A Methodology for Evaluating Geographic Profiling Software

Final Report

November 2004

Prepared for
Ronald Wilson
National Institute of Justice
810 Seventh Street, N.W.
Washington, DC  20531

Prepared by
Tom Rich
Michael Shively, Ph.D.
<table>
<thead>
<tr>
<th>Role</th>
<th>Signature</th>
</tr>
</thead>
<tbody>
<tr>
<td>Internal Review</td>
<td></td>
</tr>
<tr>
<td>Project Director</td>
<td></td>
</tr>
<tr>
<td>Technical Reviewer</td>
<td></td>
</tr>
<tr>
<td>Management Reviewer</td>
<td></td>
</tr>
</tbody>
</table>
## Contents

1. **Introduction** ................................................................................................................... 1  
   1.1. Background on Geographic Profiling ................................................................. 1  
   1.2. Approach to Methodology Development .......................................................... 3  

2. **Expert Panel Summary** .......................................................................................... 5  
   2.1. Day 1 .................................................................................................................... 5  
   2.2. Day 2 .................................................................................................................. 8  

3. **Background on the Geographic Profiling Software Applications** .................. 10  
   3.1. CrimeStat ............................................................................................................ 11  
   3.2. Dragnet ............................................................................................................... 12  
   3.3. Predator .............................................................................................................. 12  
   3.4. Rigel Analyst ...................................................................................................... 13  

4. **Evaluation Methodology** .................................................................................... 14  
   4.1. Summary ............................................................................................................. 14  
   4.2. Output Accuracy Testing ....................................................................................... 15  
       Test Inputs ........................................................................................................... 15  
       Performance Measures ....................................................................................... 16  
       Data Analysis ...................................................................................................... 17  
   4.3. User Feedback ..................................................................................................... 17  
       User Survey ......................................................................................................... 17  
       User Log ............................................................................................................. 18  
   4.4. Feature Analysis .................................................................................................. 19  
   4.5. Summary of Deliverables .................................................................................... 20  

References ......................................................................................................................... 21  

Appendix: Expert Panel Transcript ................................................................................. 26
1. Introduction

This report describes a methodology for evaluating geographic profiling software. Following a brief overview of geographic profiling (Section 1.1), Section 1.2 describes how the methodology was developed. The key component of the methodology was convening an expert panel that met in August 2004; a summary and full transcript of the panel’s discussions are in Section 2 and the Appendix, respectively. The panel focused on four geographic profiling software applications, which are described in Section 3. The actual evaluation methodology is outlined in Section 4.

1.1. Background on Geographic Profiling

Geographic profiling is a criminal investigative technique that attempts to provide information on the likely “base of operations” of offenders thought to be committing serial crimes. The base of operations could be the offender’s home, place of employment, a friend house, or some other frequented location. The predictions are based on the locations of these crimes, other geographic information about the case and the suspect, and certain assumptions about the distance offenders will travel to commit crimes.

Canter (2003) argues that geographical profiling was “born” in 1980 when a UK police investigator analyzed the locations of crime scenes of the Yorkshire Ripper and computed the “center of gravity” of the crime scenes thought to be linked to the case. It turned out that the offender lived in the town that the investigator predicted. No doubt other investigators and crime analysts have approximated such information by visual inspection ever since the advent of paper “pin maps.”

In the mid-1990s, more sophisticated models for predicting an offender’s home address were developed, building on the work of Brantingham and Brantingham (1981) and other studies of offender travel behavior (e.g., Rhodes and Conly, 1981). As summarized in Rossmo (1999), key results of these studies include:

- Most crimes occur in relatively close proximity to the offender’s home.
- Crime trips follow a distance-decay function, with the number of crime occurrences decreasing with distance from the offender’s home.
- Juvenile offenders exhibit less mobility than adult offenders
- Patterns in crime trip distances vary by crime type.


While Rigel, CrimeStat, and Dragnet are based on different types of distance-decay functions, they produce the same general type of output. In contrast to a single spatial mean (used in the Yorkshire Ripper case), these software applications create a grid over an area and then calculate the probability that the offender’s base of operations is in each grid cell based on the specified crime-related
locations. As Harries (1999) points out, law enforcement officials could use this information for: (a) suspect and tip prioritization, (b) address-based searches of police record systems, (c) patrol saturation and surveillance, (d) canvasses and searches, (e) mass DNA screening prioritization, (f) department of motor vehicle searches, (g) zip code prioritization, and (h) information request mail-outs.

It is clearly convenient to display the output of geographic profiling software on a Geographic Information System (GIS) that also shows streets, landmarks, political boundaries, and other geographic features of the areas around the crimes. The two- and three-dimensional diagrams in Exhibit 1.1 illustrate how output from geographic profiling software appears in a GIS, with color shadings (and, in the case of three-dimensional diagram, the height of the surface) representing the offender’s likely base of operations. As such, interest in geographic profiling software has grown with advances in computer mapping software and the increased use of GIS by law enforcement researchers and practitioners. Geographic profiling represents an important step in moving computerized crime mapping beyond static displays of crime locations (electronic “pin maps”) and toward more analytical mapping that help analysts interpret spatial data.

Exhibit 1.1: Illustrative Output from Geographic Profiling Software

In recent years geographic profiling has received considerable media attention. As Canter (1999) puts it, the “geographical profiling process is often presented, especially in the mass media, as an exotic,

---

1 Image available at http://www.ojp.usdoj.gov/nij/maps/briefingbook.html#tso
almost science fiction like innovation.” The Washington, DC area sniper case, in particular, led to several media stories on geographic profiling (Bowman, 2002; Lewis, 2002; Onion, 2002), including an appearance on ABC-TV’s “Good Morning America” by Dr. Kim Rossmo, whose research led to the development of Rigel. These and other articles on geographic profiling note the technique’s role in solving specific high profile cases, including serial rape cases in St. Louis (MacKay, 1999), Louisiana (Rossmo, 1999), and Las Vegas (Canter, 2003).

While there is anecdotal evidence from specific cases of geographic profiling helping to solve cases, there has not been a thorough evaluation of any of the geographic profiling software applications. Given the significant differences among currently existing applications – for the example, the cost ranges from free to $60,000 – law enforcement agencies could benefit from such an evaluation, particularly as agencies’ ability to link crimes (e.g., via DNA databases, ballistics identification systems, and sophisticated search capabilities of records management systems) improves over the years. Indeed, there is no existing source of information to which law enforcement agencies can refer to help them make decisions regarding the acquisition of geographic profiling software.

Even geographic profiling’s staunchest supporters do not claim that geographic profiling is a “magic bullet.” First, the number of cases of serial murder, rape, robbery, arson, or burglary is small for all but the largest jurisdictions, and many of these cases are not amenable to geographic profiling, in particular those cases involving offenders that Canter (2004) characterizes as “commuters.” A key input to some geographic profiling software applications is the geographic area about which the software makes calculations\(^2\): if the offender travels a long distance to commit crimes, this “search area” will likely not include the offender’s base of operations. Finally, and perhaps most importantly, the software’s performance is limited by the quality of the geographic and crime information with which it has to work. For example, the utility of geographic profiling software depends on the investigator’s ability to link cases to a single, serial offender, which can be extremely difficult. Obviously, the software will have difficulty estimating the base of operations of a serial offender if it is fed crimes committed by several offenders. That said, law enforcement agencies may view any investment in geographic profiling as worthwhile if it helps solve even a single high profile or serious case.

### 1.2. Approach to Methodology Development

NIJ contracted with Abt Associates to convene an expert roundtable to develop a methodology for evaluating geographic profiling software. NIJ selected the panelists, using two criteria. First, in an effort to make the process as fair and unbiased as possible, all panelists should have no commercial, advisory or any other direct tie to any of the four major software applications (CrimeStat, Dragnet, Predator, Rigel). Second, because criminological and geographic theory is operationalized into software, the group should have a broad range of experiences and areas of expertise. As such, panelists included both researchers and law enforcement practitioners whose expertise span criminology, crime analysis, geography, spatial analysis, and software development.

\(^2\) Rigel Analyst automatically computes the “search area”, whereas it is user-specified in CrimeStat and Dragnet.
Abt Associates prepared a briefing paper and distributed it to the panelists one week before the panel met. The briefing paper contained background material on four geographic profiling software applications and a discussion of potential evaluation methodology issues.

The expert panel met at Abt Associates’ Bethesda (MD) office on August 10-11, 2004. Discussions were recorded (see the summary of the discussions in Section 2 and a transcription in the Appendix). NIJ staff attended the meeting, but were careful not to participate in nor influence the discussions. At the beginning of the meeting NIJ briefly instructed the panel about the overall purpose of the roundtable (to develop a fair and rigorous methodology for evaluating geographic profiling software) and to remind panelists of the key audience of the evaluation (law enforcement officials). Thereafter, NIJ representatives observed the discussions and on a few occasions provided clarification about NIJ’s intent when queried directly by panelists.

After the meeting, Abt Associates organized the roundtable discussions into a general evaluation methodology and circulated an outline of the draft methodology to the panel members for comment. The methodology outlined in Section 4 below incorporates feedback on the draft methodology from the roundtable experts.
2. **Expert Panel Summary**

This section contains a summary of the expert panel meeting held at Abt Associates’ Bethesda office on August 10-11, 2004. A complete transcript of the meeting is in the appendix.

Participants were Keith Harries, Patricia Brantingham, Derek J. Paulsen, Wilpen Gorr, Jay Lee, Shari Lawrence Pfleeger, Michael Shively, Tom Rich, Robin Wilfong and Sean Bair.

2.1. **Day 1**

The meeting convened at 8:30 AM on August 10, 2004.

Debra Stoe and Ron Wilson of the National Institute of Justice opened the meeting by welcoming everyone and indicating that the purpose of the meeting was to develop a rigorous methodology for evaluating geographic profiling software. They expressed a strong desire not to influence the deliberations, and pledged to be simply “observers” during the meeting and to be available to answer questions.

The panelists then introduced themselves. The panel was designed to be diverse, bringing a range of skills and experience to the group. The diversity of the panel was immediately evident – it included both practitioners and researchers with expertise in geography, criminology, crime analysis, spatial analysis, software design and development, and evaluation. None of the panelists had any direct ties to any of the four geographic profiling software applications under consideration (i.e., commercial interests or a role in development of the software).

Following introductions, Derek Paulson gave an overview of Rigel Analyst, CrimeStat, and Dragnet; he indicated that he had tried to obtain a copy of Predator, but had been unsuccessful. (Since none of the panelists has seen or used Predator, the panel focused on the other three software applications for the remainder of the roundtable.) Derek emphasized that the applications are very different; in particular, he noted that Rigel Analyst, because it is commercial software, has more extensive input, analysis, and output capabilities than the other two. Rigel is also the only one of the three applications that automatically creates the search area; in CrimeStat and Dragnet, the search area is user specified. Rigel is also the only application that does geocoding and the only one that has any mapping or GIS capabilities. The applications also use different distance decay functions: Rigel Analyst uses one (but does not allow the user to modify it); CrimeStat allows the user to select from among five functions or, alternatively, run the software with jurisdiction-specific calibrated data; and Dragnet has one distance decay function but will allow users to create their own functions. While all three applications produce a “hit score map,” Derek also emphasized that the software applications produce different numerical statistics: Rigel Analyst produces a “hit score area” and “hit score percent,” CrimeStat produces the peak likelihood for each point in the grid; and Dragnet does not produce any numerical statistics.

The panel then asked for clarification from NIJ on what the purpose of the panel is. Debra Stoe reiterated that the goal is to provide guidance to law enforcement agencies who are interested in learning about (and possibly purchasing) geographic profiling software. She also indicated that most of the evaluations of geographic profiling software have been performed by the developers of the
software, which calls into question the objectivity of the evaluation (or, at least, creates the appearance that there may be bias in the results). Given the substantial commitment of public resources being committed to geographic profiling, an independent and unbiased evaluation of the major profiling software applications was necessary. Ron Wilson added that law enforcement agencies need information regarding “value” – for example, is the $60,000 product delivering substantially more value than the free product. NIJ also emphasized that the evaluation should focus on the software, rather than on the viability of “the concept” of geographic profiling (e.g., do offenders have “activity spaces”).

The panel discussed the issue of what the demand for this software is – particularly since 74% of law enforcement agencies serve populations under 10,000 people. There was concern that the panel was discussing something that only the very large police departments were going to consider using. Robin Wilfong commented that her agency (the Pinellas County Sheriff’s Office) offers geographic profiling services to all law enforcement agencies (many of which are very small) in her county and neighboring counties.

The frequency with which geographic profiling software was used within a law enforcement agency was also discussed. Robin indicated her agency analyzes (using Rigel Analyst) about 20 crime series a year. Sean Bair noted that his former law enforcement agency (which serves a community of 140,000) used traditional methods of analysis, rather than using geographic profiling software. (These traditional methods became more effective once the department automated their crime reporting process.) It was also noted that even small towns are victimized by serial offenders, as predicted by criminological theory (a small cohort of individuals commit the vast majority of crimes). In the end, NIJ emphasized that, regardless of what the market for the software is, the goal of the panel is to provide guidance to any law enforcement agency.

The role of expertise in solving crimes was also discussed, in particular how expertise and local knowledge is used when examining the results of a geographic profile. Panelists wondered when analysts rely on expert judgment and when they feel the need for the assistance of software. Robin Wilfong indicated that she always “folds in” her knowledge about her jurisdiction and can usually tell whether the geographic profiling results are “reasonable.”

It was also noted that the ability to properly link crimes into a series was a prerequisite to obtaining any value from geographic profiling software. Panelists noted that the extent to which an analyst can accurately link cases depends in part on their agency’s automated data systems. In general, the more sophisticated the agency’s records management system (e.g., the ability to search for a particular MO, crime type, location, and/or date range) the more successful the analyst will be at identifying crime series. It was also noted that the presence of sophisticated automated computer systems in law enforcement agencies offers the analyst a host of “data reduction” tools that may prove just as valuable (or more valuable) to crime analysts and investigators compared to geographic profiling software.

Panelists then discussed the relationship of two separate questions: (1) is the methodology valid and (2) does the software implement it correctly? Both were viewed as important, with some panelists emphasizing the need to address the first question before the second. Wil Gore made an analogy to the forecasting arena, where theory has led to the development of forecasting “principles” that have eventually found their way into forecasting software.
At this point, the discussion turned to the last agenda item before the morning break – an overview of evaluation issues. Shari Pfleeger gave an overview of the concept of quality, based on her expertise in software design and on the work of Garvin. Key questions related to quality include: is the theory correct? does the software implement the theory correctly? does the software do the things the user needs done? and, is the software a good use of your money?

After the morning break, the discussion focused on examining the theoretical foundation of geographic profiling software. Derek Paulson’s overview of the software noted that each application uses some type of distance decay function to model the offender’s travel behavior. The panelists discussed the appropriateness of these functions. Panelists commented that there were several ways to improve these functions – for example, by incorporating land use maps (e.g., lakes and other barriers to travel), travel “opportunities”, traffic patterns, and road networks. Panelists also noted that other more sophisticated models could be implemented in geographic profiling software, such as a multiple nodes approach. The panel concluded that distance decay functions were certainly not the ideal way to model offender travel behavior, but that they were relatively easy to implement and a good “baseline” that developers could use in attempts to build more effective models.

The next topic of discussion was the technical implementation of the theory underlying the software. To facilitate this discussion, Derek Paulson demonstrated the use of Rigel Analyst, CrimeStat, and Dragnet, which he had loaded on his laptop, along with sample datasets. His demonstration highlighted the major differences between the software applications, particularly in terms of the user interface. Rigel Analyst, being the only commercially available application, was clearly easier to use and had a broader array of input and output capabilities and features. Dragnet and CrimeStat, by contrast, were developed primarily for researchers; as such, their developers were less concerned about ease-of-use than implementation of the underlying algorithms.

With regard to technical implementation, the panel agreed that it was not practical to expect the evaluator to study the “inner workings” of each software application and draw conclusions as to whether the underlying algorithms had been properly implemented or whether simplifying assumptions had been made.

The discussion then moved to evaluating output accuracy. The panelists agreed that the key question was how one formulates a standard way to measure accuracy. For example, was it possible do develop a “search cost” measure for all three software applications? The issue of what crime series cases to include in the tests was also discussed. Rigel Analyst, it was noted, is only intended to be used for property crimes (and, in fact, law enforcement agencies using Rigel Analyst are required to sign a Memorandum of Understanding stating that they will only use the software on property crimes).

The panelists also discussed the possible different dimensions of output accuracy tests. One formulation discussed involved a three-dimensional graph that had axes for:
- Levels of crime pattern complexity;
- Geographic contexts (e.g., urban grid vs. suburban/rural area)
- Treatments to solve the problem (e.g., functions available and various levels of data enrichment).

The use of sensitivity analysis (varying one factor at a time) was viewed as very important in this analysis.
Panelists discussed the pros and cons of using “made up” crime data, as well as non-crime data (e.g., testing the ability of the software to predict people’s home addresses based on gas purchases and other routine activity locations). A consensus was reached, however, that the output accuracy methodology should only incorporate “real” crime series data provided by law enforcement agencies.

Following the afternoon break, the panel focused on usability issues. The panel agreed that it was important to include a “feature analysis” in the evaluation methodology. In this analysis, the evaluator would make a list of features that one would want to see in geographic profiling software, and then determine the extent to which each software application had those features. One possible “grading scale” discussed was an empty circle, a half-filled circle, and a completely filled circle, an approach often used in software reviews.

At the end of Day 1, the panel developed a list of the probable major components of the evaluation methodology. During this discussion, the major points agreed upon by the panelists is that the evaluation methodology should:

- include some type of testing related to output accuracy (either via “automated test drivers” or testing by trained analysts);
- incorporate the experiences of actual users of the software;
- include a feature analysis that examines what capabilities each software application has; and,
- include multiple definitions of performance.

The panelists agreed that geographic profiling had, to date, been tested only on very limited data sets (notably, Baltimore County crime series data), and that having substantial data sets from a number of law enforcement agencies opened up significant possibilities for analysis and evaluation. Sensitivity analysis was specifically mentioned, both involving varying input parameters and by randomly dropping specific cases from the series.

2.2. Day 2

The panel re-convened at 8:30 AM on August 11, 2004.

Prior to the first agenda item for the day (Evaluating Utility), the panel revisited the issue of trying to standardize the output accuracy tests in the methodology. Panelists recognized several potential challenges in developing unbiased tests, in particular the issue of how to standardize the search cost and the search area. With regard to search area, the panelists concluded that this issue could be handled by running any test first with Rigel Analyst (which computes the search area for the user), and then using that search area as input for Dragnet and CrimeStat. The panelists agreed that tests should involve only property crimes (since Rigel Analyst is only intended for use on these types of crimes) and that tests should involve a users with varying types and levels of experiences.

The panelists explored the pros and cons of developing “automated test drivers” to conduct the tests (i.e., developing software that controls the geographic profiling software applications and automatically runs software through a battery of tests). The alternative is to have actual users conduct the tests. While using test drivers was appealing (e.g., it eliminated the possibility of user error in conducting the tests), panelists with experience in this area later concluded that developing drivers would be expensive and would require significant collaboration with the developers and, therefore,
was not a practical approach. Finally, the panelists also recommended that specific test scenarios be developed, for both the geographic profiling applications and non-geographic profiling “control” methods.

The panel then discussed the role that actual users of these applications might play in the evaluation. Panelists felt that useful information could be gleaned from users, with the caveat that the evaluator needs to be cognizant of the potential problems of these data. For example, a user may have a negative view of a software application because it produced nonsensical results for a particular case, when in fact this occurred because the user incorrectly linked crimes thought to be part of a single series. Or, users may have an unfavorable view of an application because they were inadequately trained or used the software on a case that was not appropriate for geographic profiling.

The panelists recommended that any survey be used to learn what features users would like to see in geographic profiling software, thus using the survey as an opportunity to advance the field of geographic profiling. It was also suggested that a small group of users be recruited to keep a running log of the role that the software plays in individual cases.

Before the morning break, the panelists addressed the role of costs in the evaluation methodology. The panel agreed that documenting the costs of the software would be part of the evaluation methodology and that it was important to document the investment that law enforcement agencies make in geographic profiling software. Costs should include both direct financial costs (e.g., software licenses and the cost of training courses) and opportunity costs (e.g., time required to learn the software). The panel agreed that estimating the benefits of the software (in the sense of a doing a cost/benefit analysis) was not practical.

During the final two hours of the roundtable, the panelists attempted to resolve key outstanding issues and sketch out an overall evaluation methodology for the software applications. To begin the discussion, Abt Associates staff distributed to the panelists a ranking of the evaluation measures, based on the discussions on the previous day. The ordering was: output accuracy (most important), utility, usability, cost, implementation of theory, and theory (least important). The panelists indicated that they agreed with this ranking, although there was disagreement over whether a valid test of output accuracy could be conducted. Some panelists felt that the software applications were too different (e.g., they were based on different assumptions, allowed for different types of inputs, allowed for different options regarding distance decay functions, and computed different performance measures) to be subjected to a standard battery of tests. Other panelists felt that, while “it’s not perfect,” existing case data to perform output accuracy tests had been used in the past and was the best available and feasible methodology for output accuracy.

Proceeding on the assumption that output accuracy could be measured using extant crime data from law enforcement agencies, the panelists discussed the design components of these tests, including the minimum number of crimes in the series, the types of crimes to include, and the number and type of jurisdictions (e.g., type of road network) that should be asked to provide serial crime data. Individual panelists also presented what they felt was their “favorite” output accuracy measure. Favorite measures included the search ratio (the ratio of the optimal search area to the total search area), the profile accuracy (was the offender in the top profile area), and error distance (distance from the highest profile point to the offender’s home base).
The meeting concluded at 12:30 PM on August 11, 2004. The panel moderator and NIJ thanked the panelists for their work.

3. **Background on the Geographic Profiling Software Applications**

In developing the evaluation methodology, the expert panel focused on four geographic profiling software applications: CrimeStat, Dragnet, Predator, and Rigel Analyst. Each of these software applications is summarized in this section. It should be noted that they differ in many respects (see Exhibit 3.1), which should be expected because they were developed for different purposes and audiences. In particular, Rigel Analyst is the only one of the four that targets a commercial audience; the other three were developed either primarily for researchers or for the developer’s own use. Thus, Rigel Analyst naturally has a much more fully developed user interface. Also, only CrimeStat and Rigel Analyst are readily accessible for law enforcement agencies (Rigel Analyst is commercially available and CrimeStat can be downloaded off the Internet); Dragnet is available only by contacting the developer and Predator is currently not available at all.

**Exhibit 3.1: Feature Comparison (as of August 2004)**

<table>
<thead>
<tr>
<th>Overall Features</th>
<th>CrimeStat</th>
<th>Dragnet</th>
<th>Predator</th>
<th>Rigel Analyst</th>
</tr>
</thead>
<tbody>
<tr>
<td>Software platform</td>
<td>Windows</td>
<td>Windows</td>
<td>Windows</td>
<td>Windows</td>
</tr>
<tr>
<td>Primary target users</td>
<td>Researchers</td>
<td>Developer and his research staff</td>
<td>Developer</td>
<td>Law enforcement crime analysts</td>
</tr>
<tr>
<td>Availability</td>
<td>Free download from Web site</td>
<td>Contact developer</td>
<td>Not currently available$^3$</td>
<td>Purchase from ECRI</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Input Features</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Ability to import crime data</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Ability to manually add crime data</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Ability to geocode crime data</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Base maps included</td>
<td>No</td>
<td>No</td>
<td>Yes</td>
<td></td>
</tr>
<tr>
<td>Generation of search area</td>
<td>User specified</td>
<td>User specified</td>
<td>Automatically generated</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Analysis Features</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Type of distance decay function (DDF)</td>
<td>User defined or 1 of 5 different DDFs</td>
<td>User defined or default DDF</td>
<td>Proprietary CGT algorithm</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Output Features</th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Performance measures computed</td>
<td>Peak likelihood</td>
<td>None</td>
<td>Total hunting</td>
<td></td>
</tr>
</tbody>
</table>

$^3$ The only known user of Predatory is the developer, who has not made the software available for application or for external review. Because there is no detailed product description available and none of the expert panelists have seen Predator, no information is available on whether this software application has any of the features shown in this table.
<table>
<thead>
<tr>
<th>Ability to generate reports</th>
<th>No</th>
<th>No</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td>Types of maps produced</td>
<td>Hit score surface map</td>
<td>Hit score surface map</td>
<td>Hit score surface map</td>
</tr>
<tr>
<td>Ability to export results to other mapping software</td>
<td>Yes</td>
<td>No</td>
<td>No</td>
</tr>
</tbody>
</table>

### 3.1. CrimeStat

CrimeStat is a spatial statistics “toolbox” that contains six general types of routines, one of which is a journey-to-crime estimate (the other five are spatial distribution, distance analysis, hot spot analysis, interpolation, and space-time analysis). Dr. Ned Levine, working with other researchers and law enforcement crime analysts, developed CrimeStat with funding from the National Institute of Justice. Originally released in 1999 (Version 1.0), CrimeStat has since been upgraded three times (Version 1.1 in 2000, Version 2.0 in 2002, and Version 3.0 in 2004). The latest release contains a new approach to modeling offender travel behavior. The approach utilizes a crime travel demand model that examines crime travel behavior of repeat offenders over an entire metropolitan area. The model, which is an application of travel demand theory that is widely used in transportation planning, includes modules for predicting crime origins and crime destinations, predicting trips from each origin to each destination, estimating the travel mode used in committing a crime trip, and guessing the likely travel route taken. This model is significantly more complex than the distance decay functions in the previous versions and will require more data collection and calibration efforts from the user.

The CrimeStat application and user manual are available for downloading (at no-cost) from the National Archive of Criminal Justice Data, where it has been downloaded more than 6,000 times through March 2004. Based on e-mails received for technical support, Dr. Levine estimates that 75 percent of users are researchers (in particular students pursuing their doctorate) and 25 percent are practitioners from a variety of governmental agencies. An unknown fraction are using the journey to crime functions, as opposed to the other spatial statistics routines.

A description of CrimeStat’s approach to modeling offender travel behavior (in Version 3.0) is available at the CrimeStat Web site (see below). Briefly, users can describe the distance traveled to commit crimes in one of two ways. The first is by specifying one of five possible distance-decay functions: linear, negative exponential, normal, lognormal, and truncated negative exponential. Each function requires different user-specified parameters. The second method involves use of empirical data: CrimeStat computes a distance function based on a data set of origin (offender residence) and destination (crime location) pairs. For either of these two methods, CrimeStat outputs the probability at the offender’s base of operations is in each cell of a user-specified grid. These probabilities can be exported to a GIS to produce maps similar to Exhibit 1.1.

**For Additional Information:**

3.2. Dragnet

Professor David Canter at the Centre for Investigative Psychology at the University of Liverpool (UK) developed Dragnet in the mid 1990s as a research tool to help him and his graduate students study spatial patterns in serial crimes. As wide distribution of Dragnet has not been a priority for Dr. Canter, the software is currently not available for downloading from a Web site. Persons interested in obtaining the software should contact Dr. Canter. According to Dr. Canter, “I make it available to people who want to co-operate with us in a research partnership. If they will share data and are interested in being co-authors on any subsequent academic publications then I can e-mail the software.”

Dragnet allows any type of function to be used to model the distance that offenders travel to commit crimes, although Canter et al (2000) report that a simple negative decay function without buffer zones was the most effective for the particular data sets they tested. The software application then computes the probability that the offender’s base of operations is in each cell of a user-specified grid and displays these probabilities on a two-dimensional map surface. Dragnet is not currently linked to a GIS to allow layering of Dragnet’s output and maps of the jurisdiction. This feature may be added shortly, as Dr. Canter reports that he recently received a large grant from a law enforcement agency to further develop Dragnet.

For Additional Information:


3.3. Predator

Dr. Maurice Godwin developed Predator in the late 1990s as part of his doctoral research. He uses this software, written in Visual Basic, in his consulting business to help him solve crimes for his client agencies. Godwin (1999) describes how Predator was used in the case of five unsolved murders committed in Raleigh (NC) in 1996 and 1997. In this case, the offender lived “less than one block from the predicated home base area.”

Dr. Godwin indicated that he has not attempted to commercialize or share Predator with other researchers, although he says that it could potentially be turned into a product. The Predator Web site (see below) contains a brief description of the software application, as well as a brief discussion of what Dr. Godwin sees as the strengths of Predator’s approach to geographic profiling. No other information is available on Predator.
3.4. Rigel Analyst

Since 1997, Environmental Criminology Research Inc. (ECRI) has marketed their Rigel geographic profiling software to law enforcement agencies. Rigel is based on the doctoral dissertation of Dr. Kim Rossmo, who co-founded ECRI. Rossmo’s criminal geographic targeting (CGT) model, which produces a “probability surface” that shows the relative likelihood of the offender’s base of operations, is patented. Details of the CGT model can be found in Rossmo (1999).

Two versions of the software are currently available. The “Profiler” version sells for approximately $60,000 and is designed primarily for full-time geographic profilers. According to ECRI, “about a dozen large, national law enforcement agencies” have this version, including the Royal Canadian Mounted Police and the Bureau of Alcohol, Tobacco, and Firearms. ECRI also sells the “Analyst” version that is designed for crime analysts and costs approximately $6,000.4 Staff from ECRI emphasized the importance of training, both in the profession of geographic profiling (one full year of training is necessary to become a full-time geographic profiler) and in the use of the Rigel software. A two-week training course is available for the Analyst version. In the past, the National Law Enforcement and Corrections Technology Center (NLECTC) has provided training on Rigel Analyst, although they do not currently offer this training.

The Rigel web site includes several “success stories” that highlight how Rigel has helped solve crimes.

For Additional Information:


- Presentation by Dr. Kim Rossmo at the National Criminal Intelligence Service Conference, March 17-19, 1998 at the University of Manchester (UK) [http://les1.man.ac.uk/dass/hfc/NCIS/krossmo.pdf](http://les1.man.ac.uk/dass/hfc/NCIS/krossmo.pdf)

---

4 Note that only the Rigel Analyst – and not the Rigel Profiler – is within the scope of the evaluation.
4. Evaluation Methodology

4.1. Summary

The primary goal of the evaluation is to provide assistance to law enforcement agencies who are considering using geographic profiling software by conducting a rigorous evaluation of currently available software applications. The evaluation project should focus, at a minimum, on three such software applications: CrimeStat, Dragnet, and Rigel Analyst. The evaluator may include other software in the project, such as Predator (if available from the developer) and any other newly released software. The evaluator should include the cost of acquiring these applications in their budget. Because of its substantial cost, Rigel Profiler should not be included in the project.

The expert panel convened to develop the evaluation methodology concluded that the evaluation should consist of the following three components:

1. Output accuracy. The evaluator should assess the extent to which each software application accurately predicts the offender’s “base of operations” (e.g., home address, place of employment, an acquaintance’s residence) by conducting a battery of tests on each software application, using actual serial crime data provided by law enforcement agencies. Test results should be compared to results from control methods, such as the spatial mean and “eye-ball” estimates.

2. User feedback. The evaluator should obtain information on each software application’s utility and ease of use from crime analysts and other staff in law enforcement agencies who have actually used these software applications. Specifically, the evaluator should:
   a. Conduct a representative survey of users of each software application.
   b. Recruit a panel of users of each software application and have them keep a log or journal of the role the software plays in specific cases.

3. Feature analysis. The evaluator should conduct a “feature analysis” of each software application. Building on the list provided later in this document, the evaluator will develop a list of features (both “must haves” and “nice to haves”) for geographic profiling software and compare the available features in each software application to the list. The features should cover a wide range of measures, including cost, user friendliness, and specific software capabilities.

Additional details on each of these three components are provided below.

Contractors proposing to conduct the evaluation are free to suggest additional or alternative approaches or modifications. However, proposals must demonstrate the rigor and objectivity of the alternative approach or modification, as well as show how it benefits the methodology. The expert panel, it should be noted, did not feel that it would be cost-effective to assess the appropriateness of each software application’s underlying algorithms.

Contractors should also provide assurances that all staff working on this project have no commercial, advisory or any other direct tie to any of the software applications being evaluated. If it is found that
4.2. Output Accuracy Testing

The expert panel felt that output accuracy – defined as the proximity of the offender’s actual “base of operations” (e.g., residence, place of employment, or other frequented location) to the “top profile areas” (i.e., the predicted most likely region containing the base of operations) predicted by the software applications – is the most important evaluation criterion. To date, these software applications have only been tested against very limited data sets. CrimeStat, for example, has only been tested with crime series data from Baltimore County, Maryland (see Levine, 2002).

The evaluator should specify who will actually perform the output accuracy tests. Options include (1) evaluation staff who have been thoroughly trained in all the software applications and (2) law enforcement analysts who are familiar with all the applications. In either case, the testers should have no direct ties to any of the developers, as noted earlier. The evaluator must also guard against the possibility of tester bias due to different levels of knowledge about the software and familiarity with the jurisdiction where the crimes occurred.

Test Inputs

To conduct these tests, the evaluator should use actual crime series data from several law enforcement agencies. Evaluators are responsible for obtaining these data and should provide evidence (e.g., letters of support from law enforcement agencies) that they can successfully complete this task.

The expert panel recommended that data be obtained from law enforcement agencies that vary in size, urbanicity, type of road network (e.g., grid, non-grid, and sprawl), and presence of major travel obstacles (e.g., major rivers bisecting the jurisdiction).

The expert panel also recommended the following with respect to the crime series data:
- Crime series should be comprised of at least 3 crimes.
- To the extent possible, the data should closely resemble what was actually available to analysts during the investigation (i.e., include, if possible, the locations of other case-related items, just as places of employment or addresses of suspect acquaintances).
- Cases can include both “marauders” (i.e., offenders whose base of operations is within the minimum bounding rectangle of the locations where they commit crimes) and “commuters,” but the evaluator should recognize that some software applications (e.g., Dragnet and CrimeStat) were not designed for commuter-type cases.
- Cases should only involve property crimes, since Rigel Analyst is intended only for use on property crimes.
- Cases should involve a variety of types of offenders, including youths.

The evaluator will be required, as part of an initial project deliverable, to specify in detail the testing protocol, including test scripts. This will help ensure that the tests are as unbiased as possible.
Performance Measures

The evaluator should specify which performance measure(s) will be used in the tests. The panel expert suggested a number of different measures for output accuracy, each of which has advantages and disadvantages (see the table below). In doing so, the panel recognized that, given the nature of the problem that the software is attempting to solve, there are no existing standards for measuring output accuracy, and that it will be up to the evaluator to define these standards. Again, the proposal must demonstrate the rigor and objectivity of any standard and discuss its advantages and disadvantages.

The different measures of output accuracy that the panel suggested are summarized below:

<table>
<thead>
<tr>
<th>Performance Measure</th>
<th>Definition</th>
<th>Advantages</th>
<th>Disadvantages</th>
</tr>
</thead>
<tbody>
<tr>
<td>Error Distance</td>
<td>- Distance from the actual to the predicted base of operations</td>
<td>- Can be applied to all software applications and control methods</td>
<td>- Focuses on a single point, rather than an area (which is more useful to investigators). - Not all of the software applications produce a point for which to compare one another.</td>
</tr>
<tr>
<td>Search Cost / Hit Score</td>
<td>- The percentage of cells, in an overlaid grid, that need to be searched to locate the cell that contains the offender’s base of operations</td>
<td>- Indicates how much the profile reduced the search area, which is a key goal of the software applications.</td>
<td>- Highly dependent on how the search area is defined. - Subject to severe changes in output display based on method of thematically mapping the output</td>
</tr>
<tr>
<td>Profile error distance</td>
<td>- The distance from actual base of operations to the nearest point in the top profile region</td>
<td>- Takes into account the whole profile, rather than a single point</td>
<td>- Ignores the size of the profile area. - Not all of the software applications produce a point for comparison purposes</td>
</tr>
<tr>
<td>Top profile area</td>
<td>- The ratio of the total area of the top profile region to the total search area. Often used in combination with Profile error distance.</td>
<td>- Provides an area on which to focus the search</td>
<td>- Subject to severe changes in output display based on method of thematically mapping the output.</td>
</tr>
<tr>
<td>Profile accuracy</td>
<td>- Whether the offender’s base of operations is within the top profile area (yes/no)</td>
<td>- Provides a simple measure of whether or not the profile was “correct”</td>
<td>- Provides no information on the size of the top profile region. - No ‘cost’ built into the models to reflect the accuracy of ease of travel.</td>
</tr>
</tbody>
</table>

The panel suggested that the evaluator use control measures in the analysis, such as simple spatial distribution measures (e.g., spatial mean) or “eye ball” estimation.
Data Analysis

The evaluator should propose a plan for analyzing the output accuracy test results. The expert panel recommended that the evaluator determine how output accuracy of the different software applications – including the control methods – varies by:

- Crime type (e.g., burglary, robbery, auto theft, and arson)
- Type of jurisdiction (e.g., size, population, road network type)
- Number of crimes in the series. The evaluator should vary the number of crimes in the series by randomly dropping crimes from the series.
- The spatial distribution of the crime series (e.g., highly dispersed or clustered)
- The temporal distribution of the crime series
- The level of experience / expertise of the tester

The evaluator should also conduct additional sensitivity analyses, including varying the values of different input parameters and changing the method by which the output are thematically mapped.

4.3. User Feedback

The expert panel felt that the experiences of existing users of geographic profiling software could contribute important information for the evaluation, and recommended that the evaluator undertake two related tasks:

1. User Survey. The evaluator will conduct a representative survey of geographic profiling software users that focuses on the software’s utility and the user’s overall experience with the software.

2. User Log. The evaluator will recruit law enforcement agency staff that are using the software on a regular basis, and have them keep a log that records (1) the details of each case for which the software is used and (2) the impact, if any, that the software had in the case.

In conducting these two tasks, the evaluator must develop a plan for addressing the potential problems with incorporating user bias. For example, a user may have a negative view of a particular software application because it produced nonsensical results for a particular case, when in fact this occurred because the user incorrectly linked crimes thought to be part of a series. Additionally, users may have an unfavorable view of a software application because they were inadequately trained or used the software on a case that was not appropriate for geographic profiling. The evaluator also needs to control for the extent to which the user has access to other “data reduction” techniques (e.g., sophisticated database querying based on offender characteristics), which can increase the utility of geographic profiling software.

User Survey

The evaluator will conduct a survey of users of each software application. In all likelihood, no comprehensive list of users exists for some of the software applications; thus, the evaluator should propose a method for obtaining a representative sample of users for each application.

The evaluator will design and pilot test a survey instrument. The instrument should build on the following items, which the expert panel recommended for the survey:
• Background questions on the respondent (e.g., position at the agency, # year of experience)
• Background questions on the agency (e.g., size)
• Which geographic profiling software applications have they used
• For each geographic profiling software application used:
  o When the software was acquired
  o Amount and type of training they received
  o Start up time required to fully learn the software
  o Frequency of use (e.g., number of cases by crime type per year)
  o Are they still using the software? If not, why not?
  o Overall satisfaction with the software
  o Degree to which investigators in the agency are receptive to geographic profiles
  o Features they would like to see in a new version of the software
  o Details on success and failure stories
  o Did they calibrate the software with local data
  o Details on the data that was used, including the appropriateness, completeness, and scope
  o What other investigative tools and methods do they use in conjunction with geographic profiling software

The evaluator should propose additional survey items, as appropriate.

User Log

In addition to the user survey, the evaluator will recruit a group of law enforcement personnel who are currently using one or more of the software applications and have them record on a on-going basis information about (1) the cases in which they used the software and (2) the impact that software application had on the investigation or outcome of the case.

The evaluator will design and pilot test a data collection instrument for the log. The instrument should build on the following items, which the expert panel recommended:

• What were the specifics of the case (e.g., crime type, number of crimes in the series)
• What was the hypothesized search area prior to doing the geographic profile?
• What inputs were used (e.g., the size of the search area).
• Did the profile results “look reasonable”?
• To what extent did local knowledge of the jurisdiction affect interpretation of the profile?
• How did the hypothesized search area prior to doing the geographic profile compare to the software application’s output?
• Were the profile results shared with investigators? If yes,
  o How did investigators initially react to the profile?
  o Did they alter any investigative strategy as a result of the profile?
• In what way, if any, did the geographic profile help with the investigation?

The evaluator should propose additional data for the user log, as appropriate.
4.4. Feature Analysis

To assist law enforcement officials who are considering using geographic profiling software, the evaluator will develop a list of desirable features for geographic profiling software and judge each software application against those features. The list should include, at a minimum, those features that the expert panel recommended (see below). A secondary objective of the feature analysis is to determine what new features law enforcement personnel would like to see in these software applications. Discussing how each application would benefit from having new features would assist in efforts to advance the field of geographic profiling.

Since much of the feature analysis focuses on usability issues, the evaluator should recognize – and it should be emphasized in project deliverables – that the software applications were developed for different purposes and audiences. In particular, the degree to which each applications’ user interface has been developed varies widely.

The following features should be included in the feature analysis:

- **Usability-related**:
  - **Installation procedure**. The evaluator shall report the ease with which the software can be installed and any errors that occurred during the installation procedure. The installation procedure should be tested on a variety of Windows platforms.
  - **Documentation**. The evaluator shall report whether the software has on-line help and written documentation. The evaluator shall also assess the overall quality of the documentation.
  - **Technical support**. The evaluator shall report on the availability and quality of technical support.
  - **Reliability**. As part of the output accuracy testing, the evaluator shall record the frequency and seriousness of error messages and crashes.
  - **Performance**. Also as part of the output accuracy testing, the evaluator shall record the approximate time required to compute the probability or risk surface values.
  - **Data management capabilities**. The evaluator shall judge the software’s ability to streamline and simplify data management tasks, such as entering or importing crime series data.
  - **Output / performance measure display**. The evaluator shall report on the range of features available for viewing and analyzing outputs and performance measures.
  - **Overall ease of use**. The evaluator shall judge the software’s overall ease of use. Inasmuch as this measure is subjective, the evaluator shall justify any proposed rating scale.

- **Cost-related**
  - **License cost**. The evaluator shall document the cost of the software license (if any).
  - **Hardware and software requirements**. The evaluator shall detail the hardware and software requirements for operating and fully utilizing (e.g., if a GIS is needed to visualize results) the software.
  - **Training requirements and costs**. The evaluator should determine the cost (and time commitment required) of the level of training that the developer recommends or requires to use the software effectively.
  - **Technical support / maintenance costs**. The evaluator shall document the cost of technical support and maintenance agreements.
• **Assumptions on offender travel behavior**
  
  o  **Type of distance decay or other function.** The evaluator, via a review of the software’s documentation and discussions with the developers, shall summarize the type(s) of functions used to model the distance that offender’s travel to commit crimes.
  
  o  **Assumptions on distances between points.** The evaluator shall describe each software applications’ assumptions regarding travel distances. Two commonly used travel distance metric are the Manhattan (or right-angle) and crow-flies metrics.
  
  o  **Assumptions and modeling of actual travel ease.** The evaluator shall describe other assumptions in each software application regarding offender travel characteristics.

As noted earlier, the evaluator can suggest other features by which to judge the software applications.

### 4.5. Summary of Deliverables

The evaluator will prepare three project deliverables:

• **Revised evaluation design document, including:**
  
  o  The output accuracy testing plan, including details on test scripts and data to be used for the software applications.
  
  o  The user survey instrument and the sampling plan for the survey
  
  o  The user log data collection instrument and the identity of law enforcement personnel recruited for the user log task.
  
  o  List of features included in the feature analysis

• **Draft final report**

• **Final report**
References


Paulsen, Derek (2004). *Geographic Profiling: Hype or Hope?: Preliminary Results into the Accuracy of Geographic Profiling Software*. A presentation at the UK Crime Mapping Conference.  


http://www.ojp.usdoj.gov/nij/maps/Conferences/02conf/Tuesday.html.


http://www.ojp.usdoj.gov/nij/maps/Conferences/02conf/Tuesday.html.


Appendix: Expert Panel Transcript

The appendix contains a complete transcript of the expert panel meeting held at Abt Associates’ Bethesda office on August 10-11, 2004. The transcript is available upon request.