist. Maybe the foremost common and easiest method to rank crimes is by legislative definition. Criminal offenses
will be classified from least serious to most serious, as violations, misdemeanors, or felons. Similarly, distinctions are drawn between standard and aggravated or dangerous crimes. Generally, agents of the justice can offer additional attention to serious and dangerous offenses than for less serious or standard crimes.

Criminals are classified in a variety of ways. Ways of classification embrace offense seriousness (status of offenders, misdemeanants and felons) and a danger to the general public (non-threatening, threatening, doubtless harmful, or harmful). Criminals are classified as first-time offenders, repeated offenders, or career offenders. Also, they will be classified primarily based upon the kinds of offenses that they commit as burglar, rapist, thief, con artist, shoplifter, hatchet murderer, serial sniper, etc..

1.7 CLUSTERING

Data clustering has been filled with life area of study in statistics, pattern recognition, neural networks and enormous databases but with totally and completely different stress and methods. When observed from a machine learning outlook, clusters correspond to hidden patterns and conjointly the search of clusters corresponds to unsupervised learning and conjointly the following system represents an information concept. Therefore, clustering is an unsupervised learning of a hidden info concept.

Clustering algorithms are attractive for problems that require minimal domain information regarding the class identification and see clusters of arbitrary kind with good efficiency. It is the strategy of forming groups in large databases. The task is aimed to identify clusters embedded among the info, i.e., to cluster the info points into a collection of groups such that, among each cluster the info points are identified as different and are totally and completely different from the info points contained in alternative
teams. The clusters are fashioned on the premise of similarity criterion that is either the closeness of knowledge points or their characteristics.

The larger similarity between the clusters and bigger the excellence between the teams, the upper and more distinct is the clustering. Once a dataset has been partitioned into clusters, it is attention grabbing to understand what the key characteristics of each cluster are. This defines the applying of various techniques, like statistical analysis, classification or association rules. Thus, clustering is typically only one single step of a wider analysis task.

To discover the profiles that share an equivalent pattern of temporal behavior and finding the profiles that are near one another is equivalent to clustering. The clusters of profiles that share similar temporal behavior patterns and are neighbors spatially discovered. Temporal clustering is clustering of all the profiles that show similar temporal patterns or behavior. The profiles that exhibit some closeness and are located near one another end in spatial clustering. Further, Clustering techniques can be categorized into three major groups.

Basic representation of this category of algorithms is k-means. It is based mostly on the concept that a middle purpose will represent a cluster. A middle known as centroid is computed because the mean or median of a cluster of points and therefore, sometimes it is not some extent of the dataset. Then, every information purpose is assigned to the cluster whose centroid is the closest. Once all points are assigned, the positions of k-centroids are calculated once more. This is often repeated until all the centroids for two consecutive iterations stay equivalent. The foremost common strategies of partitioning algorithms are k-medoids and k-means and these tend to find spherical shapes in knowledge.
Hierarchical clustering is predicated on the concept of performing a collection of division operations. The two balancing approaches to the computation of a hierarchical clustering are (a) Agglomerative Clustering (bottom-up) considers every individual data information purpose as a separate individual cluster and then uses the combining strategies recursively to induce the specified range of separate clusters. Therefore, essentially, it starts with singleton clusters and at every step the two clusters are merged. (b) Divisive Clustering (top-down) that, conversely, starts from an all inclusive cluster and at every step a cluster is split into two elements.

This relation among algorithms fully changes the concept of the cluster. At this point, the clusters are inhabited not by sets of objects near one another however by objects which may reach one another through densely populated regions. This includes density-based and grid-based algorithms. Primarily Density based algorithm DBSCAN considers dense region of point’s united attention-grabbing cluster or region. Primarily, Grid based methodology quantizes the clustering area knowledge points into grids and then performs operations on these grid cells to create clusters. The wants of clustering strategies in knowledge mining embrace scalability, ability to modify totally different attributes, ability to find clusters of arbitrary shapes, minimal demand of domain information to seek input parameters, ability to modify noise and outliers, insensitivity to the order of input records and high dimensionality.

1.7.1 Spatial Clustering

Spatial clustering is obtained by constructing a spatial hierarchy. Quad-tree and R-tree are typical hierarchical structures. In spatial data sets, clustering permits a generalization of the spatial element that enables successful data mining. There are quite a variety of algorithms that are developed for spatial clustering PAM (Partition Around Medoids) could be a
k-clustering algorithm that works well for small data sets. To search out clusters, k-clustering on medoids is used. This is often terribly expensive for large data sets. CLARA (Clustering giant Applications) deals with large data sets. It works by making multiple samples of the info and then applies PAM to every one of the samples. CLARANS (Clustering giant Applications primarily based on Randomized Search), motivated by PAM and CLARA relies on randomized search. It has been well studied and its potency and effectiveness is proved comparing with different algorithms. In CLARANS, a cluster is represented by its medoids or the foremost centrally located knowledge purpose within the cluster. CLARANS was followed by DBSCAN that could be a locality-based algorithm betting on the density of objects for clustering. BIRCH and CURE algorithms create use of each hierarchical techniques and grouping of connected things. BIRCH is appropriate for large databases and finds smart clusters with one scan of the database. It is capable of dynamically clustering multi-dimensional knowledge points.

It is often seen from the on top of sections that a major subset of analysis within the field data mining appearance at the actual semantics of the area and time and also the manner within which they will be accommodated into data mining algorithms. Most of the work is often placed in one in every three categories: 1. Temporal Association Rule Mining, 2. Spatial Association Rule Mining, 3. Spatial clustering and Time-series Analysis. Most of the on top of work deals with either spatial or temporal semantics, with only a few handling each. In fastened boundary cells are used to model spatial-temporal knowledge. During this approach, the matter is simplified by dividing it into a standardized grid of cells that have predated R-trees. R-trees are usually used for indexing spatial-temporal knowledge. However, the overlapping nature of R-trees, that causes backtracking in search, is not best suited to each application.
1.8 SPATIAL DATA MINING

A large range of techniques is contributing to the mining of spatial databases. These techniques give results regarding spatial classification, spatial trend detection and spatial clustering. The aim of spatial data mining is to reveal unknown data. Two issues associated with clustering that have received attention in the recent past years include:

1. generalization-based characterization applied to spatial information proposed in;

2. a cluster characterization primarily based on spatial options like lakes, resorts etc.

1.8.1 Temporal Data Mining

It is a rapidly evolving space of analysis that is standard in many disciplines as well as statistics (e.g., time series analysis), temporal pattern recognition, temporal databases, optimization, high-performance computing, and parallel computing. The temporal data mining element of the Knowledge Discovery in Temporal Databases method is bothered with the algorithmic means that those temporal patterns are extracted and counted from temporal knowledge. A comprehensive overview of the techniques is presented for the mining of temporal data. Main temporal data mining techniques that are commonly derived by either bottom-up or top-down induction are: temporal classification (basic goal is to predict temporally connected fields in an exceedingly temporal database primarily based on different fields); and temporal clustering. In temporal analysis, several temporal data mining applications create use of clustering. There are two basic approaches to temporal clustering and analyzing it. One is that the live temporal similarity approach and the other is a temporal optimal partition approach. If the amount of clusters is given, then clustering techniques will be divided into three
classes: (i) Metric-distance primarily based technique, (ii) Model-based technique and (iii) Partition-based technique. If the amount of clusters is not given, then we are able to use Non-hierarchical Clustering Algorithms to seek out their k.

1.9 GEOGRAPHIC SIMULATION MODELS WITH GIS

Many researchers had observed the importance of mixing GIS with dynamic and analytical models. In recent times, GIS has been extensively used for sustaining dynamic geographic model building and analysis in many geographically connected fields, like ecology, urban designing, and hydrology. The role of GIS in exploring space-time processes as a platform, a controlled surroundings or laboratory that stressed its capability in visualizing and knowledge managing, (Takeyame 1997) building crime analysis modeling into GIS directly, by showing GIS-based simulation as a sort of map dynamics that are formulated as a map equation.

The existing GIS applications are restricted to mapping, data locating or querying, and low level spatial analysis, like distance calculation, path optimizing, overlay and buffering. When reviewing this standing of linking GIS and crime forecasting models, special data structure and mathematical functions to implement the crime forecasting models are needed. Yet, it is conjointly time consuming and unnecessary to re-develop spatial data structure and spatial analysis functions with simulation models. The two major approaches are identified by their use and not fully their economics GIS software.

Economic GIS software

In this software, there are two selections of developing spatial simulation models, loosely-coupled integration or tightly-coupled integration with GIS. If the economic GIS software meets the necessities of building a
simulation model, a tightly-coupled integration has the best resolution. Beneath this instance, spatial simulation model is about to be represented in GIS macro language, like Arc data AML, or Arc read Avenue. When economic GIS software could not handle the complexity of the spatial simulation model, additionally the model also wants some basic spatial information management show and analysis, a loosely-coupled approach generally is sometimes counseled. Loose-coupled integration develops the simulation model with C, C++ or different programming languages and connects it with economic GIS software. GIS saves the efforts to develop a spatial information reading and analyzing system.

**Non Economic GIS Software**

It additionally has two sub-categories embedding GIS-like functionalities into simulation modeling packages, or coming up with the simulation model with simulation software. Whereas system developers relish lots of freedom in system vogue and model construction, they need to write down programming code not only for the simulation model itself, but additionally the spatial information visualization and analysis which may be done by GIS software. Generally, the GIS-like functionalities developed at intervals of the modeling packages, are very restricted. Another technique is to form use of simulation software, like STELLA, EXTEND, Simulab and VenSim, which could facilitate scientists merely to form their simulation models.

A simulation model needs more time and testing because it must accurately replicate the real processes or case of flow happening in your modification system. This includes identifying all possible sources of new access and reassess into the system and each important decision point that could change the preservation or release of an individual. The length of time between decision points must also be specified in order to know when these
would occur in a given time period. The necessities of knowledge grow geometrically because the variety of subpopulation characteristics (e.g., people by offense sort, by risk classifications, by gender, etc.) increase since every one may be a separate element of an overall system. Once a valid simulation model is developed, it is used for assessing varied alternatives that have an effect on how briskly and in what manner private moves through the corrections system.

The following two reasons recommend for a Crime forecasting and crime pattern simulation model and GIS.

The complexity of crime forecasting model demands additional resourceful functions that are not offered in GIS. Thus, model illustration and major simulation spatial computation need to be done outside GIS.

It is possible. GIS software will be customized because the simulation interface, which may call the external simulation program to execute behind the interface. It makes the mixing additionally convenient for researchers to check the model. No additional data conversion, input and output are required. The model interface works independently.

1.10 PROBLEM DEFINITION

Crime may be a behavior deviating from traditional violation of the norms giving people’s losses and harms. Social, psychological, economic and environmental factors are to be thought-about in crime issues. These ideas have an effect on the occurrence of crime in numerous ways that, the Stakeholders who have roles in crime prediction are police, native governments, law enforcement agencies and folks exposed to crime and offenders.
Crime is tried to be explained by numerous theories from totally different sciences. Social and psychological theories contemplate the foundation causes of crime noticing factors like social disorganization, temperament disorders and inadequate parenting etc. However, such theories are inefficient to be used for crime prediction and clarification. The new advance in criminology science has necessary ideas which will guide crime prevention and crime analysis efforts.

Crime prevention could be a vital issue that people have been managing for hundreds of years. Preventing crime could be a necessity to create a more peaceful world for individuals to live in. To realize additional peace and secure life, police is the most accountable foundation for crime prevention by deploying the resources optimally. Police use strategic, tactical and administrative policies that assist the required precautions before the occurrence of a criminal activity. To create effective policies and to improve prevention techniques, police ought to strengthen the use of criminological theories and crime analysis.

In the scope of this problem classification are used to identify the crime attributes based on temporal activities like space and time factors and mapping the crime attribute events into hot and cold spot maps, additionally it was used for preprocessing crime mapping incidents. The clustering analyses are used to identify hot spots. Cluster analysis aims at collecting data into groups according to several algorithms. The two main groups of clustering occurring in spatial data analysis are clustering and partitioning approaches. Partitioning approaches use optimization procedures to divide data into meaningful groups. Additionally, to find the exact hotspot locations where crime occurs more, Brich, Clarns, DBScan and clique are the examples of partitioning subspace approach with different algorithms. Also, there are new clustering techniques generated for specific purposes. One of them is Spatial-
Temporal Analysis of subspace partitioning approaches collects all the density hot spots from the KDE map and find the exact hot spot based on space, time and size of the hotspot. After finding the exact crime hotspot, three types of micro, me-so and macro computer simulations to find the long and short time forecasting crime offenders are applied.

Another issue in crime analysis is crime forecasting, whether it directly supports crime prevention and law enforcement or not. Developing highly reliable ways of forecasting future crime trends and issues is one amongst the foremost and most popular ways in which to boost crime prevention and reduction measures. With the advance of crime forecasting, spatial and temporal predictions of crime have been used to form long and short term planning. Within the scenario of obtaining correct predictions, it is potential to manage security resources efficiently. Police provide attention on highlighted areas, target patrols, allocate resources and do alternative police interventions to prevent crime.

Crime forecasting is often examined in three different concepts; spatial, temporal and spatial-temporal. Spatial-temporal crime forecasting could be a new and a preferred subject studied by geographers and crime analysts which incorporates methods from modeling to data mining.

1.11 STRUCTURE OF THE THESIS

This thesis has been arranged in seven chapters. A brief description of each chapter is given below.

Chapter 1: Provides an introduction to the crime analysis, crime forecasting and its techniques GIS and how it supports to the crime mapping and why it gives more importance to the hotspot and how the crime attributes are classified and clustered and how it works in the spatial clustering, spatial data

**Chapter 2:** This chapter reviews different types of subspace clustering algorithms, classification algorithms which are more related to this study. A number of classification and clustering algorithms are briefly reviewed to show their parameters and application areas. The advantages and disadvantages of such algorithms are also described. The procedure involved in forecasting simulation is reviewed.

**Chapter 3:** Provides a brief review of literature about the Crime Analysis, Crime Forecasting, Crime Classification, Spatial Clustering Algorithm, Geographic Simulation, Hot spotting techniques and Visualization that are presented and discussed, and also provides the scope of the research work.

**Chapter 4:** It demonstrates the basic concepts for classification algorithm with police crime data. The proposed structure classification algorithm is a data mining based classification algorithm. It is used to find classes of crime quickly. Performance evaluation of the algorithm is presented.

**Chapter 5:** It demonstrates the proposed spatial clustering based CLIQUE Optimization algorithm for identifying crime hotspots. Optimization based splitting is used for handling of the problem. In this algorithm, recursive optimization technique is used for identifying crime hotspots based on crime density and presenting an evaluation of the algorithm.

**Chapter 6:** It provides new forecasting simulation functions of police decision making and supporting crime analysis and crime mapping. The proposed macro, micro, me-so level approaches simulated for crime forecasting short, medium, long term analysis.
Chapter 7: This chapter provides a conclusion of the study by summarizing the findings of the research. And it also reviews how well the aim and objectives have been fulfilled. The chapter finally looks up to the possibilities for future research and ends with some concluding statements.