Team Control Number

# 8595

For office use only
T1 ________________
T2 ________________
T3 ________________
T4 ________________

For office use only
F1 ________________
F2 ________________
F3 ________________
F4 ________________

Problem Chosen

**B**

---

**2010 Mathematical Contest in Modeling (MCM) Summary Sheet**
(Attach a copy of this page to each copy of your solution paper.)

Type a summary of your results on this page. Do not include
the name of your school, advisor, or team members on this page.

A criminal act is not committed by a perpetrator in a vacuum; rather a criminal act displays an intimate relation between a person and their environment.

A geographical profile, as a tool, allows investigators to quickly identify probably regions of criminal activity and to try to catch a serial criminal before innocents come to harm.

Our final result was the culmination of two distinct models: a center-of-mass model that took into account the geographical data associated with previous incidents, and a patch-foraging model that took into account the demographic (and geographic) data associated with previous incidents.

Key to the center of mass model was the assumption that there is a set point in space from which a criminal's past incidents are centered and from which a criminal will not stray too far due to the psychology of a criminal. In practice, a set of points that are fixed to a coordinate system yielded an average radius about a polar axis that gave an idea of an area for future crimes.

The patch model was based in foraging theory, that a criminal may be described by a vector of information describing what kind of victim they are attracted to, and the city may be described as a matrix of potential victims. This model included information about the size and density of each patch, the value of each patch for the criminal, the risks to the criminal, center of mass information from the previous model, and a few other variables that would be left for an expert in crime to determine chance of return to a specific patch.

In practice data was obtained about various serial offenders in the city of Baltimore as well as demographic information about the city of Baltimore. The center of mass model was tested through the creation of an upper and lower bound for the likelihood of incidents based on previous incidents, which could then be visually inspected (location of incidents inside or outside the predicted area). The patch model was tested with population data about the city of Baltimore (though any demographic could be modeled with this analysis). This yielded a topographical map which was visually inspected for its relation to the criminal's previous incidents.

The strength of our model is that is gives a deeper understanding of the situation beyond a pure center of mass model, as certain areas within the set geometry are more likely than others, and this model highlights those areas. The main weakness of this model is the sheer amount or detail of information that must be known in order for the model to truly come into its own.

# Executive Summary : Team 8595

## What is it?

*This is a method for predicting where a serial criminal will strike next.*

More specifically, the method takes a map of your region of interest, be it a city, a county, or even a state, and breaks it down into patches based on common traits with what the criminal has shown himself to be interested in. Depending on how good the demographic information about your region is, this model allows an analyst or technician to get results and estimates of the next target more accurately than existing models.

The predictions made and the values used by this patch method are backed up by a secondary method that relies only on knowing the location of where crime events took place

## Why use this one?

The two major benefits of using our methods instead of traditional methods are improved isolation of high-risk areas, allowing you to prioritize force distribution, and better resolution of predictions over older techniques.

## How much information can it process?

Any information that forensic analysts at a crime scene can find can be handled with a detailed implementation. More importantly, the model is flexible enough to allow weighting of different characteristics (for example, age, race, and income of victims) as the criminal continues to attack, updating itself to account for the criminal's possibly changing preferences.

## What is the downside?

A technician or analyst may require some brief period of time to become properly acquainted with the model in order to use it effectively.

## When is this model appropriate?

We have found that even with poor demographic data and poor data collected on-site, the methods work reasonably well in targeting the next location. The model is general-purpose enough for it to be applicable for any relevant data from the analysts so long as the user can reasonably operate the implementation.

## When is this model not appropriate?

If demographics of your city are not available or well-recorded, the center of mass portion of the model can still reduce the search area but the patch model will not be able to isolate a particular area where the criminal is most likely to appear.

# ADAPTING COMPOSITE MODELS TO GEOPROFILING FOR CRIME PREDICTION

TEAM 8595

## Contents



## 1. Introduction

We seek to predict the next location where a serial criminal will strike by forming two models which we combine to develop an estimated target area. The use of two different techniques to predict information about a criminal's movements should yield a more accurate result than either method individually. We also wish to develop techniques that improve upon existing crime-prediction methods.

Initially we use only the location of past crime events to make a naïve first guess for location of the next crime. Motivated by mathematical models for foraging behaviors [5], if demographic data and information on crime type or characteristics of the crimes are present, we use machine learning algorithms to partition the region of interest into patches and then determine likelihood of the next appearance occurring in each patch based on its attractiveness to the criminal, difficulty to reach, and similarity to past crimes. We create a map of the region divided into patches with color-coded 'hot spots' of high likelihood for a simple visual aid to administrators deciding where to focus searches or patrols.

## 2. Major Assumptions

*The serial criminal is fairly rational.* We assume that the serial criminal will attempt to seek the most attractive target within nearby location of his anchor point. From the literature [7] it is apparent that almost all serial criminals operate consistently, and while they may go to some effort to disguise their home or base, they will not perform highly illogical acts, such as traveling to a different state to perform a single crime and then resume activities in their primary city.

*Time is not an important factor.* Contrary to expectation, research [4] has shown that time is an unreliable predictor. We therefore formulate our model primarily time-independent, but with enough flexibility that if time is found to be important in a case, the model can be updated to depend upon time.

*Ease of travel, particularly distance, is an important factor.* Canter and Hammond [2] indicate that serial killers have a strong preference for committing crimes near their homes. We adopt the logarithmic distance decay model for how willing criminals are to visit potential event sites based on distance from an anchor point.

*The majority of crimes in consideration are actually attributable to the serial criminal.* Although the center of mass method does take into account noise and reduces the effect of attributing crimes to the serial criminal that were actually committed by some other deviant, when the number of false hits grows too large, predictive models cannot account for the seeming changes in methodology. The center of mass method was shown empirically to be capable of handling over 40% of the data points being false positives.

## 3. Center of Mass Model

Center of mass has been found [1] to be a highly effective means of predicting the anchor point (often the home) of a serial criminal based upon the locations of crime events. We develop a model for estimating probable central locations via center of mass, adjusted for empirical data where we do not know exactly which crimes were committed by whom. The core of this method is to create a cluster of coordinate means by removing points from the data and recomputing the center of mass.

This allows an investigator to make a list of probable crimes committed by a serial criminal and have some way of distinguishing the actual serial crimes from those with similar methodology.

### 3.1. Center of Mass Mean.

For the sake of simplicity, when this model was implemented for testing we limited it to removing up to three points simultaneously. Removed points cycled through the entire data list and results that agreed with expectations were found for data on the activities of serial criminals. This allows the automatic removal of some noise from a data set when it is not known exactly who committed a crime. The major restriction on this method is that there must be more events due to the serial criminal than due to unpredictable background events.

When the distribution and placement of means are found, we take the mean coordinates of that cluster and examine the standard deviation. Smaller standard deviations indicate that there is less probability of events from sources other than our serial criminal having been considered. At this point we consider the criminal's anchor point found for the purposes of calculating distance to potential sites.

### 3.2. Predictions from Center of Mass.

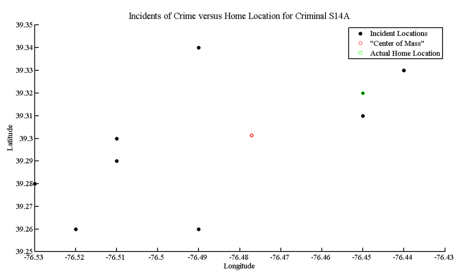
We now wish to isolate an area where the criminal is most likely (95%) to strike next. Our method is to place a polar coordinate system at the central point (the mean of means) and then find mean Euclidean distance from central point to each known event, and the standard deviation of this distance. We simplify here by assuming that with relatively few points, a normal distribution is a good enough approximation, and although it should be skewed so that there is a bias towards closer distances, this was not found empirically to be of importance.

Two circles are then drawn about the central point, each at two standard deviations from the mean distance. It is possible for the inner circle to be too close and thus the end appearance is of a disc rather than a donut, but whether the area between the outer and inner circles is a disc or donut in appearance, there is a 95% probability that the next event will happen in this area. On its own this is not an especially useful prediction, but when combined with the foraging patch model we should be able to make accurate predictions with confidence.

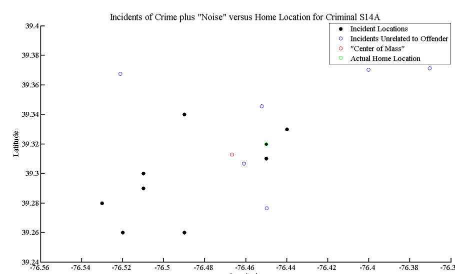
It has been found [1] that there is often a 'buffer zone' around a criminal's home where they tend not to operate, viewing the risk as too great, so such a donut-shaped area is consistent with the literature.

3.3. **Center of Mass Implementation.** Data we were able to use for implementation and verification procedures had low resolution and all points were mapped onto a grid of points, given by latitude and longitude rounded to nearest hundredth. Despite this major limitation, the center of mass method returned with reasonable results for centering the actual crime events.

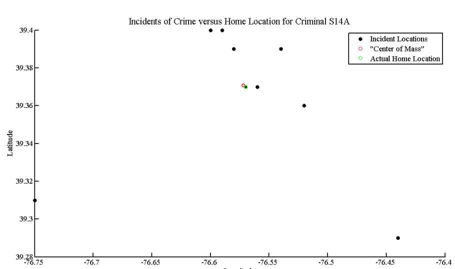
FIGURE 3.1. Center of Mass Reliability with Noisy Data



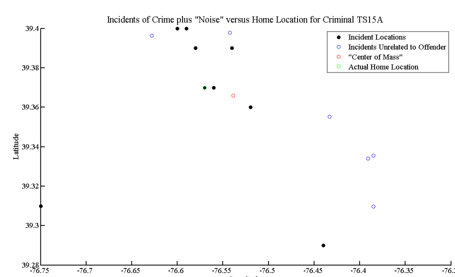

(A) S14A Center of Mass




(B) S14A Center of Mass with Random Noise




(C) TS15A Center of Mass

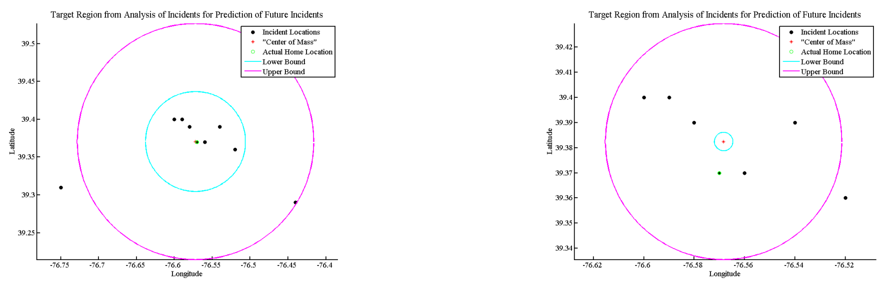



(D) TS15A with Random Noise

*We notice that biased noise perturbs the center of mass calculated by this method only slightly.*

Cheered on by these results we implemented our prediction algorithm outlined in 3.2 and found 95% confidence intervals. The results of the implementation with these data sets are somewhat less appealing, and show how a more refined method is necessary in the event of an outlier case. Particularly indicative of this is the data set `TS15A`, where two outlier points out of 14 otherwise consistent events make the mean and standard deviation extremely out of scale.

FIGURE 3.2. Effect of Outliers on Center of Mass Predictions




(A) $TS15A$ with Anomalous Points (B) $TS15A$ without Anomalous Points

*The center of mass prediction is far better without the two inexplicable anomalies.*

3.4. **Weaknesses of Center of Mass Model.** Clearly a large weakness of the center of mass method is sensitivity to outliers. This difficulty plagues any algorithm that depends heavily on means and standard deviations. By taking many varying centers of mass the effect of single outliers is reduced in the computation of the central anchor point for the data set.

Kent and Leitner [3] detail an approach that we would like to adapt to our center of mass algorithm. Their research and results indicate that creating elliptical areas of interest rather than circular is a better approximation to the actual behavior of a criminal. Time did not permit the adaptation of Kent and Leitner's model to our purposes but it would be beneficial to incorporate their refinements.

## 4. Foraging Patch Model

This model seeks to divide a region of possible crime events of a serial criminal into patches. We then compare past events to establish similarities, through which we make predictions of specific neighborhoods (or patches) of high likelihood.

We consider individual targets within patches, allowing for the possibility of target movement, although in general, the difference between targets and patches is fairly unimportant. In practice, it is unlikely that targets change patches and thus time-dependence is usually unimportant.

4.1. **Patch Model Parameters.** There are $i$ patches chosen from $n$ general patch types. A target $j$ is usually considered stationary for computational simplicity as well as matching real-world expectations (few rapes happen at noon in a courthouse), but for improved accuracy at the cost of additional complexity in implementation and computation targets can be allowed to be considered vulnerable at multiple locations. For some crimes, such as bank robbery or arson, it does not make sense at all for targets (buildings) to move.

- $P_{in'}$ is the $i^{th}$ patch of type $n'$. We represent the size of the patch with $|P_i|$ and the density of targets by $\rho_i$.
- $k$ is a reinforcing term, adjustable depending on whether in expert opinion the criminal is more or less likely to return to the same location.
- $B_{n'}$ is the proportion of previous events that have occurred in patches of type $n'$.
- $\alpha$ is the minimum threshold value for a target/patch.
- $\alpha_{ijn'}$ is the calculated threshold value for target $j$ in patch $i$ of type $n'$.
- $E_j$: expected value of target $j$. Similarly, $E_i$ is the expected value of patch $i$.
- $\psi_{ij}$ is a measure of how difficult it is for the criminal to get to target $j$ in patch $i$.
- $R_{ijn'}$ is a measure from 0 to 1 of the risk of target $i$ in patch $j$ of type $n'$.

If time-dependence is important for a situation, all parameters can be made to change depending on the time except $P_{in'}$, which is determined empirically through comparison of previous events.

4.2. **Patch Model.** First we check to see if a particular patch and target combination is worth considering.

$$\alpha_{ijn'} = E_i\psi_{ij}(1 - R_{ijn'}). \tag{4.1}$$

If $a_{ijn'}(t) \leq \alpha$, we set $c_{i,n'} = 0$ and say that this patch and target combination is not attractive enough to our rational serial criminal - it is either too dangerous, too difficult to get to, or there is too low of a payoff to be worth considering. Otherwise we continue:

$$c_{i,n'} = AkB_{n'}(|P_i|\rho_i E_j)\psi_{ij}(1 - R_{ijn'}). \tag{4.2}$$

The parameter $A$ is the probability of another attack occurring, which for the purpose of predicting the next attack is assumed to be 1. The matrix $C = [c_{i,n'}]$ is an $i \times n'$ matrix of probabilities. Every $c_{i,n'}$ not specified previously is now given the value 0, typically since patch $i$ is not of type other than $n'$. Once normalized so that $\Sigma c_{i,n'} = 1$, the matrix will give us the probabilities of the next event occurring in each patch, indexed by patch number and also by patch type. Summing all values in a column tells us the probability of the next crime taking place in patches of that type.

4.3. **Application of Patch Model.**

4.3.1. *Initialization.* As events take place, as many useful characteristics as possible of the events are recorded. The region of interest is then broken up into small areas (the user must decide on the scale of accuracy here). Based on existing demographics each small area is given a composite value of the same characteristics as the event. The characteristics of previous events are then transformed by the same process into a composite value and their mean is taken to yield a scalar value, by which we divide the entire composite value matrix. We call this matrix of the region's composite score divided by criminal's interest score the *similarity matrix*. We then partition the entire region of interest into contiguous patches where the composite value proportions are near to each other.

Patch properties of interest are calculated, such as difficulty of travel from the estimated home location of the criminal or density of targets in that patch.

4.3.2. *Parameter Calculation.* The major calculation here is to run the center of mass procedure on known events to try to estimate a central node or anchor point of operation. From here $\psi_{ij}$ can be calculated, typically through distance from the central mean predicted by center of mass, keeping distance decay [2] in consideration.

Other parameters such as patch size and density as well as proportions of events in patch types may now be calculated. These will have to be updated each time patches are redrawn.

4.3.3. *Probability Matrix and Solution.* Once all parameters have been calculated, it is a simpled if computationally intensive process to calculate the entries of $C$. Once $C$ has been computed and normalized, police investigators should be given some descriptive statistics on the matrix. Suggested important properties are the highest three probabilities and in which patches they occur, as well as the spread of probabilities. This will allow law enforcement to prioritize high-risk areas for the next crime.

4.4. **Weaknesses of Patch Model.** The patch model as outlined above requires skilled investigators with good knowledge of the demographics of their target region and which factors are important possible motives for the crimes they are investigating. The model is only as good as the distribution of patches, so the factors important to the criminal must be criteria for patch selection.

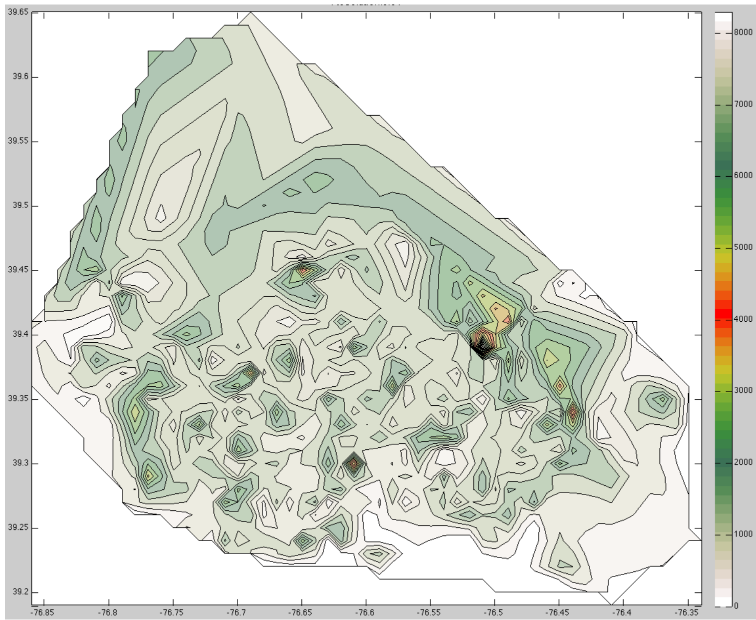
The composite value used to establish the similarity matrix of the region's elements to the pattern of crimes requires a norm. For many purposes a weighted dot product should suffice, although weighting is again dependent on the individual situation and requires the input of an experienced police investigator.

Estimation of the criminal's home location is dependent on the validity of center-of-mass estimates. Many parameters are also dependent on best guesses from experts in serial crimes of a specific type. While burglary can be assigned an expected value based on the value of items typical to a house, apartment, or business in an area, expected values of murder or arson are much harder to come by. Furthermore, no model can predict the next crime if the serial criminal has an arbitrary list of targets without apparent connection.

## 5. Example Implementation of Patch Model

Population data for the city of Baltimore, MD and its surrounding locality were obtained from [6]. The data were latitudinal and longitudinal coordinates as well as demographic data. MATLAB scripts were written to import this data as a matrix that could be manipulated according to the model. This data constituted the topography or 'landscape' of the patch-foraging model. They were then visualized as a contour map.

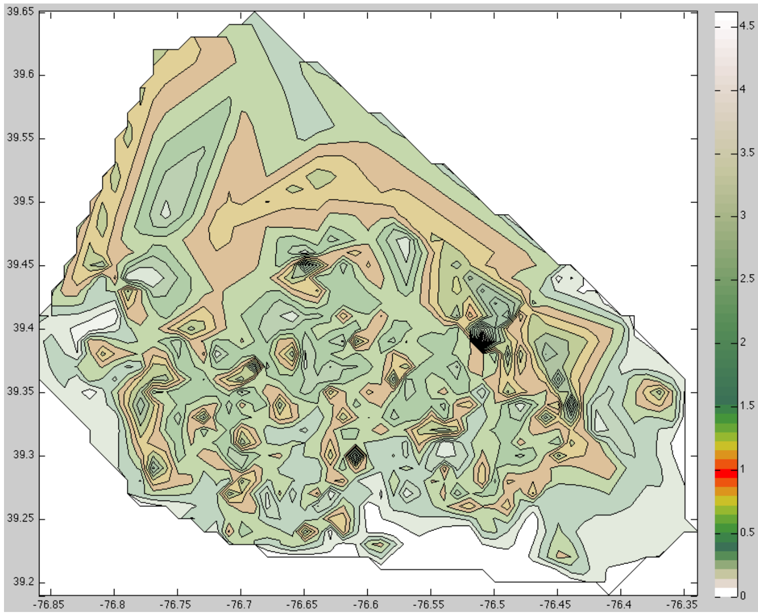
Figure 5.1. Baltimore Area Population by Latitudinal and Longitudinal Coordinates



From this figure we have a general sense of the distribution of population throughout the landscape. Note that the color map is bilinear, highlighting red locations as those with median population values and extreme population going to green and eventually white. This emphasizes the areas of median population, which we hypothesize to be of the most interest to most serial criminals. While we must wait for actual preferences of the criminal to be analyzed, areas that are too dense might be too difficult to escape unseen from, and areas of too low population often do not have enough of interest. This depends, of course, on the type of crime and the individual.

With this topography in place, the crimes of a particular Baltimore serial criminal, `S14A`, were imported and analyzed in the context of the topography. Due to the rudimentary data available, the model was forced to only consider population as the predictive property. We then located each previous crime event and determined the population at that event's coordinates, allowing us to compute the mean population value where `S14A` struck. This value was then compared to the entire landscape by element division. We interest ourselves in areas of closest similarity to the preferences of `S14A`.

FIGURE 5.2. Baltimore Area as Similarity Patches




We now see how 'attractive' the Baltimore area is to S14A and where the most likely next crimes will occur (red areas).

With the topography in place, we applied the results from the center of mass model to predict the anticipated range of S14A, including Euclidean distance decay. The resulting predictions, zoomed in to the areas of highest interest, are shown below. One feature of this application of distance decay is that it drastically reduces the probability of the criminal targeting an area very far away from the anchor point, allowing significant reduction in the search area. The color map below is also coded to emphasize the top 50% of probabilities of attack, making the resultant map more human-readable.

FIGURE 5.3. Anticipated Patches of Interest


*Best candidates based on previous activity and closeness to anchor point.*

## 6. Ideal Implementation

Our implementation was greatly limited by the time we had to write code. Under good circumstances, such as implementation by individuals with authorized access to city government or police data and records, many parameters could be estimated or determined algorithmically to a much higher degree of accuracy.

### 6.1. **Improved Parameter and Metric Determination.**

Several metrics that the patch model 4.2 depends upon are difficult if not impossible to generalize since they require specialized knowledge of local conditions. In particular, $\psi_{ij}$ depends on knowing the ease of travel conditions to patches and determining a norm for turning the vector of crime data at each event into a composite value. The suggested distance metrics of least Euclidean distance and decreasing logarithmic [2] may also not be the best choices for a particular region, city, or social consciousness.

The threshold value $\alpha$ is again dependent on the local user examining values of $\alpha_{ijn'}$ and determining what $\alpha$ is a reasonable threshold value. The reinforcement or avoidance term $k$ depends on the individual criminal and thus must be scaled for each serial criminal.

### 6.2. **Additional Flexibility and Power.**

Time dependence in the entire method could be of value. One place in particular where this would be important in the time scale of a typical serial criminal spree is the value of targets: due to lack of expertise in criminal psychology we weighted each previous case equally when determining the common traits of crime events. However, we would have liked to introduce a system where more recent events are weighted more heavily, to represent the target evolution of a criminal, in particular those dealing with monetary values: thieves might start small but as they gain in confidence begin to steal larger or more well-defended items and consider their beginning thefts no longer worth their time.

The similarity matrix of the 'cityscape' or region of interest to the interests of the criminal in question itself might be transformed from merely looking at scalar values to a vector-valued matrix with a comparison operator to compensate for difficulties with comparison when some values (gender) are binary and others (income) have a wide range.

Additionally, programmers, analysts, and investigative personnel attached to law enforcement should be able to better utilize geographic data stored by local government, which would allow for a much better demographic view of the region in concern and thus much better patch selection than we were able to make. Not only would they be able to obtain far more characteristics of a subunit area in the region, but the resolution would be significantly higher. Due to the limitations of our locational data, estimates were made more imprecise.

Furthermore, with more time, we would interface our predictive patch model with Google Maps or some similar tool that could determine distances along roads or other transportation methods, as well as traffic densities. Combined with time variability, this would make for an extremely powerful method of determining distance with non-Euclidean geometry and including the effect of traffic at varying times of day.

Finally we would have liked to combine the general area prediction made by the center of mass model with the patch model's predictions to find out not only which patches really are the most likely to be targeted next, but also if specific areas in a patch are more likely than others to be the site of the next crime event.

## 7. Comparison to Existing Models

There are a variety of forecasting models available to law enforcement personnel, ranging from simple geometric spatial models to complex Bayesian methods. One relevant example is Crimestat's spatial analysis module, which gives users the ability to forecast the next location where a serial criminal will strike, and then find areas within that region which might be attractive. This model seems to embrace a kind of "on the ground" paradigm to forecasting, narrowing down where a criminal will be so that law enforcement personnel can reconnoiter in the area and get a feel for what zones might be attractive.

In contrast to Crimestat's spatial analysis model, our model works from the converse perspective, first characterizing an area along established predictive features, and then using this characterization to drive the movement of the criminal. The axiom of foraging theory, that an organism will want to maximize the efficiency with which it can obtain what it is seeking, makes knowing what the organism, or criminal, is looking for more important than where it is, or where it resides.

## 8. Conclusion

Finding criminals as quickly as possible is key in the prevention of not only theft of property, but also loss of life. Although geoprofiling techniques have been in existence for some time, they continue to be refined as they are shown to be useful in locating dangerous serial criminals.

We have developed and demonstrated the use of a composite method for predicting subsequent crime events composed of an adjusted center of mass model and a foraging patch model. The combination of these techniques appears to be effective in drastically reducing search areas and hot spots of potential criminal activity when compared with existing models.

Ultimately the ability to locate criminals before they do too much harm will require more than just the coordinated efforts of excellent on-site information, demographic information about where they operate, and probabilistic models that accurate represent the criminal motivations. It will also require skilled analysts who can determine which factors are the most important, with the aid of mathematical methods.

## References

1. E. Beauregard, J. Proulx, and D.K. Rossmo, *Spatial patterns of sex offenders: Theoretical, empirical, and practical issues*, Aggression and Violent Behavior **10** (2005), no. 5, 579–603.
2. D. Canter and L. Hammond, *A comparison of the efficacy of different decay functions in geographical profiling for a sample of US serial killers*, Journal of Investigative Psychology and Offender Profiling **3** (2006), no. 2, 91–103.
3. J. Kent and M. Leitner, *Efficacy of standard deviational ellipses in the application of criminal geographic profiling*, Journal of Investigative Psychology and Offender Profiling **4** (2007), no. 3, 147–165.
4. M. Laukkanen, P. Santtila, P. Jern, and K. Sandnabba, *Predicting offender home location in urban burglary series*, Forensic Science International **176** (2008), no. 2-3, 224–235.
5. J. Lawrance, R. Bellamy, M. Burnett, and K. Rector, *Using information scent to model the dynamic foraging behavior of programmers in maintenance tasks*, (2008).
6. Ned Levine, *Crimestat: A spatial statistics program for the analysis of crime incident locations: Sample data sets*, `http://www.icpsr.umich.edu/icpsrweb/CRIMESTAT/download.jsp`, 2009, This is an electronic document. Date of publication: [Date unavailable]. Date retrieved: February 19, 2010. Date last modified: October 1, 2009.
7. M. Tonkin, J. Woodhams, J.W. Bond, and T. Loe, *A theoretical and practical test of geographical profiling with serial vehicle theft in a UK context*, Behavioral Sciences & the Law **9999** (2009), no. 9999.