ILLINOIS
CRIMINAL JUSTICE
INFORMATION AUTHORITY

Research Bulletin

Spatial and Temporal Analysis of Crime

April 1987

A computerized crime analysis program and accompanying methodology are being developed by the Illinois Criminal Justice Information Authority to give Illinois law enforcement agencies an easier and more accurate means of finding patterns in the times and locations of crimes.

This *Research Bulletin* describes the first stage of the Spatial and Temporal Analysis of Crime (STAC) project: the development of the theories and methods behind STAC and the first tests of the computer programs on selected data. The next stage of the project will involve testing the programs in a working law enforcement environment and further refinement of the methods and programs.

Many law enforcement agencies use *pin maps* as their chief crime analysis tool. Pins or markers are placed on a map of the community to represent the locations of crimes and to show other information, such as the times that incidents occur. The advantage of pin maps is that they enable law enforcement agencies to examine the distribution of different types of crime throughout their communities.

But pin maps have shortcomings as well. One of the most important is the considerable time they take to prepare and maintain. In addition, the analysis of patterns on these maps is purely visual; patterns of information other than location can be difficult to identify. Finally, the amount of information on a pin map can be so extensive that it defies easy analysis.

The Illinois Criminal Justice Information Authority undertook its Spatial and Temporal Analysis of Crime (STAC) project to develop clearer, more efficient ways of analyzing and representing the kind of data traditionally shown on pin maps, and even to go beyond what pin maps can show. With a grant from the federal Bureau of Justice Statistics, the Authority has created a group of computer programs and analytical methods for detecting patterns of crime in a community, using both geographic *and* time data.

Both the programs and the methods could be very useful to law enforcement agencies. The computer programs condense large amounts of crime information into a manageable form. The analytical methods could be used in resource allocation, crime analysis, beat definition, and other applications.

The STAC program and methods were developed and tested using data on six types of crime (based on definitions taken from the Uniform Crime Reports) in four communities in the Chicago area[1] (see Table 1). The following crime types were chosen because they are all observable and thus possibly preventable by police patrol: robbery, criminal damage to property, motor vehicle theft, nonresidential burglary, residential burglary, and theft from motor vehicles. The four communities, all of which are users of the Authority's Police Information Management System (PIMS), cross a wide range of population densities and crime levels. Thirty months of data were obtained for each crime and town, spanning the period between January 1983 and June 1985. The data used for each incident consisted of the location of the incident[2] and the day and time of the incident's occurrence.[3]

This bulletin describes the programs written to analyze time and space data and discusses the results of the test analyses for each type of data. Methodological issues and potential applications for law enforcement are also discussed.

Table 1

Populations of Test Communities

Town #1	73,000
Town #2	52,000
Town #3	66,000
Town #4	22,000

Crime Types Analyzed

Robbery
Criminal Damage to Property
Motor Vehicle Theft
Nonresidential Burglary
Residential Burglary
Theft from Motor Vehicles

Temporal Analysis

The basic purpose of the STAC time analysis program is to categorize incidents by their time of occurrence. The program analyzes two data categories — the day of the week and the hour of the day.

The program describes the distribution of a given group of incidents over an entire week by calculating the percentage of those incidents that occurred on each day of the week (Sunday, Monday, etc.). Similarly, the program calculates the percentage of incidents that occurred during each of 12 two-hour intervals (midnight to 2 a.m., 2 a.m. to 4 a.m., and so on).[4] This way, it is possible not only to analyze the distribution of incidents over the entire day, but also to compare levels of activity at different times of the day (see Figures A and B).

The temporal analysis program can also combine these two categorizations and produce a time-of-day distribution for each day of the week. For example, if 25 percent of all incidents occurred on Wednesday, the program can show what percentage of that 25 percent occurred in each two-hour period on Wednesday. It will do the same for the incidents that occurred on each of the other days as well. This yields a distribution of incidents over the week at the hourly level, giving a more detailed picture than the day-of-week distribution.

Note that because no prior assumptions were made about the distribution of incidents over time, the test analyses investigated only the extent to which the crimes were randomly distributed. Any departure from a random distribution (in other words, a large concentration of crimes in a short period) was therefore considered to signify a pattern. When the temporal analysis is combined with other information about the crimes, however, more significant results may be achieved. In particular, if a rough distribution over time is already known for some type of crime, the temporal analysis can be used to check whether a specific group of incidents matches that distribution.

This research bulletin was written by Authority research analyst Samuel Bates. It was edited by Maureen Hickey.

Results

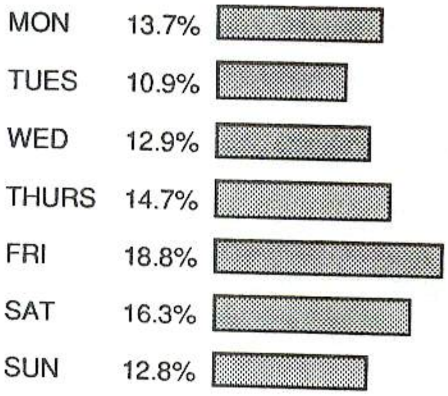
Results of the temporal analysis were calculated for each of the six crimes in each of the four communities. The 2.5-year period was analyzed as a whole and, when the number of incidents was large enough for reliable analysis, each three-month quarter (January-March, April-June, etc.) was analyzed separately. Figures A and B provide examples of the output generated by the temporal analysis program.

The results discussed are illustrative of the capabilites of the temporal analysis program. Actual analyses can only be performed after further testing of the program.

Robbery. The small number of robberies made interpretation of the results difficult. Only town #1 had enough robbery incidents for a reliable pattern of incidents per hour of the day to emerge. The pattern was nevertheless clear: more than half of all robberies occurred in the late afternoon and evening (see Figure B). Although the robbery data from the other three communities were too scanty for reliable analysis, the time-of-day data available for those towns fit the pattern for town #1.

Figure A

Sample output of the temporal analysis program shows day-of-week distribution for robberies in town #1 between January 1983 and June 1985.



267 incidents were used in this analysis. Percentages do not add up to 100 because of rounding.

No pattern for the day-of-week distribution of incidents was obvious. Only two towns had enough data for analysis (see "Quantity of Data," page 4). In one, 52 percent of the incidents occurred between Thursday and Saturday (see Figure A). In the other, 40 percent occurred on Monday and Tuesday.

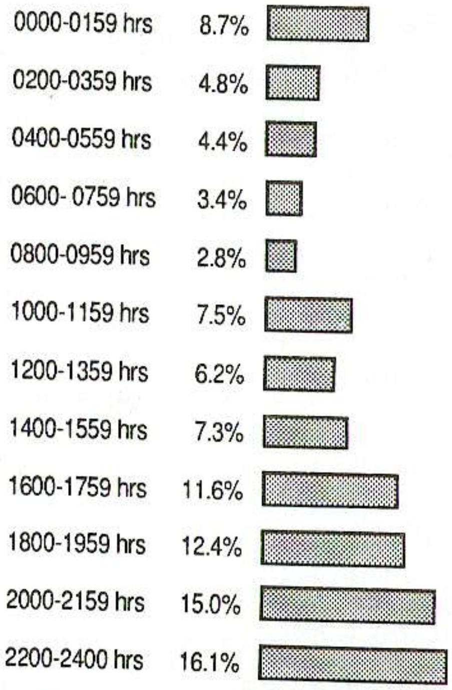
Criminal Damage to Property. Unlike the results for robbery, the time-of-day distribution for criminal damage to property in each town showed no particular pattern; the incidents were scattered fairly evenly over the entire day.

The day-of-week distributions were somewhat less uniform than those for time of day. Nearly half of all incidents took place on the weekend in all four communities.

Motor Vehicle Theft. The distributions of motor vehicle thefts also presented no clear pattern. The spread across hours of the day and across days of the week was nearly uniform.

Figure B

Sample output of the temporal analysis program shows time-of-day distribution for robberies in town #1 between January 1983 and June 1985.



267 incidents were used in this analysis. Percentages do not add up to 100 because of rounding.

Nonresidential Burglary. The analysis revealed a clear pattern in the occurrence of nonresidential burglaries: at least half of all incidents in each town took place in the late evening and early morning.

There were some differences among the towns. In towns #1 and #3, more than 55 percent of the incidents happened between 10 p.m. and 6 a.m., whereas in the other two towns the peak period was between 2 a.m. and 10 a.m. One explanation for this difference may be that in towns #1 and #3, a smaller percentage of the times recorded for incidents was precise than in towns #2 and #4 (see "Data Precision," below).

The overall patterns for nonresidential burglary, however, were consistent over the entire period of analysis in each town, supporting the hypothesis that burglars try to pick times during which the property is likely to be vacant. The day-of-week distributions were fairly even in all four communities.

Residential Burglary. The time-of-day distributions for residential burglary showed a definite pattern. In each of the towns, between 57 percent and 70 percent of all incidents took place in the afternoon and early evening. This pattern complemented the one for nonresidential burglary and further supported the hypothesis that burglars tend to operate when property is likely to be vacant. Once again, the day-of-week distributions were quite uniform, although they were slightly skewed toward the weekend.

Theft from Motor Vehicles. The results were mixed for this crime. In town #3, and in towns #1 and #2 in 1984 and 1985, most incidents occurred in the late evening and early morning. Incidents in town #4, and in town #2 in 1983, took place more frequently in the morning and early afternoon. In all cases the day-of-week distributions were quite smooth.

Overall Patterns. The temporal analyses yielded a few clear patterns. The results for both residential and nonresidential burglary support the reasonable hypothesis that burglars tend to pick times when property is likely to be vacant.

Table 2

Percentage of crimes for which the period between earliest and latest possible occurrences is less than one hour.

	Town #1	Town #2	Town #3	Town #4
Robbery	**87%**	**100%**	**100%**	**100%**
Criminal Damage to Property	**59**	**62**	**64**	**87**
Motor Vehicle Theft	**21**	**74**	**46**	**65**
Nonresidential Burglary	**24**	**68**	**41**	**60**
Residential Burglary	**47**	**55**	**36**	**54**
Theft from Motor Vehicles	**18**	**55**	**34**	**45**

Residential burglary tends to take place in the late afternoon and early evening, before many people arrive home from work, and nonresidential burglary is more common in the late evening and early morning, when businesses are generally closed. A strong pattern also emerged for robberies in the one town examined: more than half of all incidents took place in the late afternoon and evening hours.

The remaining three crimes (criminal damage to property, motor vehicle theft, and theft from motor vehicles) displayed no immediate patterns. The study was restricted to looking only for deviations from randomness, however, and did not include any assumptions about likely distributions.

None of the patterns observed may be surprising. But they indicate that temporal analysis has the potential to yield more significant results, especially when combined with other analyses or performed on more strictly defined data.

Issues

Data Precision. One problem in analyzing crimes for temporal patterns is that often the exact time of occurrence is not known. For instance, if a burglar breaks into a house at 10 a.m., the crime may not be discovered until the resident returns from work at 5 p.m. The police then record the crime as having happened between, say, 8 a.m. (when the resident left for work) and 5 p.m.

This time period between the earliest and latest possible occurrences of a crime should be taken into account in any analysis of temporal patterns. For this analysis, if the time period between the earliest and latest possible occurrence of an incident was less than one hour, it was considered "precise."

As might be expected, different crimes had different levels of precision (see Table 2). In each of the four towns, at least 87 percent of all robberies were precise. By contrast, at most 55 percent of all thefts from motor vehicles were precise. In three of the towns, less than 46 percent of the times of thefts from motor vehicles were known precisely.

Not only did the precision of the data differ among crimes, but there were further variations from town to town. For some crimes, this variation was small; the amount of precise robbery data ranged from 87 percent to 100 percent. For others, the variation was quite large; with motor vehicle theft, it ranged from 21 percent in town #1 to 74 percent in town #2. There are many factors that could explain the variation, such as different recording practices or varying numbers of incidents.

These percentages were based on a one-hour time range as the definition of precision. A longer time range would yield higher percentages, as more incidents would be included.

Table 3

Percentage of crimes for which the period between earliest and latest possible occurrences is less than one day.

	Town #1	Town #2	Town #3	Town #4
Robbery	98%	100%	100%	100%
Criminal Damage to Property	81	94	92	96
Motor Vehicle Theft	80	95	88	91
Nonresidential Burglary	74	95	81	90
Residential Burglary	78	92	79	87
Theft from Motor Vehicles	74	94	89	90

In Table 3, the maximum time period between the limits of occurrence of an incident was one day. As the table shows, at least 74 percent, and usually more than 90 percent, of all crimes are positively known to have occurred within a 24-hour span.

There are several ways of dealing with imprecise data. One of the common methods is to assume that the crime occurred at the middle of the specified time interval. Thus, for example, an auto theft that took place sometime between 8 a.m. and 4 p.m. would be treated as if it had taken place precisely at noon. This method not only ignores the issue of the quality of the data, since it treats all incidents as equally precise, but can also give seriously misleading results for crimes that have a low degree of precision.

The method used in STAC assumes that the crime could have happened with equal likelihood at any time within the interval, rather than assigning it to any one time. Each hour during the interval is assigned a probability that the crime occurred in that hour.

For instance, if a crime occurred between 10 a.m. and 1 p.m., a .33 probability would be assigned to each of the three hours during that interval (10 a.m. to 11 a.m., 11 a.m. to noon, and noon to 1 p.m.). This has the effect of giving less weight to imprecise data than to crimes whose times are known more exactly.

A crime that took place in an eight-hour span contributes only a .125 probability to each hour in that span, whereas a crime that is known to have occurred during a specific hour contributes a probability of 1 to that hour.

Quantity of Data. The small amount of data available also creates a problem in temporal analysis. Crime analysts are often interested in quick detection of patterns, based on data for a week or a month. For some crimes, this time restriction results in very few incidents for analysis.

The test data for nonresidential burglary provides a good example. In the largest of the four communities, the average number of incidents per month was 28, while the smallest town averaged only one per month. Such small numbers can produce highly variable results, since one or two unusual cases can radically change the distribution of incidents over time.

How many incidents, then, are necessary to avoid this problem? There is a common statistical rule of thumb for analyzing data that falls into categories, like days of the week. According to the rule the analyst needs enough data to expect at least five cases in each category.[5]

A minimum of 35 incidents would therefore be required in order to analyze the distribution of incidents over the days of the week. At least 60 incidents would be required for the time-of-day analysis (based on 12 two-hour periods per day), and at least 420 incidents would be required to analyze the time-of-day distribution for all seven days.[6]

It is clear that there will often be too few incidents for each of the possible analyses, as Table 4 illustrates. A possible solution is to reduce the number of categories, in this case time intervals, by combining one or more of them into one category. Thus in the day-of-week analysis, Friday and Saturday could be combined into one category and Tuesday and Wednesday into another. This would reduce the number of categories to five: Sunday, Monday, Tuesday-Wednesday, Thursday, and Friday-Saturday. According to the rule of five cases per category, a minimum of 25 incidents could then be analyzed, instead of the 35 needed to analyze all seven days as separate intervals. There is a tradeoff, though; with fewer categories, the analysis is less detailed and the subsequent conclusions more general. In this case, it would be impossible to distinguish a crime that occurred early Friday morning from one that occurred late Saturday night.[7]

Applications to Law Enforcement

There are many potential applications of temporal analysis to law enforcement. The simplest would be to use the program to illustrate the distribution of various types of crime over the hours of the day or the days of the week in a particular time period. (The period chosen for the test analyses was three months.) These charts could be produced for several consecutive time periods, yielding a picture of how the distribution of crimes changes over time.

For example, Figure C shows the distribution of thefts from motor vehicles in town #1 over the four quarters of 1984. Although the distribution over the whole year was quite uniform, the breakdown into quarters shows how the peak period moves between time periods. The information garnered from such charts could help officials allocate resources at various times of the day or week.

Table 4

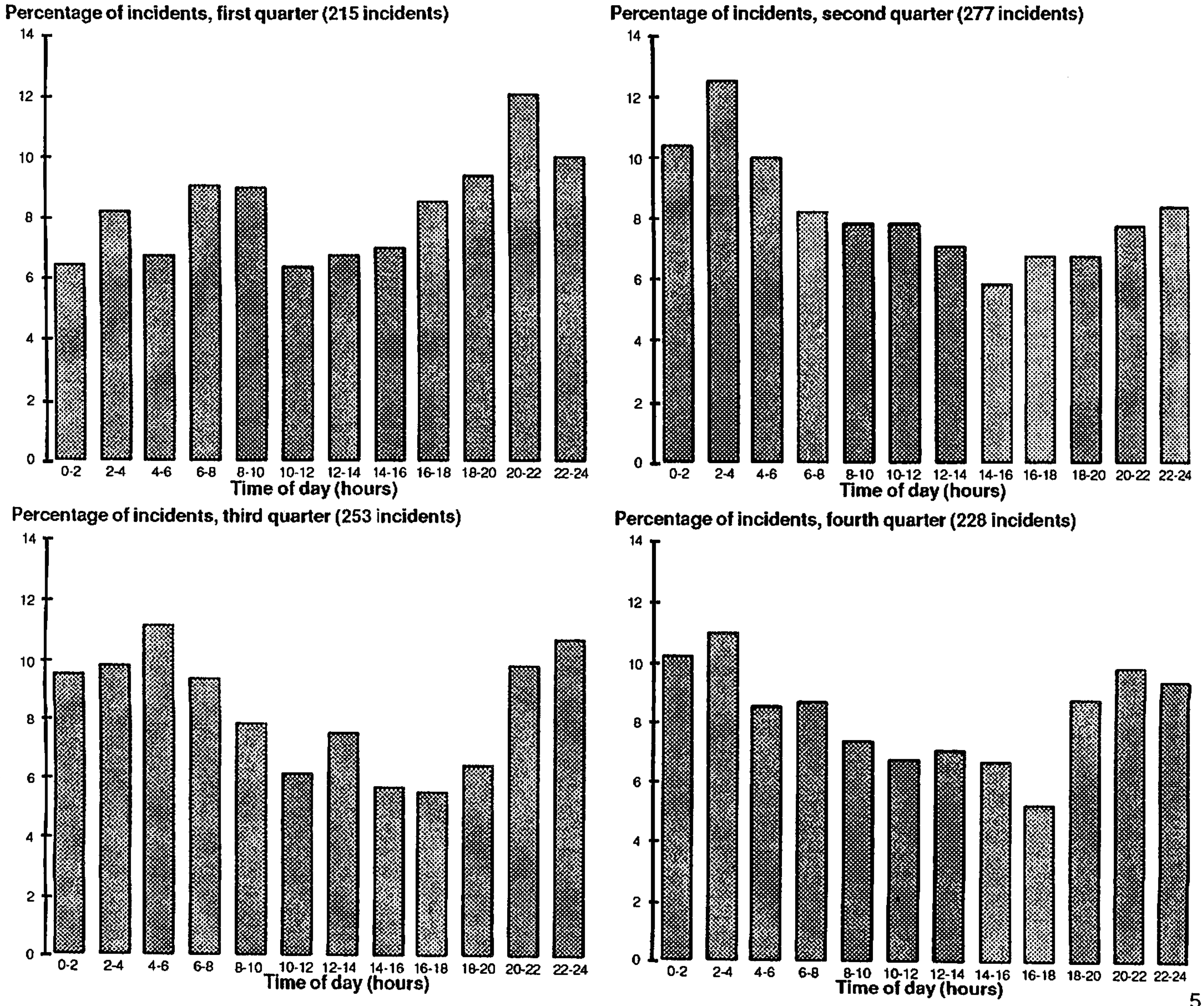
Number of incidents in each town between January 1983 and June 1985.

	Town #1	Town #2	Town #3	Town #4
Robbery	269	35	28	9
Criminal Damage to Property	2,365	2,050	2,204	796
Motor Vehicle Theft	436 *	337	309	57
Nonresidential Burglary	509 *	179	329	40
Residential Burglary	2,589	563	740	125
Theft from Motor Vehicles	1,226 *	1,036	918	215

**Does not include 1983 data, since there appears to have been a change in reporting practices.*

Figure C

Temporal analysis results for thefts from motor vehicles in town #1.



Percentage of incidents, first quarter (215 incidents)

Percentage of incidents, second quarter (277 incidents)

Percentage of incidents, third quarter (253 incidents)

Percentage of incidents, fourth quarter (228 incidents)

This information could also be used to determine the best times to change shifts. The distributions show the periods of least activity, at which times any loss of service due to a shift change could be minimized. The analysis could have further administrative uses as well, such as investigating the effectiveness of a particular police program in reducing crime levels.

Spatial Analysis

The spatial analysis program examines the geographic distribution of incidents and displays information about those crimes that law enforcement officials can use for strategic and tactical crime analysis. Unlike the temporal analysis program, there are many different types of spatial analysis, and the program reflects this diversity.

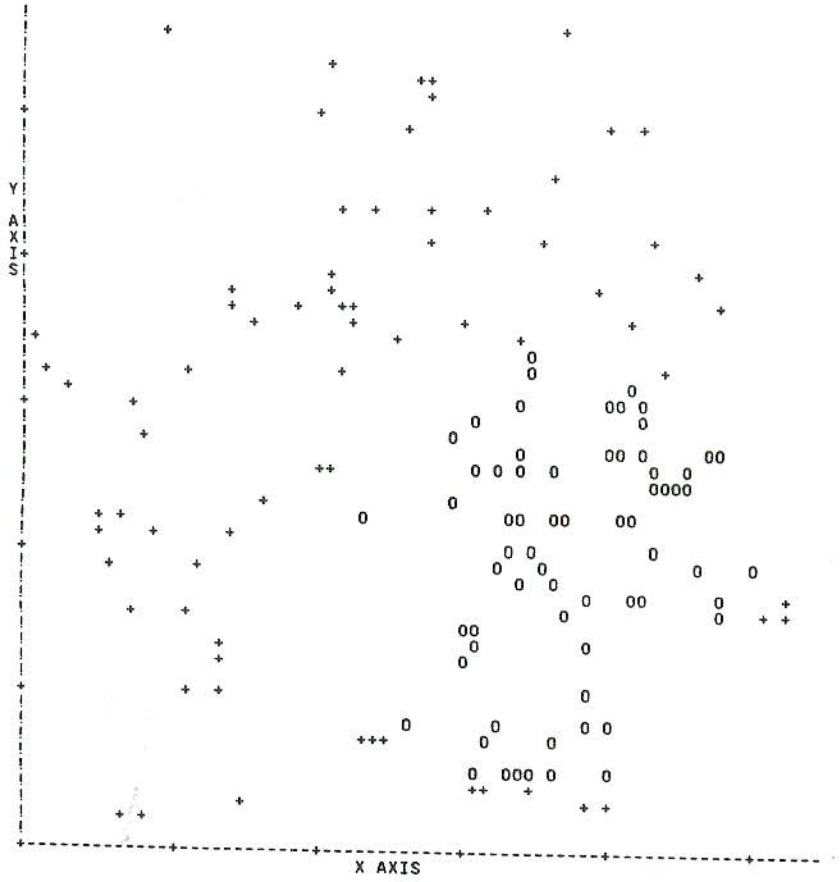
The spatial analysis program has many distinct capabilities, all geometric in nature. These include outlining the boundary of a group of incidents, finding out how much crime has occurred in a given area, and the *hot spot* procedure, which finds the portion of a given area in which the largest number of incidents is concentrated. This section of the bulletin concentrates on the types of information that can be obtained from the hot spot procedure and its component sub-procedures, and describes several potential uses of these procedures.

Capabilities

The hot spot procedure consists of a combination of two separate procedures, *radial search* and *scanning*. Each of these procedures can be used separately to provide certain types of information, or can be used in combination with other procedures for more complex analyses.

Figure D

Sample output of the spatial analysis program: ovals mark the "hot spot" for incidents in town #1 between January and March 1983, plus signs indicate incidents outside the hot spot. (Scale = 2,000 feet/inch.)



Radial search is the fundamental component of many of the spatial analysis procedures. It allows the user (or another STAC procedure) to specify a location and a distance, then finds all the incidents within that distance from the location. The program creates a circle of a given size and centers it on the specified point. It then prints out the locations of all incidents that occurred within the circle.

For example, the program can list all incidents happening within 100 yards of a train station or within half a mile of a schoolyard. These incidents can also be printed on a map showing their positions relative to other incidents in the area. Figure D is an example of one type of output of the spatial analysis program, with the incidents within a 3,000-foot radius of a given point marked by ovals and the incidents outside the circle marked by plus signs.[8]

The user can also instruct the program to vary the radius, thereby producing a set of concentric circles centered around a specified point. The user can specify that the outermost circle contain a certain maximum percentage of incidents or have a certain radius. The program then gives the number of incidents and percentage of total incidents in each of the concentric circles (see Figure E).

The scanning procedure allows the user to investigate the distribution of incidents over a wide area. This procedure lays a grid of points over a selected area. It then does a radial search centered at each point of the grid. The program displays the number of incidents in each circle, giving an overview of the amount of crime in the area.

Figure F shows an example of the scanning procedure output. The program first created circles with a radius of 2,000 feet around every point on a grid laid over a selected area and then counted the number of incidents that had occurred within each circle. The incidents are marked by plus signs, and at each grid point either a small circle or a letter appears denoting the number of incidents in the circle. A small circle indicates no incidents, an "A" indicates one incident, a "B" two incidents, and so on.[9] The area to be scanned can be as large as an entire town or as small as a city block. The spacing between grid points and the radius chosen for the radial searches, both of which can be chosen by the user, determine the level of detail.

Various program capabilities can also be combined to get different types of information. The hot spot procedure is the result of one such combination. Given a radius, this procedure finds the one circle containing the largest number of incidents. In other words, it locates the highest crime density, or hot spot, within a larger area (see "How STAC Finds Hot Spots," page 8).

The hot spot procedure was performed for each crime in each town in each three-month period over the 2.5 years studied. This yielded 10 hot spots for each crime within each town.[10] The analysis indicated not only the area of highest crime density in each three-month period, but also shifts in this hot spot over time.

Issues

Length of Radius. The most important factor in the hot spot procedure is the radius given for the search circle. This radius determines the size and spacing of the search grid. More importantly, however, the length of the radius strongly influences the ultimate location of the hot spot. As the procedure was being tested, it was observed that different radii found different hot spots. In Figure G, for example, hot spots found by searches with radii of 1,500 feet and 2,000 feet, respectively, are more than a mile apart. This does not indicate that the hot spot procedure is inconsistent in its results; rather, searches with different radii are in essence uncovering different aspects of the data. For instance, further examination of the data reveals that the 1,500-foot radius found a hot spot surrounded by an area of low crime density compared to the hot spot itself. When the search radius was extended to 2,000 feet, the total area covered by the search circle had a crime density that was not exceptionally high compared to other areas. The program therefore found a different hot spot with a 2,000-foot search radius.

Figure E

Sample output of the spatial analysis program: results of concentric circle searches in town #1 between January and March 1983.

Maximum percentage of incidents per circle specified

```
70.0%  (  148 incidents)  fall within  5640.  feet
60.0%  (  127 incidents)  fall within  4664.  feet
50.0%  (  106 incidents)  fall within  3419.  feet
40.0%  (   84 incidents)  fall within  2850.  feet
30.0%  (   63 incidents)  fall within  1917.  feet
10.0%  (   21 incidents)  fall within  1579.  feet
```

Radius of each circle specified

```
43.1%  (   91 incidents)  fall within  3000.  feet
35.1%  (   74 incidents)  fall within  2700.  feet
28.0%  (   59 incidents)  fall within  2400.  feet
23.2%  (   49 incidents)  fall within  2100.  feet
15.6%  (   33 incidents)  fall within  1800.  feet
 9.5%  (   20 incidents)  fall within  1500.  feet
 8.5%  (   18 incidents)  fall within  1200.  feet
 5.2%  (   11 incidents)  fall within   900.  feet
 1.9%  (    4 incidents)  fall within   600.  feet
  .5%  (    1 incidents)  fall within   300.  feet
```

Figure F

Sample output of the spatial analysis program: letters mark the centers of radial searches. A = 1 incident within the circle, B = 2 incidents, etc. (Scale = 2,000 feet/inch.)

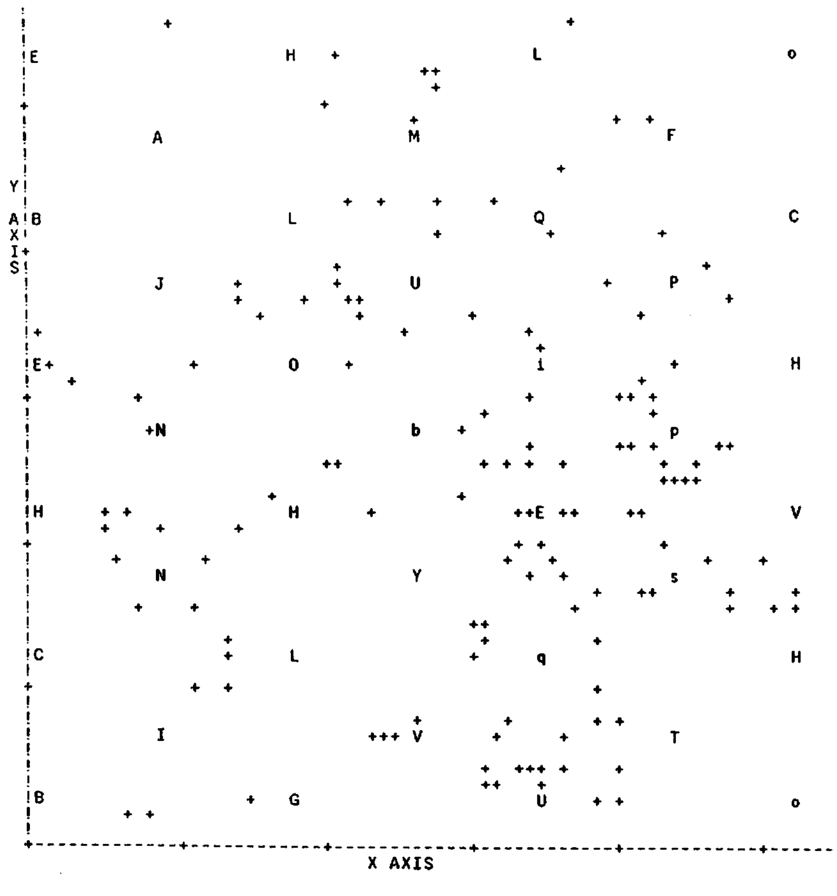
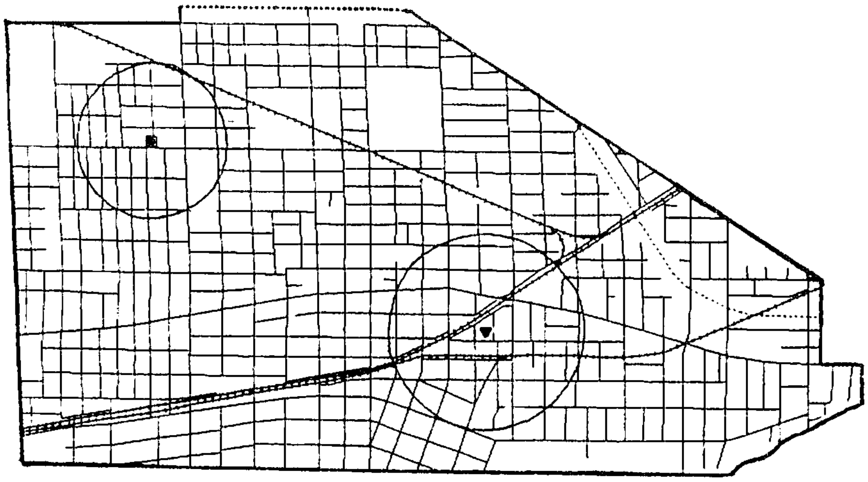


Figure G

Different radii can result in different hot spots. The square denotes the center of a hot spot found using a 1,500-foot radius, the triangle the center of a hot spot found using a 2,000-foot radius.



How STAC Finds "Hot Spots"

The simplest possible spatial pattern is a high concentration of incidents in a particular spot. The STAC program uses the following procedure to search for such a "hot spot," given a search radius. The procedure looks for the point or points at which a circle of the given radius encloses the largest number of incidents. The length of the radius determines the size of the hot spot.

1. The program overlays a triangular grid or lattice on the area in question (for example, a beat or an entire city), with the distance between the nodes of the lattice equaling the specified radius (see Figure H).

2. A circle of that radius is then centered at each node, and the number of incidents in each circle is determined. This lattice of incident totals provides a rough summary of the distribution of crime in the area. Since the radius equals the distance between circle centers, the circles are forced to overlap; this minimizes the likelihood of a high-crime or "hot" area being divided into two pieces and hence overlooked.

3. The program then ranks the nodes according to the incident totals calculated in Step 2 and determines the range of incident totals. The program uses the nodes with incident totals in the top 25 percent of this range as first approximations of the hot spot.

4. This step closes in on the actual hot spot by shifting the circle in an attempt to increase the number of incidents within it. Six more nodes are placed in a hexagon surrounding each of the approximate nodes found in Step 3 (see Figure I). This hexagon is calculated to be closer to its central node than the six closest nodes in the original grid; in effect, this shrinks the scale on which the search is conducted.

5. Circles are now centered at the corners of each hexagon, and the new incident totals are determined.

6. The program then ranks the nodes from Steps 3 and 5 in descending order by number of incidents. The one with the most incidents is taken as the hot spot or place of highest crime density. If there is a tie (if more than one node has the highest number of incidents) all the nodes with this number are given as hot spots.

Figure H

Location of lattice points for the hot spot procedure.

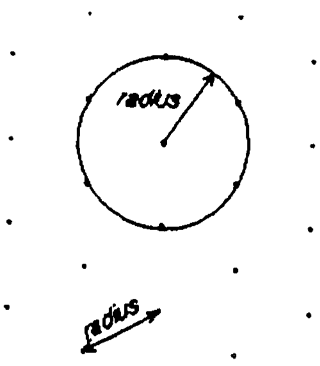
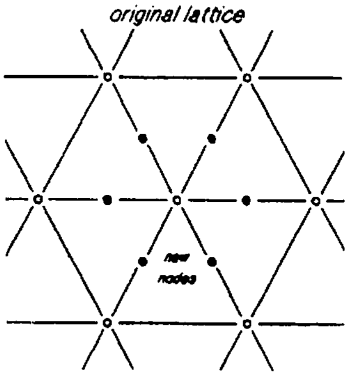


Figure I

Location of new nodes relative to the original lattice.



Lattice Shape. The shapes made by the points of the grid, the *lattice*, are another factor that may influence the search. The program has two shapes available: squares and triangles (see Figure H for an example of the latter). The hot spot procedure currently uses the triangular lattice, which is more compact than the square lattice. The triangular lattice also makes the procedure less likely to miss a hot spot by dividing it between two lattice points.

However, the square lattice may be more useful in a city where the streets are laid out at right angles. Since the incident data are placed by address, the incidents may cluster along streets and thus be more easily pinpointed using a square lattice.

Quantity and Significance of Data. As with the temporal analysis, it is reasonable to assume that there is some minimum number of incidents below which the hot spot procedure gives potentially misleading results. Establishing a minimum number of incidents for spa tial analysis is more complex, however, since the hot spot procedure does not divide incidents into obvious categories as does the temporal analysis.

One possibility is to divide the area of the town by the area of the search circle and multiply by some minimum number of incidents per circle, in effect requiring a minimum density of crime for analysis. The next step is to decide what this minimum number of incidents per circle should be. The rule of five incidents per category (in this case, within a circle) used in the temporal analysis is probably too small for adequate analysis in this case. Further testing of the program will be necessary before a minimum number of incidents for spatial analysis can be firmly established.

Another problem with spatial analysis is the question of whether the distribution of incidents is random or due to some cause. The program will find a "hot spot" regardless of whether the distribution is random or not. In order to be certain the hot spot actually shows an area with a higher incidence of crime, and isn't just an accident of the distribution, the hot spot must contain a significantly larger percentage of incidents than it would if the distribution were simply random (see "Hot Spot: High Crime Zone or Just Coincidence?," page 10).

Applications to Law Enforcement

Spatial analysis has several potential applications to law enforcement. The most obvious is plotting the geographic distribution of various types of crime. Areas ranging from a city block to an entire town can be examined using the radial search and the scanning procedure. These maps can be used to investigate patterns of crimes suggested by other information, such as *modus operandi*, and could help connect several incidents to a single suspect or group of suspects.

The analysis could also have administrative uses for law enforcement. An examination of the spatial distribution of incidents over time could aid in deciding how to allocate resources at various levels. At the patrol level, an officer's beat could be analyzed to find the areas of greatest activity, indicating the areas where free patrol time should be concentrated. At a somewhat higher level, the analysis could help in defining the beat structure for an entire town. The hot spot procedure could be used to find the principal centers of crime, and the beats could be defined or re-defined to take these centers into account.

The spatial analysis program, in addition to supporting crime analysis, could be used to analyze other types of geographical data. For example, the hot spot procedure could identify intersections with the highest number of traffic accidents. The program could also be used to track the locations of various types of calls for service, such as calls about youth gang activity.

Next Steps

The production and testing of these analytic techniques have been an exploratory process to investigate the feasibility of analyzing space and time data. Not only is the analysis feasible, but it promises to provide useful information for crime analysis and other law enforcement activities. For example, patterns for burglary found by the temporal analysis program, while not surprising, are confirmed by police officers. The identification of these patterns by the program suggests that it will be accurate when future more sophisticated, analyses are conducted.

The next step in the STAC project (which will also be supported by a grant from the federal Bureau of Justice Statistics) is to apply the analytic techniques in practical situations. This will require collaboration with one or more law enforcement agencies that use PIMS and the PIMS mapping capability.

The analyses will be modified to address more directly the needs of these agencies, and specific applications of the programs, such as traffic accident analysis, will be tested and revised for maximum usefulness. The specific requirements for, and the form of, these applications will be determined, and the spatial analysis program will be revised to incorporate them. One of the most important requirements that will be addressed is law enforcement agencies' need to be able to analyze small amounts of data in a short period of time.

At the same time, the programs will be made easier to use. At present, there are many different options available to the user at each step of the analysis. As the applications become more clearly defined, more of the information required for the analysis will be written into the programs, simplifying the decisions to be made by the user at each step. The output of the programs will also be revised to provide the information in the most useful format possible, as maps, tables, or other types of graphics.

The theoretical development of the spatial and temporal analytical techniques will continue concurrently with their application and field testing. One direction for further analysis is the combination of temporal and spatial analysis. This is already possible to a certain extent, since the results of one type of analysis can be used as the data for the other. For instance, the temporal analysis program can analyze the incidents in a hot spot. This merging of analytic techniques will continue to be developed, as will the individual techniques. The ultimate goal of the project is to produce a package of computer programs that local law enforcement agencies, both in Illinois and across the country, can use to apply spatial and temporal analysis to a variety of law enforcement concerns.

Hot Spot: High-Crime Zone or Just Coincidence?

The hot spot procedure is of no practical use if it is impossible to distinguish between a hot spot that represents abnormally high criminal activity and one that is simply the highest concentration of incidents that occurred randomly. On a technical level, the hot spot procedure simply finds the place of highest crime density in a given area, regardless of whether or not that density is purely coincidental. Comparison of hot spots produced from the test data with those produced from a "control set" of randomly distributed data shows, however, that the geographical distribution of crime in the test communities was indeed not random. The hot spots produced for these crimes are therefore probably due to unusually high criminal activity in those places in the test communities.

To make this comparison, it was necessary to determine what kind of hot spot would be produced by a completely random distribution of incidents. The level of randomness of a sample can be measured in terms of how much it deviates from complete uniformity.

If incidents are uniformly distributed over an area, the hot spot procedure cannot find a "hot spot" with a higher crime density than any other. In such a case, the percentage of total incidents within a hot spot circle, no matter where it is drawn on the map, will nearly equal the percentage of the total area covered by the circle. If a hot spot of a given radius covers 10 percent of the area, then it will contain roughly 10 percent of the incidents in that area. The percentage of total incidents within a hot spot is the *hot spot percentage*; the hot spot percentage for a uniform distribution of incidents is the *uniform hot spot percentage*. A hypothetical uniform hot spot percentage serves as the benchmark against which the relative uniformity of an area's hot spot percentage can be measured.

In order to compare hot spots from different areas, the *hot spot ratio* is used. The hot spot ratio equals an area's hot spot percentage *as a percentage* of the area's *uniform* hot spot percentage.

In the case of a completely random distribution, the hot spot percentage will exceed the uniform hot spot ratio, because some places will have more incidents than others. For example, a hot spot covering 10 percent of the area could contain 16 percent of the incidents. The hypothetical uniform hot spot percentage for that area would be 10 percent, but the actual hot spot percentage is 16 percent. In other words, the hot spot ratio would equal 160 percent (16 percent divided by 10 percent).

To find the range of the hot spot ratios for completely random distributions, 25 random data sets were generated for an imaginary town and the hot spot procedure was performed on each one. The radius chosen (3,000 feet) gave a circle enclosing 7.8 percent of the town. The uniform hot spot percentage was therefore 7.8 percent.

Each of the 25 hot spots contained at least 10 percent, and sometimes as much as 13.1 percent, of the incidents in the town. The mean hot spot percentage was 11.4 percent, or nearly 150 percent of the uniform hot spot percentage. The random hot spot ratios ranged between 126 percent and 165 percent.

Continued on page 11

Table 5

Uniform hot spot percentage: percentage of uniformly distributed incidents found by the hot spot procedure with a radius of 3,000 feet.

Town # 1	13.0%
Town # 2	10.4%
Town # 3	6.6%
Town # 4	22.0%

Table 6

Range of the hot spot ratios for each crime and town given as percentages of the corresponding values in Table 5.

	Town #1	Town #2	Town #3	Town #4
Robbery	237-346%	303-641%	467-505%	227-455%
Criminal Damage to Property	171-253	170-276	198-264	115-177
Motor Vehicle Theft	198-364	220-346	365-589	114-303
Nonresidential Burglary	253-448	222-650	112-615	114-145
Residential Burglary	186-275	246-361	186-378	136-292
Theft from Motor Vehicles	179-287	204-298	222-344	173-256

Continued from page 10

The next step was to compare these results with the hot spot results from the test data. First, the hypothetical uniform hot spot percentages were found for each of the four communities. Table 5 shows the uniform hot spot percentage for each town, given a radius of 3,000 feet. These numbers vary because the towns have different areas. The hot spot ratio was then found for each of the 10 quarters of data for each of the crimes in each of the towns.

The ranges of hot spot ratios for each case appear in Table 6. The hot spot ratios were then averaged for each case (see Table 7).

The hot spot ratios for the test crime data are generally higher than those for the random distributions. Nineteen of the 24 *lowest* hot spot ratios in the test data were greater than the *highest* of the random distributions (165 percent). After the results were averaged (Table 7), they were even further removed from the 126 percent to 165 percent random distribution range. In 22 of 24 cases, the average hot spot ratio was at least 200 percent, and a few times it was greater than 400 percent. Such results indicate that the incidence of these crimes is probably not randomly distributed.

One of the two remaining cases, nonresidential burglary in town #4, is based on an analysis of only 43 incidents. The low average hot spot ratio (133 percent) may in this case be due to the small number of incidents.

Table 7

Average hot spot ratios for each crime and town as percentage of the corresponding values in Table 5.

	Town #1	Town #2	Town #3	Town #4
Robbery	**284%**	**438%** *	**495%** *	**328%** *
Criminal Damage to Property	**207**	**201**	**223**	**147** **
Motor Vehicle Theft	**277**	**283**	**463**	**234** **
Nonresidential Burglary	**347**	**355**	**392**	**133** *
Residential Burglary	**237**	**290**	**276**	**213** **
Theft from Motor Vehicles	**224**	**258**	**264**	**225** **

**Only three hot spot ratios averaged (yearly data).*

***Only eight hot spot ratios averaged (missing first half of 1983).*

It is interesting to note that both the highest and lowest percentages in Table 7 come from very small data sets (see Table 4). The other instance of a low hot spot ratio (criminal damage to property in town #4) cannot be explained in this way, as there were 796 incidents for the whole period. In this case, the incidents actually may have been scattered randomly across the town.

It is by no means certain that 200 percent is the appropriate hot spot ratio for distinguishing a non-random scattering; another percentage may be more accurate. The distinguishing ratio may also vary with the crime or the town. After further testing of the procedure, it may be possible to determine the distinguishing hot spot ratios for random and non-random distributions for different crimes.

Notes

[1] This bulletin does not identify the four communities because its purpose is to report the development of a *method*, not the findings for any particular community. Since the program is still under development, the results of any analysis are still far from conclusive.

[2] The location is given in state-plane coordinates, a coordinate system based on latitude and longitude and widely used for maps and mapping applications.

[3] See Figure B for an example of the time-of-day distribution.

[4] The grouping of hours in pairs is for reasons of display; the analysis is actually carried out at the hourly level.

[5] See Larsen and Marx, *An Introduction to Mathematical Statistics and Its Applications* (Englewood Cliffs, N.J.: Prentice-Hall, 1981) Chapter 9.

[6] These numbers are based on the assumption of a uniform distribution of incidents, or that an incident has an equal chance of occurring in any one of the time intervals.

[7] Combining categories gives rise to the problem of deciding which categories to combine. This is difficult because, as with the question of data precision, the solutions will vary from crime to crime and from town to town. One general rule might be to combine time periods that frequently have low numbers of incidents. For example, if most nonresidential burglaries occur at night, it may make sense to combine the morning hours into a single category.

[8] The maps can be printed to a wide range of scales, allowing them to be used as overlays for conventional maps; this can greatly speed the production of pin maps.

[9] This radius is set by the user, as is the distance between grid points.

[10] Data were unavailable for the first two quarters of 1983 in town #4, so there are only eight hot spots for the crimes in this town. Also, some of the cases (robbery in towns #2, #3, and #4, and nonresidential burglary in town #4) had so few incidents that they required analysis of a year at a time rather than a quarter, thus yielding only three hot spots for these cases.

Further Reading

The following publications provide more information about the analytical techniques described in this bulletin.

Aldenderfer, Mark and Roger Blashfield, 1984, *Cluster Analysis*. Beverly Hills, California: Sage Publications.

Bartlett, M.S., 1982, *The Statistical Analysis of Spatial Pattern*. London, England: Chapman and Hall.

Buck, George A., Robert Austin, Gary Cooper, Don Gagnon, John Hodges, Kai Martensen, Michael O'Neal, 1973, *Police Crime Analysis Unit Handbook*. Washington, D.C.: U.S. Government Printing Office.

Cliff, A.D. and S.K. Ord, 1973, *Spatial Autocorrelation*. London, England: Pion.

Cliff, A.D. and S.K. Ord, 1981, *Spatial Processes: Models & Applications*, London, England: Pion.

Everitt, B.S., 1980, *Cluster Analysis*, 2nd ed. New York: Wiley.

Georges-Abeyie, Daniel E. and Keith Harries, 1980, *Crime: A Spatial Perspective*. New York: Columbia University Press.

Larsen, Richard J. and Morris L. Marx, 1981, *An Introduction to Mathematical Statistics and Its Applications*. Englewood Cliffs, New Jersey: Prentice-Hall.

Ripley, Brian D., 1981, *Spatial Statistics*. New York: Wiley.

Spath, Helmut, 1985, *Cluster Dissection and Analysis*. Chichester, England: Ellis Horwood.

Unwin, David, 1982, *Introductory Spatial Analysis*. London, England: Chapman and Hall.

Printed by authority of the state of Illinois.
Printing order number 87-89.
2,000 copies.

ILLINOIS
CRIMINAL JUSTICE
INFORMATION AUTHORITY

120 South Riverside Plaza
Chicago, Illinois 60606

Research Bulletin