Mathematical models for the geographic profiling problem

Mike O'Leary

Department of Mathematics
Towson University

UCLA Applied Mathematics Colloquium
November 12, 2009

Acknowledgements

- Towson University Applied Mathematics Laboratory
  - Undergraduate research projects in mathematics sponsored by companies and government agencies
  - Student participants in 2005-06, 2006-07, and 2007-08:
    - Lauren Amrhine, Brooke Belcher, Brandie Biddy, Colleen Carrion, Chris Castillo, Paul Corbitt, Gregory Emerson, Adam Fojtik, Yu Fu, Natasha Gikunju, Laurel Mount, Kristopher Seets, Jonathan Vanderkolk, Grant Warble, Ruozhen Yao, Melissa Zimmerman
  - Faculty participants:
    - Coy L. May (Towson University) (2005-2007)
    - Andrew Engel (SAS) (2005)
- National Institute of Justice
  - Supported by 2005-IJ-CX-K036, 2007-DE-BX-K005 and 2009-SQ-B9-K014.
  - Special thanks to Ron Wilson for his support
  - Thanks to Stanley Erickson, Iara Infosino, and Tommy Sexton
- Phil Canter, Baltimore County Police Department

The Geographic Profiling Problem

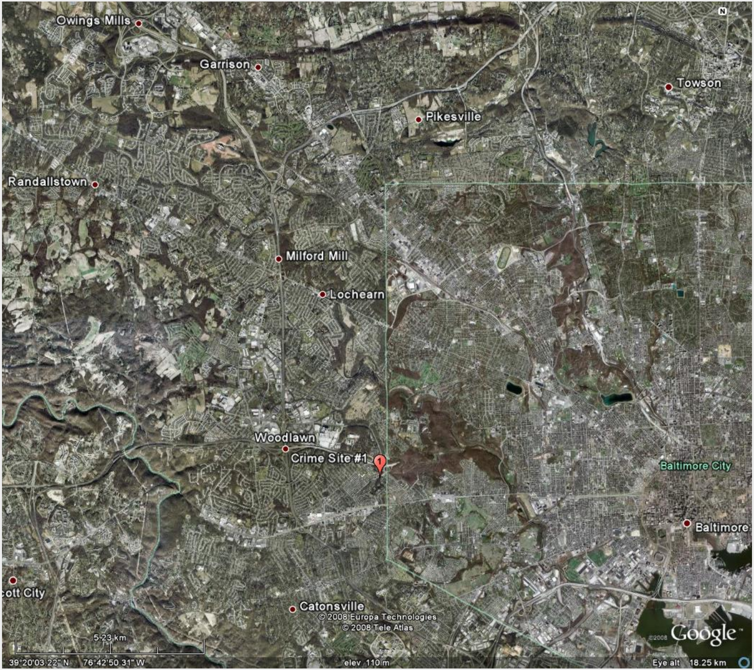
The Geographic Profiling Problem

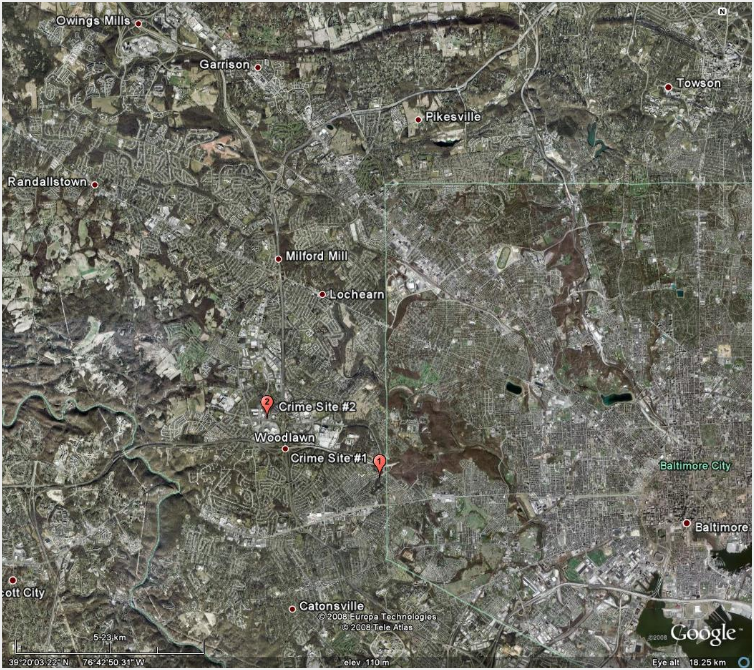
How can we estimate for the location of the anchor point of a serial offender from knowledge of the locations of the offender's crime sites?

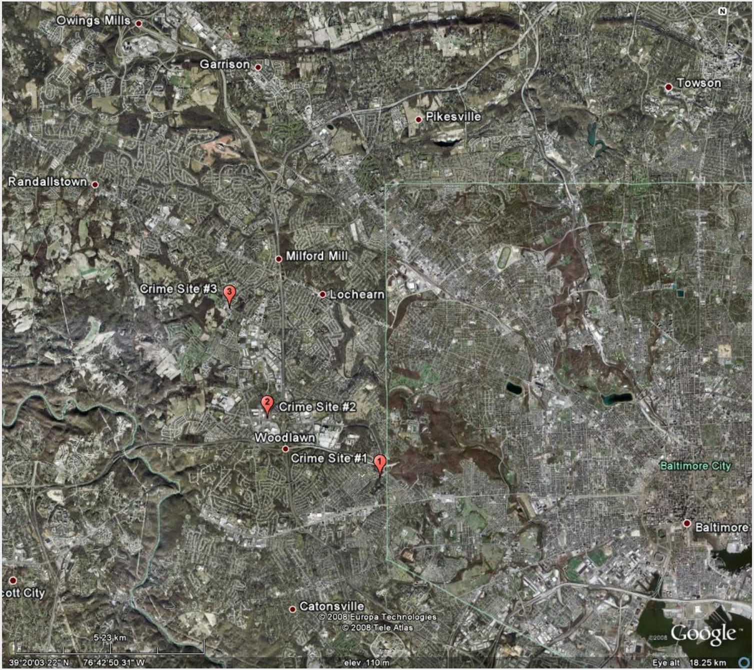
- The anchor point can be the offender's place of residence, place of work, or some other location important to the offender.

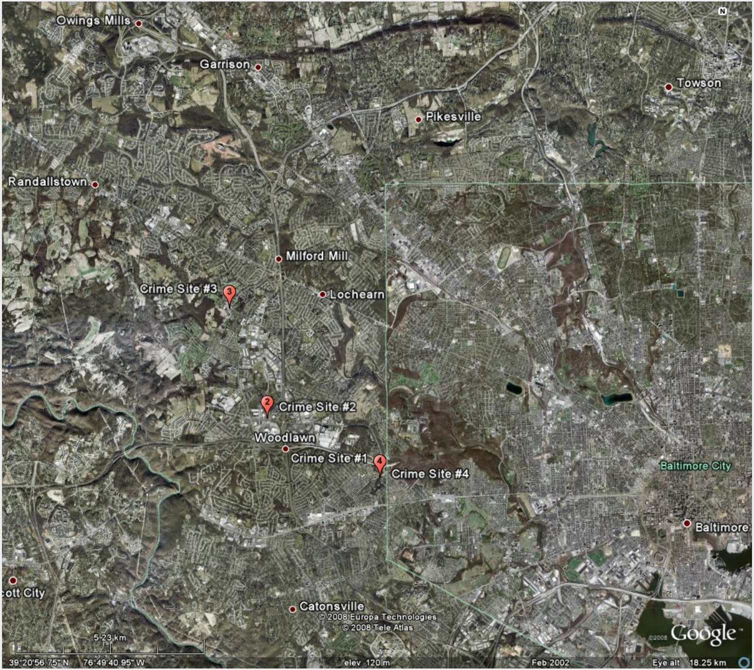
Example- Convenience Store Robberies

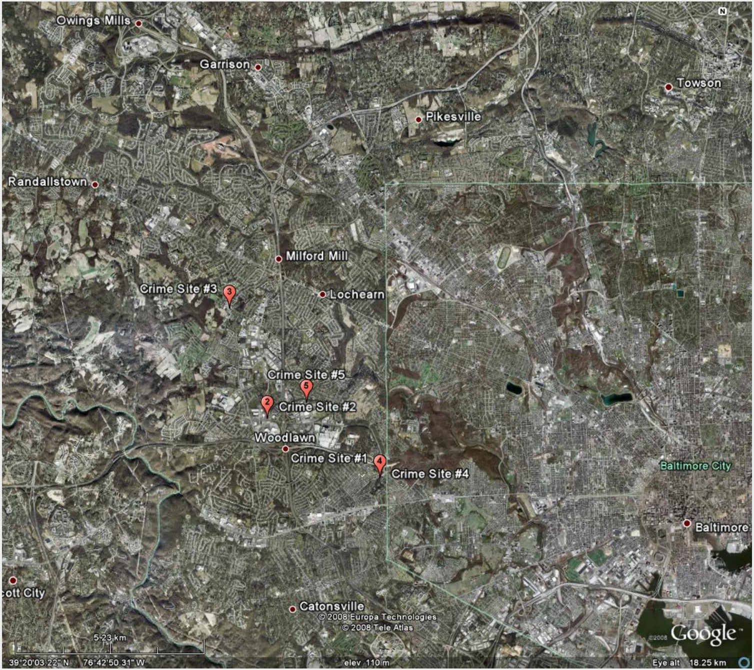
Date	Time	Location		Target
		Latitude	Longitude	
March 8	12:30 pm	-76.71350	39.29850	Speedy Mart
March 19	4:30 pm	-76.74986	39.31342	Exxon
March 21	4:00 pm	-76.76204	39.34100	Exxon
March 27	2:30 pm	-76.71350	39.29850	Speedy Mart
April 15	4:00 pm	-76.73719	39.31742	Citgo
April 28	5:00 pm	-76.71350	39.29850	Speedy Mart

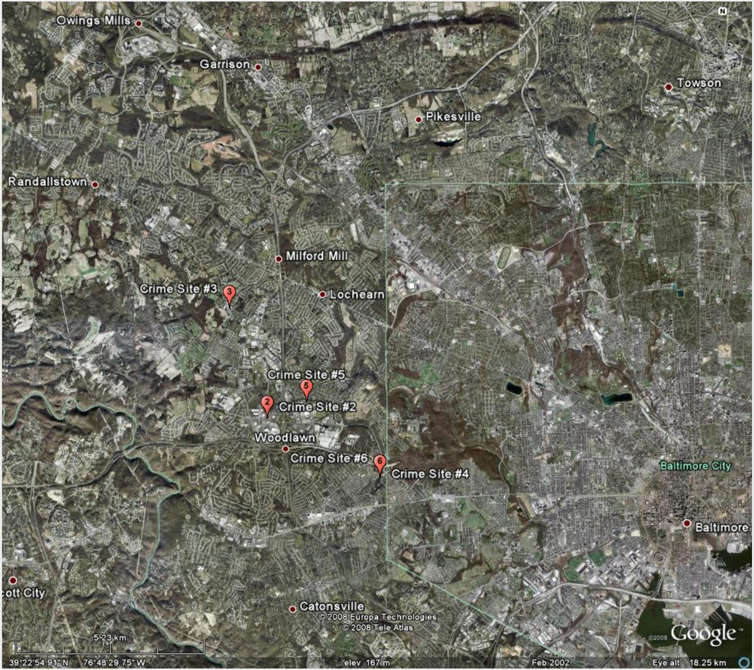

Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Lochearn
Woodlawn
Crime Site #1
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°20'03.22" N 76°42'50.31" W
elev 110 m
Eye alt 18.25 km
Google


Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Lochearn
Crime Site #2
Woodlawn
Crime Site #1
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°20'03.22" N 76°42'50.31" W
elev 110 m
Eye alt 18.25 km
Google


Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Crime Site #3
Lochearn
Crime Site #2
Woodlawn
Crime Site #1
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°20'03.22" N 76°42'50.31" W
elev 110 m
Eye alt 18.25 km
Google


Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Crime Site #3
Lochearn
Crime Site #2
Woodlawn
Crime Site #1
Crime Site #4
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°20'56.75" N 76°49'40.95" W
elev 120 m
Feb 2002
Eye alt 18.25 km


Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Crime Site #3
Lochearn
Crime Site #5
Crime Site #2
Woodlawn
Crime Site #1
Crime Site #4
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°20'03.22" N 76°42'50.31" W
elev 110 m
Eye alt 18.25 km
Google



Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Crime Site #3
Lochearn
Crime Site #5
Crime Site #2
Woodlawn
Crime Site #6
Crime Site #4
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°22'54.91" N 76°48'29.75" W
elev 167 m
Feb 2002
Eye alt 18.25 km
Google

Existing Methods

- A number of software packages have been developed to help this problem.
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)

Existing Methods

- A number of software packages have been developed to help this problem.
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)
- There are a number of controversies surrounding geographic profiling.

Existing Methods

- A number of software packages have been developed to help this problem.
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)
- There are a number of controversies surrounding geographic profiling.
  - How should you evaluate the effectiveness of a geographic profiling tool?

Existing Methods

- A number of software packages have been developed to help this problem.
  - CrimeStat (Ned Levine)
  - Dragnet (David Canter)
  - Rigel (Kim Rossmo)
- There are a number of controversies surrounding geographic profiling.
  - How should you evaluate the effectiveness of a geographic profiling tool?
    - Rich, T., & Shively, M. (2004). *A methodology for evaluating geographic profiling software*
    - Rossmo, K. (2005). *An evaluation of NIJ's evaluation methodology for geographic profiling software*
    - Levine, N. (2005). *The evaluation of geographic profiling software: Response to Kim Rossmo's critique of the NIJ methodology*

Existing Methods

- Some researchers suggest that the best solution is simply to provide humans with some simple heuristics.
  - Snook, B., Canter, D., & Bennell, C. (2002). Predicting the home location of serial offenders: A preliminary comparison of the accuracy of human judges with a geographic profiling system. *Behavioral Sciences & the Law, 20*, 109-118.
  - Snook, B., Taylor, P., & Bennell, C. (2004). Geographic profiling: The fast, frugal, and accurate way. *Applied Cognitive Psychology, 18*(1), 105-121.
  - Snook, B., Taylor, P., & Bennell, C. (2005). Shortcuts to geographic profiling success: A reply to Rossmo. *Applied Cognitive Psychology, 19*, 655-661.
  - Bennell, C., Taylor, P., & Snook, B. (2007). Clinical versus actuarial geographic profiling strategies: A review of the research. *Police Practice and Research, 8*(4), 335-345.
  - Bennell, C., Snook, B., Taylor, P., Corey, S., & Keyton, J. (2007). It's no riddle, choose the middle. *Criminal Justice and Behavior, 34*(1), 119-132.

Existing Methods

- Some researchers suggest that the best solution is simply to provide humans with some simple heuristics.
  - Snook, B., Canter, D., & Bennell, C. (2002). Predicting the home location of serial offenders: A preliminary comparison of the accuracy of human judges with a geographic profiling system. *Behavioral Sciences & the Law, 20*, 109-118.
  - Snook, B., Taylor, P., & Bennell, C. (2004). Geographic profiling: The fast, frugal, and accurate way. *Applied Cognitive Psychology, 18*(1), 105-121.
  - Snook, B., Taylor, P., & Bennell, C. (2005). Shortcuts to geographic profiling success: A reply to Rossmo. *Applied Cognitive Psychology, 19*, 655-661.
  - Bennell, C., Taylor, P., & Snook, B. (2007). Clinical versus actuarial geographic profiling strategies: A review of the research. *Police Practice and Research, 8*(4), 335-345.
  - Bennell, C., Snook, B., Taylor, P., Corey, S., & Keyton, J. (2007). **It's no riddle, choose the middle**. *Criminal Justice and Behavior, 34*(1), 119-132.

Existing Methods

- Spatial distribution strategies
  - Estimate the anchor point with the centroid of the crime series locations
  - Estimate the anchor point with the center of miniumum distance from the crime locations
  - Canter's Circle hypotheses:
    - The anchor point is contained in a circle whose diameter is formed by the two crime locations that are farthest apart
    - Canter D. & Larkin, P. (1993). The environmental range of serial rapists. *Journal of Environmental Psychology, 13*, 63-69.
- Probability distance strategies
  - These have been implemented in software
    - CrimeStat (Ned Levine)
    - Dragnet (David Canter)
    - Rigel (Kim Rossmo)

Developing a Model

To understand how we might proceed let us begin by adopting some common notation

- A point $\mathbf{x}$ will have two components $\mathbf{x} = (x^{(1)}, x^{(2)})$.
  - These can be latitude and longitude
  - These can be the distances from a pair of reference axes
- The series consists of n crimes at the locations $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$
- The offender's anchor point will be denoted by $\mathbf{z}$.
- Distance between the points $\mathbf{x}$ and $\mathbf{y}$ will be $d(\mathbf{x}, \mathbf{y})$.

Mathematical Review of Existing Methods

- Existing algorithms begin by first making a choice of distance metric d; they then select a decay function f and construct a hit score function $S(\mathbf{y})$ by computing

$$S(\mathbf{y}) = \sum_{i=1}^{n} f(d(\mathbf{x}_i, \mathbf{y})) = f(d(\mathbf{x}_1, \mathbf{y})) + \cdots + f(d(\mathbf{x}_n, \mathbf{y})).$$

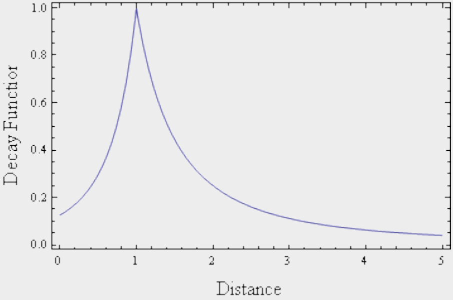
  - Essentially, a probability density function is centered at each crime site, and the result summed.

- Regions with a high hit score are considered to be more likely to contain the offender's anchor point $\mathbf{z}$ than regions with a low hit score.

Rossmo's method:

- The distance metric is the Manhattan distance
- The distance decay function f is

$$f(d) = \begin{cases} \frac{k}{d^h} & \text{if } d > B, \\ \frac{kB^{g-h}}{(2B-d)^g} & \text{if } d \leqslant B. \end{cases}$$



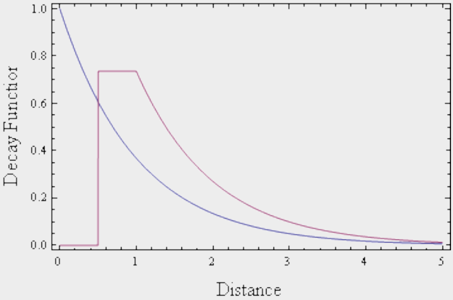
- From Rossmo, K. (2000). *Geographic profiling*, CRC Press

Mathematical Review of Existing Methods

Canter's method:

- The distance metric is the Euclidean distance
- The decay function is either $f(d) = e^{-\beta d}$ or

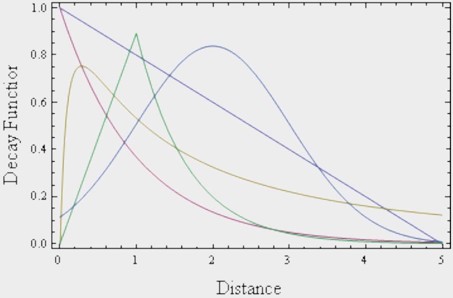
$$f(d) = \begin{cases} 0 & \text{if } d < A, \\ b & \text{if } A \leqslant d < B, \\ Ce^{-\beta d} & \text{if } d \geqslant B. \end{cases}$$



- From Canter, D., Coffey, T., Huntley, M., & Missen, C. (2000). Predicting serial killers' home base using a decision support system. *Journal of Quantitative Criminology, 16*(4), 457-478.

Levine's method:

- The distance metric is the Euclidean distance
- The decay function can be linear, exponentially decaying, truncated exponentially decaying, normal, lognormal, or a function fit to decay data.



- The latest version of CrimeStat (3.1) has a new Bayesian algorithm, significantly different from this approach.

- From Levine, N. (2000). *CrimeStat: A spatial statistics program for the analysis of crime incident locations (v 3.1).*

A New Mathematical Approach

- Suppose that we know nothing about the offender, only that the offender chooses to offend at the location x with probability density $P(\mathbf{x})$.
  - The probability density does not mean that the offender chooses randomly (though he may), rather we are modeling our lack of complete information about the offender.
  - Probabilistic models are common in modeling deterministic phenomena, including
    - The stock market
    - Population dynamics
    - Genetics
    - Epidemiology
    - Heat flow

A New Mathematical Approach

- On what variables should the probability density $\mathrm{P}(\mathbf{x})$ depend?
  - The anchor point $\mathbf{z}$ of the offender
    - Each offender needs to have a unique anchor point
    - The anchor point must have a well-defined meaning- *e.g.* the offender's place of residence
    - The anchor point needs to be stable during the crime series
  - The average distance α the offender is willing to travel from their anchor point
    - Different offender's have different levels of mobility- an offender will need to travel farther to commit some types of crimes (*e.g.* liquor store robberies, bank robberies) than others (*e.g.* residential burglaries)
    - This varies between offenders
    - This varies between crime types
  - Other variables can be included
- We are left with the assumption that an offender with anchor point $\mathbf{z}$ and mean offense distance α commits an offense at the location $\mathbf{x}$ with probability density $\mathrm{P}(\mathbf{x} \mid \mathbf{z}, \alpha)$

A New Mathematical Approach

- Our mathematical problem then becomes the following:
  - Given a sample $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ (the crime sites) from a probability distribution $P(\mathbf{x} \mid \mathbf{z}, \alpha)$, estimate the parameter $\mathbf{z}$ (the anchor point).
- This is a well-studied mathematical problem
- One approach is to use maximum likelihood techniques to estimate the parameters.
  - Construct the likelihood function

$$L(\mathbf{y}, a) = \prod_{i=1}^{n} P(\mathbf{x}_i \mid \mathbf{y}, a) = P(\mathbf{x}_1 \mid \mathbf{y}, a) \cdots P(\mathbf{x}_n \mid \mathbf{y}, a)$$

  - Then the best choice of $\mathbf{z}$ is the choice of $\mathbf{y}$ that makes the likelihood as large as possible.
  - This is equivalent to maximizing the log-likelihood

$$\lambda(\mathbf{y}, a) = \sum_{i=1}^{n} \ln P(\mathbf{x}_i \mid \mathbf{y}, a) = \ln P(\mathbf{x}_1 \mid \mathbf{y}, a) + \cdots + \ln P(\mathbf{x}_n \mid \mathbf{y}, a)$$

    - The log-likelihood has a similar structure to the hit score method
    - Rossmo mentions the possibility of constructing hit scores by

Bayesian Analysis

- Suppose that there is only one crime site $\mathbf{x}$. Then Bayes' Theorem implies that

$$P(\mathbf{z}, \alpha \,|\, \mathbf{x}) = \frac{P(\mathbf{x} \,|\, \mathbf{z}, \alpha)\pi(\mathbf{z}, \alpha)}{P(\mathbf{x})}$$

  - $P(\mathbf{z}, \alpha \,|\, \mathbf{x})$ is the *posterior* distribution
    - It gives the probability density that the offender has anchor point $\mathbf{z}$ and the average offense distance α, given that the offender has committed a crime at $\mathbf{x}$
  - $\pi(\mathbf{z}, \alpha)$ is the *prior* distribution.
    - It represents our knowledge of the probability density for the anchor point $\mathbf{z}$ and the average offense distance α before we incorporate information about the crime
    - If we assume that the choice of anchor point is independent of the average offense distance, we can write

$$\pi(\mathbf{z}, \alpha) = H(\mathbf{z})M(\alpha)$$

      where $H(\mathbf{z})$ is the prior distribution of anchor points, and $M(\alpha)$ is the prior distribution of mean offense distances
  - $P(\mathbf{x}) = \iint P(\mathbf{x} \,|\, \mathbf{z}, \alpha)\pi(\mathbf{z}, \alpha)\, d\mathbf{z}\, d\alpha$ is the *marginal* distribution

Bayesian Analysis

- A similar analysis holds when there is a series of n crimes; in this case
$$P(\mathbf{z}, \alpha \,|\, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \frac{P(\mathbf{x}_1, \ldots, \mathbf{x}_n \,|\, \mathbf{z}, \alpha)\pi(\mathbf{z}, \alpha)}{P(\mathbf{x}_1, \ldots, \mathbf{x}_n)}.$$
- If we assume that the offender's choice of crime sites are mutually independent, then
$$P(\mathbf{x}_1, \ldots, \mathbf{x}_n \,|\, \mathbf{z}, \alpha) = P(\mathbf{x}_1 \,|\, \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n \,|\, \mathbf{z}, \alpha)$$
giving us the relationship
$$P(\mathbf{z}, \alpha \,|\, \mathbf{x}_1, \ldots, \mathbf{x}_n) \propto P(\mathbf{x}_1 \,|\, \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n \,|\, \mathbf{z}, \alpha) H(\mathbf{z}) M(\alpha).$$
- Because we are only interested in the location of the anchor point, we take the conditional distribution with respect to α to obtain the following

Fundamental Result

Suppose that an unknown offender has committed crimes at $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, and that

- The offender has a unique stable anchor point $\mathbf{z}$
- The offender chooses targets to offend according to the probability density $P(\mathbf{x} \mid \mathbf{z}, \alpha)$ where α is the average distance the offender is willing to travel
- The target locations in the series are chosen independently
- The prior distribution of anchor points is $H(\mathbf{z})$, the prior distribution of the mean offense distance is $M(\alpha)$ and these are independent of one another.

Then the probability density that the offender has anchor point at the location $\mathbf{z}$ satisfies

$$P(\mathbf{z} \mid \mathbf{x}_1, \ldots, \mathbf{x}_n) \propto \int_0^\infty P(\mathbf{x}_1 \mid \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n \mid \mathbf{z}, \alpha) H(\mathbf{z}) M(\alpha)\, d\alpha$$

Using the Fundamental Theorem

- For the mathematics to be useful, we need to be able to:
  - Make some reasonable choice for our model for offender behavior
  - Make some reasonable choice for the prior distribution of anchor points
  - Make some reasonable choice for the prior distribution of the average offense distance, and
  - Be able to evaluate the mathematical terms that appear

Models of Offender Behavior

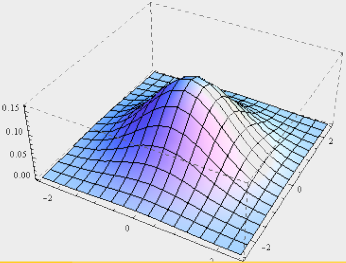
- Suppose that we assume that offenders choose offense sites according to a normal distribution, so that

$$P(\mathbf{x} \,|\, \mathbf{z}, \alpha) = \frac{1}{4\alpha^2} \exp\left(-\frac{\pi}{4\alpha^2}|\mathbf{x} - \mathbf{z}|^2\right).$$

- If we also assume that all offenders have the same average offense distance $\tilde{\alpha}$, and that all anchor points are equally likely, then

$$P(\mathbf{z} \,|\, \mathbf{x}_1, \ldots, \mathbf{x}_n) = \left(\frac{1}{4\tilde{\alpha}^2}\right)^n \exp\left(-\frac{\pi}{4\tilde{\alpha}^2}\sum_{i=1}^{n}|\mathbf{x}_i - \mathbf{z}|^2\right).$$

  - The mode of this distribution- the point most likely to be the offender's anchor point- is the mean center of the crime site locations.



Models of Offender Behavior

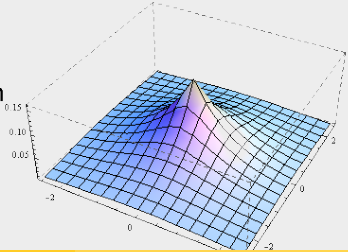
- Suppose that we assume that offenders choose offense sites according to a negative exponential distribution, so that

$$P(\mathbf{x} \,|\, \mathbf{z}, \alpha) = \frac{2}{\pi\alpha^2} \exp\left(-\frac{2}{\alpha}|\mathbf{x} - \mathbf{z}|\right).$$

- If we also assume that all offenders have the same average offense distance $\tilde{\alpha}$, and that all anchor points are equally likely, then

$$P(\mathbf{z} \,|\, \mathbf{x}_1, \dots, \mathbf{x}_n) = \left(\frac{2}{\pi\tilde{\alpha}^2}\right)^n \exp\left(-\frac{2}{\tilde{\alpha}} \sum_{i=1}^{n} |\mathbf{x}_i - \mathbf{z}|\right)$$

  - The mode of this distribution- the point most likely to be the offender's anchor point- is the center of minimum distance of the crime site locations.



Models of Offender Behavior

- What would a more realistic model for offender behavior look like?
  - Consider a model in the form

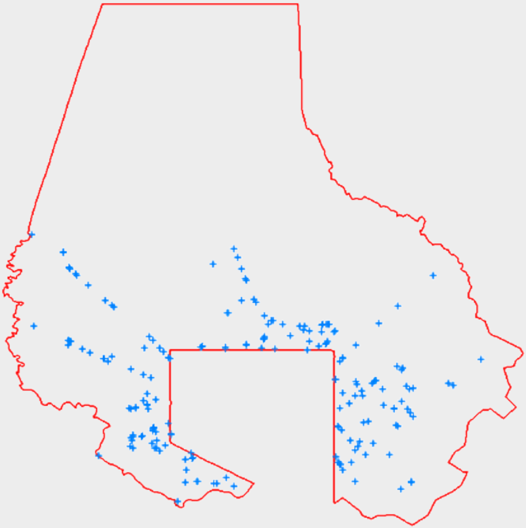
$$P(\mathbf{x} \,|\, \mathbf{z}, \alpha) = D(d(\mathbf{x}, \mathbf{z}), \alpha) \cdot G(\mathbf{x}) \cdot N(\mathbf{z})$$

  - D models the effect of distance decay using the distance metric $d(\mathbf{x}, \mathbf{z})$
    - We can specify a normal decay, so that $D(d, \alpha) = \frac{1}{4\alpha^2} \exp\left(-\frac{\pi}{4\alpha^2} d^2\right)$
    - We can specify a negative exponential decay, so that $D(d, \alpha) = \frac{2}{\pi\alpha^2} \exp\left(-\frac{2}{\alpha} d\right)$
    - Any choice can be made for the distance metric (Euclidean, Manhattan, *et.al*)
  - G models the geographic features that influence crime site selection
    - High values for $G(\mathbf{x})$ indicate that $\mathbf{x}$ is a likely target for typical offenders;
    - Low values for $G(\mathbf{x})$ indicate that $\mathbf{x}$ is a less likely target
  - N is a normalization factor, required to ensure that P is a probability distribution
    - $N(\mathbf{z}) = \left[\iint D(d(\mathbf{y}, \mathbf{z}), \alpha) G(\mathbf{y}) dy^{(1)} dy^{(2)}\right]^{-1}$
    - N is completely determined by the choices for D and G.

Geographic Features that Influence Crime Selection

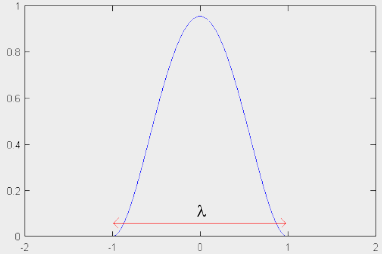
- G models the geographic features that influence crime site selection, with high values indicating the location was more likely to be targeted by an offender.
- How can we calculate G?
  - Use available geographic and demographic data and the correlations between crime rates and these variables that have already been published to construct an appropriate choice for $G(\mathbf{x})$
    - Different crime types have different etiologies; in particular their relationship to the local geographic and demographic backcloth depends strongly on the particular type of crime. This would limit the method to only those crimes where this relationship has been well studied
  - Some crimes can only occur at certain, well-known locations, which are known to law enforcement
    - For example, gas station robberies, ATM robberies, bank robberies, liquor store robberies
    - This does not apply to all crime types- *e.g.* street robberies, vehicle thefts.
  - We can assume that historical crime patterns are good predictors of the likelihood that a particular location will be the site of a crime.

Convenience Store Robberies, Baltimore County



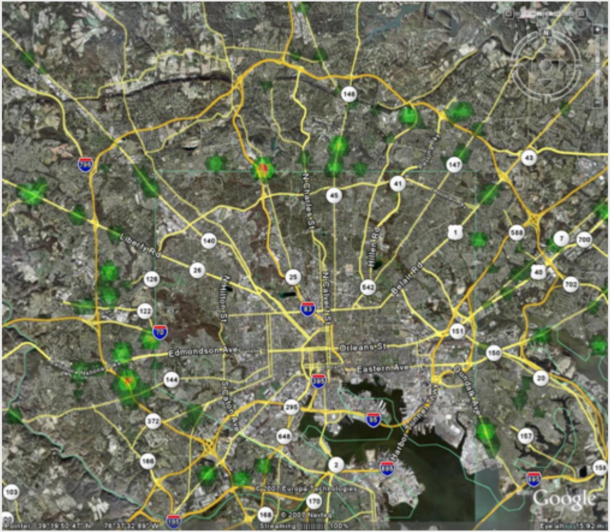
Geographic Features that Influence Crime Selection

- Suppose that historical crimes have occurred at the locations $\mathbf{c}_1, \mathbf{c}_2, \dots, \mathbf{c}_N$.
- Choose a kernel density function $K(y \mid \lambda)$
  - λ is the bandwidth of the kernel density function



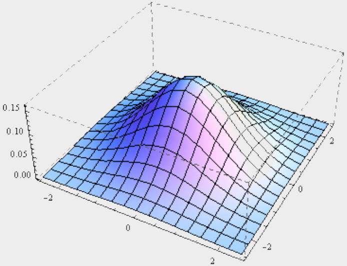
- Calculate $G(\mathbf{x}) = \sum_{i=1}^{N} K(d(\mathbf{x}, \mathbf{c}_i) \mid \lambda)$
  - The bandwidth λ can be *e.g.* the mean nearest neighbor distance
  - Effectively this places a copy of the kernel density function on each crime site and sums

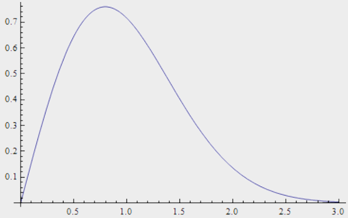
Convenience Store Robberies, Baltimore County



Distance Decay: Buffer Zones

- A buffer zone is a region around the offender's anchor point where they are less likely to offend, presumably due to a fear of being recognized.
- Consider the following models of offender behavior:

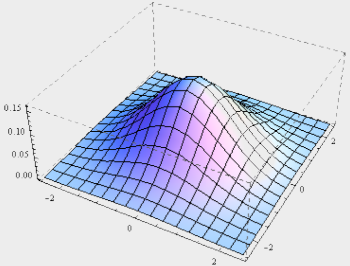


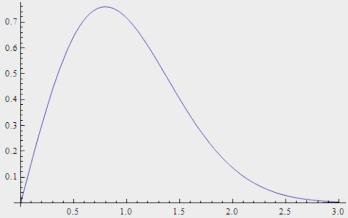


- Which shows evidence of a buffer zone?

Distance Decay: Buffer Zones

- A buffer zone is a region around the offender's anchor point where they are less likely to offend, presumably due to a fear of being recognized.
- Consider the following models of offender behavior:

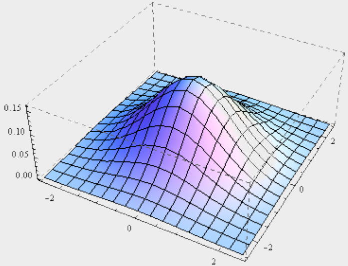


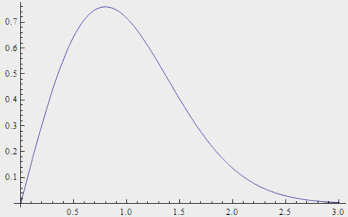


- Which shows evidence of a buffer zone?
  - These are two views of the same distribution

Distance Decay: Buffer Zones

- A buffer zone is a region around the offender's anchor point where they are less likely to offend, presumably due to a fear of being recognized.
- Consider the following models of offender behavior:





- Which shows evidence of a buffer zone?
  - These are two views of the same distribution
  - If the offender is using a two-dimensional normal distribution to select targets, then the appropriate distribution for the offense distance is the *Rayleigh* distribution.

Distance Decay

- Suppose that the (two-dimensional) distance decay component $D(d(\mathbf{x}, \mathbf{z}) \mid \alpha)$ is modeled with a Euclidean distance d
- Then the (one-dimensional) distribution of offense distances $D_{\text{one-dim}}(d \mid \alpha)$ is given by

$$D_{\text{one-dim}}(d \mid \alpha) = 2\pi d \cdot D(d \mid \alpha)$$

- In particular, $D_{\text{one-dim}}(d \mid \alpha) \to 0$ as $d \to 0$, regardless of the particular choice of $D(d \mid \alpha)$, provided $D(0 \mid \alpha) < \infty$.

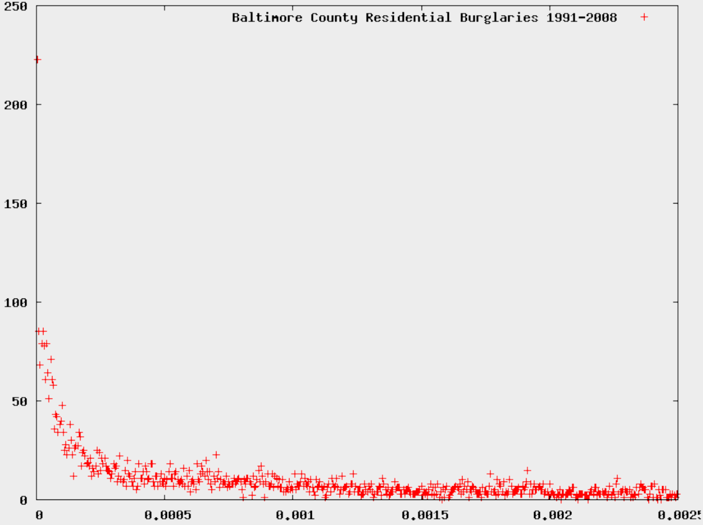
Distance Decay

- Suppose that the (two-dimensional) distance decay component $D(d(\mathbf{x}, \mathbf{z}) \mid \alpha)$ is modeled with a Euclidean distance d
- Then the (one-dimensional) distribution of offense distances $D_{\text{one-dim}}(d \mid \alpha)$ is given by

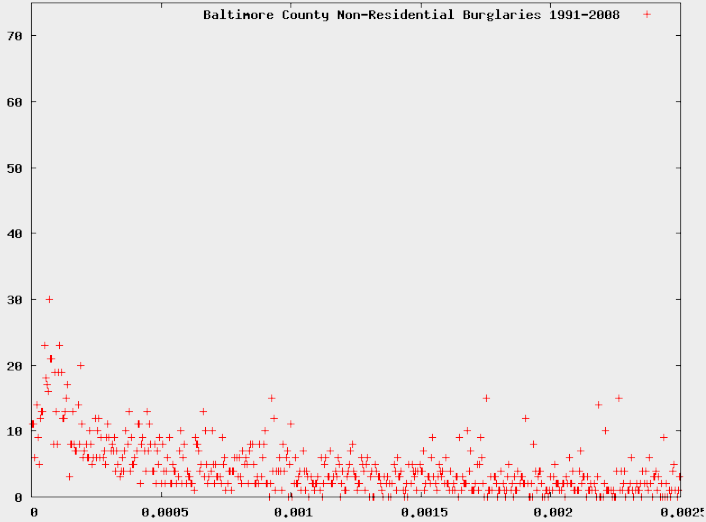
$$D_{\text{one-dim}}(d \mid \alpha) = 2\pi d \cdot D(d \mid \alpha)$$

- In particular, $D_{\text{one-dim}}(d \mid \alpha) \to 0$ as $d \to 0$, regardless of the particular choice of $D(d \mid \alpha)$, provided $D(0 \mid \alpha) < \infty$.
- When considering the effect of distance, it is essential to be aware of the dimension of the underlying function

Distance Decay: Residential Burglaries in Baltimore County



Distance Decay: Non-Residential Burglaries in Baltimore County



Distance Decay: Data Fitting

- Suppose that we measure the aggregate number of offenders who commit a crime at a distance d from their anchor point; call the relative fraction $A(\mathrm{d})$.
- Different offenders are willing to travel different distances to offend; $M(\alpha)$ was defined to be the probability distribution for the mean offense distance across offenders.
- Suppose that each offender chooses targets according to $D_{\text{one-dim}}(\mathrm{d} \,|\, \alpha)$
- Then

$$A(\mathrm{d}) = \int_0^\infty D_{\text{one-dim}}(\mathrm{d} \,|\, \alpha) M(\alpha) \, d\alpha$$

- Since $A(\mathrm{d})$ can be measured and $D_{\text{one-dim}}(\mathrm{d} \,|\, \alpha)$ modeled, we can solve this equation for the prior mean offense travel distance $M(\alpha)$

Distance Decay: Solving the Integral Equation

- The operator $M \mapsto A$ given by

$$A(d) = \int_0^\infty D_{\text{one-dim}}(d \mid \alpha) M(\alpha) d\alpha$$

is smoothing; we expect that the inverse operator $A \mapsto M$ to be ill-posed.

- If we choose a normal form for the two-dimensional decay function (and so a Rayleigh form for the one-dimensional decay function), then

$$A(d) = \int_0^\infty \frac{\pi d}{2\alpha^2} \exp\left(-\frac{\pi d^2}{4\alpha^2}\right) M(\alpha) d\alpha$$

  - If we make the changes of variables $p = \pi/4\alpha^2$, $\alpha = \sqrt{\pi/2p}$, $\omega(p) = \alpha M(\alpha)$, $s = d^2$, we obtain

$$\frac{A(\sqrt{s})}{\sqrt{s}} = \int_0^\infty e^{-sp} \omega(p)\, dp = \mathcal{L}(\omega)(s)$$

Distance Decay: Solving the Integral Equation

- Choose a step size $\delta > 0$, and suppose choose N so that
  - $A(d) \approx 0$ for $d \geqslant N\delta$; then
  - $M(d) \approx 0$ for $d \geqslant N\delta$.
- Suppose that $A(d)$ is not known exactly, but that a sample $\{\rho_1, \rho_2 \ldots, \rho_S\}$ of size S has been drawn.
  - Define $a_j = \#\{s \,|\, d_{j-1} \leqslant \rho_s < d_j\}$
  - Then $A(d_j)\delta \approx a_j/S$
- Apply collocation at the points $d_k^* = (k + \frac{1}{2})\delta$, $1 \leqslant k \leqslant N$ and approximate the integral by the midpoint rule at the nodes $\alpha_j^* = (j + \frac{1}{2})\delta$, $1 \leqslant j \leqslant N$, to find the linear discretization of the integral equation

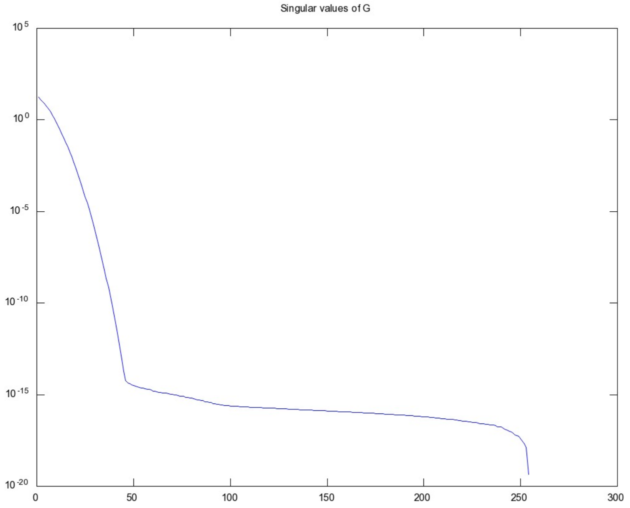
$$\mathbf{a} = G\mathbf{m}$$

  - $G = G_{jk} = \dfrac{\pi S \delta}{2} \dfrac{(j - \frac{1}{2})}{(k - \frac{1}{2})^2} \exp\left(-\dfrac{\pi}{4} \dfrac{(j - \frac{1}{2})^2}{(k - \frac{1}{2})^2} \right)$
  - $\mathbf{a} = (a_1, a_2, \ldots, a_N)$
  - $\mathbf{m} = (M(\alpha_1^*), M(\alpha_2^*), \ldots, M(\alpha_N^*))$

Distance Decay: Solving the Integral Equation

- Attempts to directly solve the equation $\mathrm{G}\mathbf{m} = \mathbf{a}$ for $\mathbf{m}$ fail due to numerical instability; though G is analytically non-singular, it is not numerically non-singular.
- Attempts to solve the equation using the pseudo-inverse $\mathrm{G}^{\dagger}$ so that $\mathbf{m} = \mathrm{G}^{\dagger}\mathbf{a}$ still fail due to numerical instabililty.
  - Write $\mathrm{G} = USV^{\top}$ with $S = \text{diag}(s_1, s_2, \ldots, s_N)$, then $s_j \to 0$ with no appreciable gaps.
  - G has ill-defined numerical rank.

Distance Decay: Solving the Integral Equation



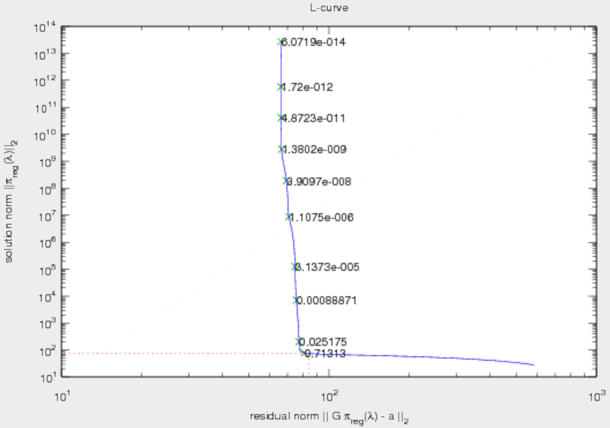
Distance Decay: Solving the Integral Equation

- We can apply Tikhonov regularization; i.e. replace $S^\dagger$ with

$$S_\lambda^\dagger = \text{diag}\left(\frac{s_1}{s_1^2+\lambda^2}, \frac{s_2}{s_2^2+\lambda^2}, \dots, \frac{s_N}{s_N^2+\lambda^2}\right)$$

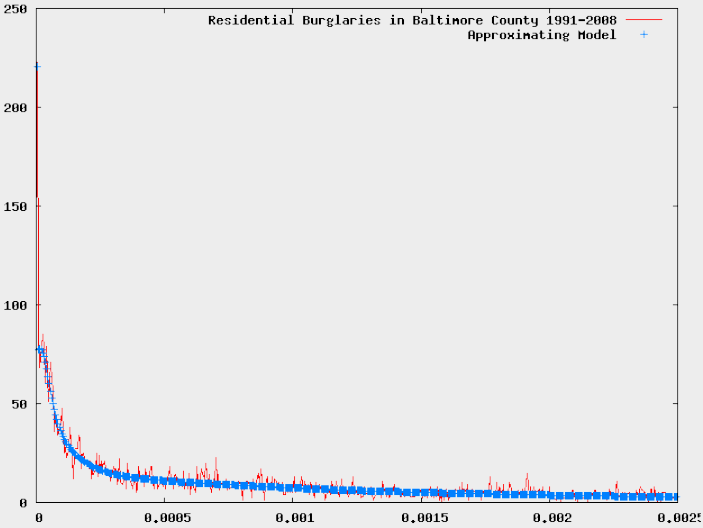
  then $\mathbf{m} = G_\lambda^\dagger \mathbf{a}$ can be calculated.
- The parameter λ can be calculated via the L-Curve method; this locates the point on the graph of log $\|G\pi - \mathbf{a}\|$ versus log $\|\pi\|$ with maximum curvature.

Distance Decay: The L-Curve

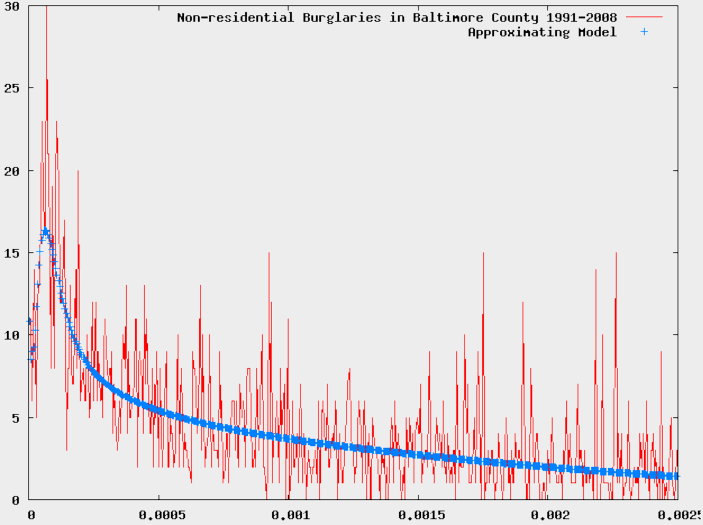


- The L-curve, plotted using data for Baltimore County residential burglaries

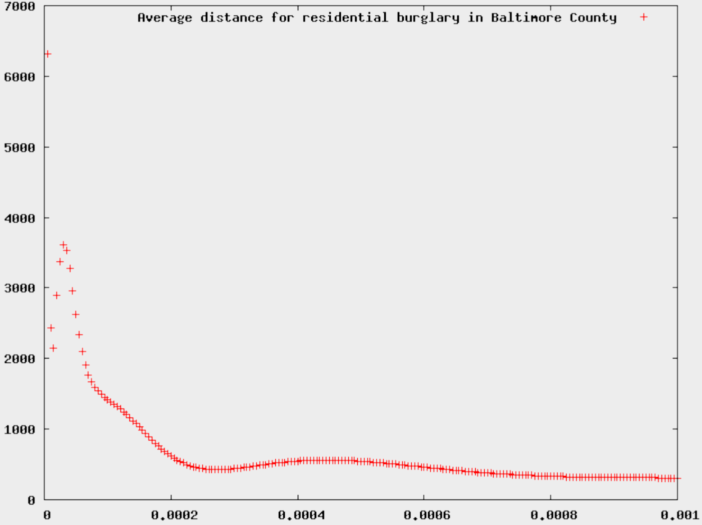
Distance Decay: Residential Burglaries in Baltimore County- Model Fit



Distance Decay: Non-Residential Burglaries in Baltimore County- Model Fit



Distance Decay: Residential Burglaries in Baltimore County- Prior Distribution

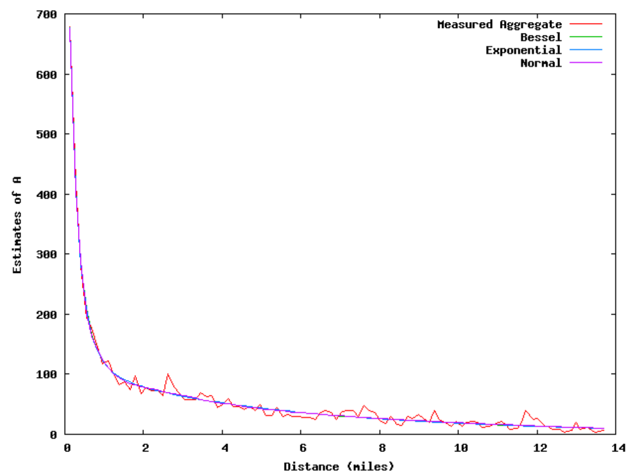


Distance Decay: Solving the Integral Equation

- The fact that the model fits the observed aggregate data does not provide any evidence for the validity of the particular choice of $D(d(\mathbf{x},\mathbf{z}) \,|\, \alpha)$ or $D_{\text{one-dim}}(d \,|\, \alpha)$, as these simply result in different choices for the prior distribution $M(\alpha)$ of average offense distance.
- Indeed, reasonable choices of $D(d(\mathbf{x},\mathbf{z}) \,|\, \alpha)$ include:
  - Normal:
$$D(\mathbf{x} \,|\, \mathbf{z}, \alpha) = \frac{1}{4\alpha^2} \exp\left(-\frac{\pi}{4\alpha^2}|\mathbf{x}-\mathbf{z}|^2\right)$$
  - Negative Exponential:
$$D(\mathbf{x} \,|\, \mathbf{z}, \alpha) = \frac{2}{\pi\alpha^2} \exp\left(-\frac{2|\mathbf{x}-\mathbf{z}|}{\alpha}\right)$$
  - Bessel K:
$$D(\mathbf{x} \,|\, \mathbf{z}, \alpha) = \frac{\pi}{8\alpha^2} K_0\left(\frac{\pi|\mathbf{x}-\mathbf{z}|}{2\alpha}\right)$$

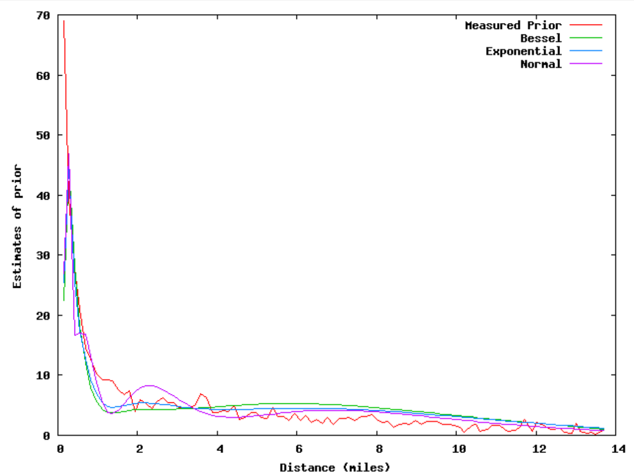
Distance Decay, Residential Burglary

- Comparing the measured aggregate $A(d)$ with calculated values $A_{\text{calculated}}(d) = \int_0^\infty D_{\text{one-dim}}(d \mid \alpha) M(\alpha)\, d\alpha$ for different choices of $D(\mathbf{x} \mid \mathbf{z}, \alpha)$.



Distance Decay, Residential Burglary

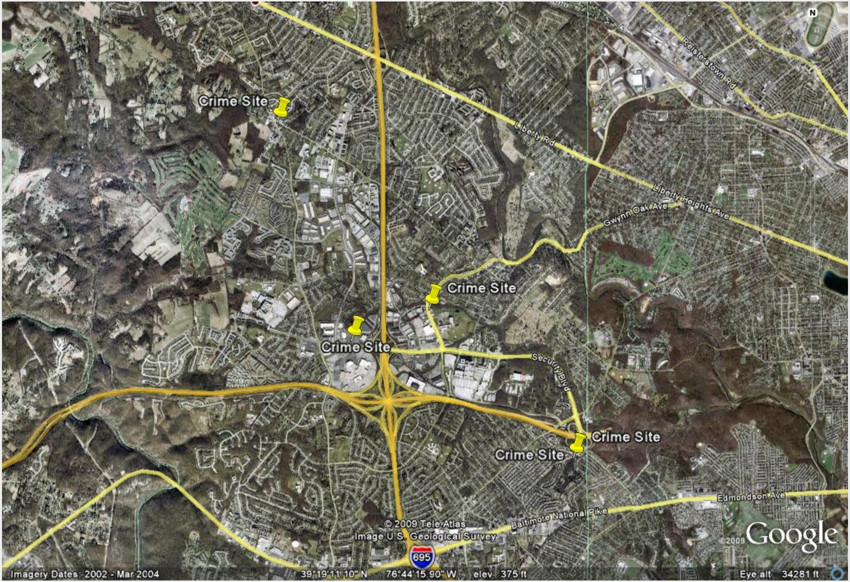
- Comparing the measured prior $\mathrm{M}(\alpha)$ (2889 offenders, 5863 offenses) with calculated values for different choices of $\mathrm{D}(\mathbf{x} \mid \mathbf{z}, \alpha)$.



Search region

- Not all locations are equally likely to be the sites of an offender's home base.
  - The offender's home base is unlikely to be located in a body of water, on a mountain, or in some other unpopulated area.
  - Other, densely populated regions are much more likely to contain an offender's anchor point.
- Consider the geography near the elements of our crime series

Search region

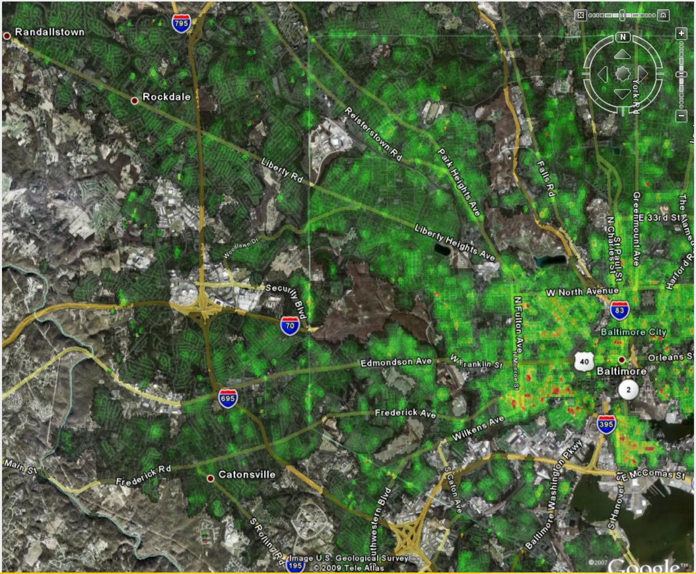


Anchor Points

- We have assumed
  - Each offender has a unique, well-defined anchor point that is stable throughout the crime series
  - The function $H(\mathbf{z})$ represents our prior knowledge of the distribution of anchor points before we incorporate information about the crime series.
- What are reasonable choices for the anchor point?
  - Residences
  - Places of work
- Suppose that anchor points are residences- can we estimate $H(\mathbf{z})$?
  - Population density information is available from the U.S. Census at the block level, sorted by age, sex, and race/ethnic group.
    - We can use available demographic information about the offender
    - Set $H(\mathbf{z}) = \sum_{i=1}^{N_{\text{blocks}}} = p_i K(\mathbf{z} - \mathbf{q}_i \mid \sqrt{A_i})$
    - Here block i has population p_i, center $\mathbf{q}_i$, and area A_i.
  - Distribution of residences of past offenders can be used.
    - Calculate $H(\mathbf{z})$ using the same techniques used to calculate $G(\mathbf{x})$

Anchor Points

- Population density near the crime series

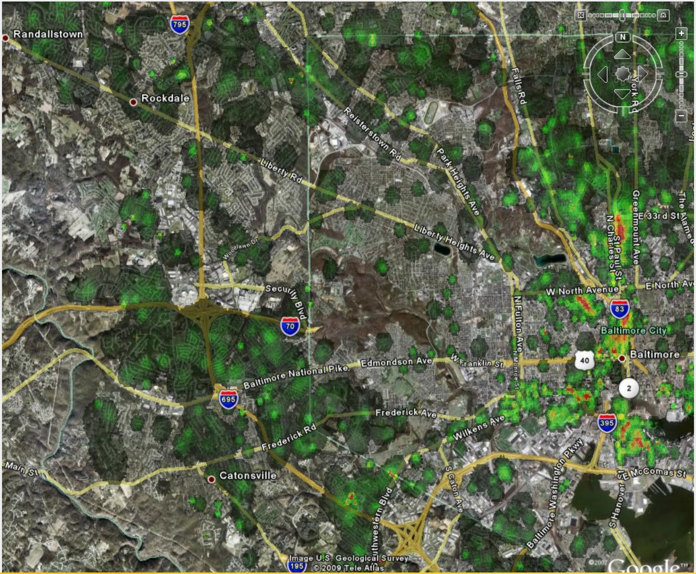


Offender information

- If information about the offender is available, we can use it to refine the search area.
- Census data is available at the block level for
  - Race / ethinic group
  - Age
  - Sex

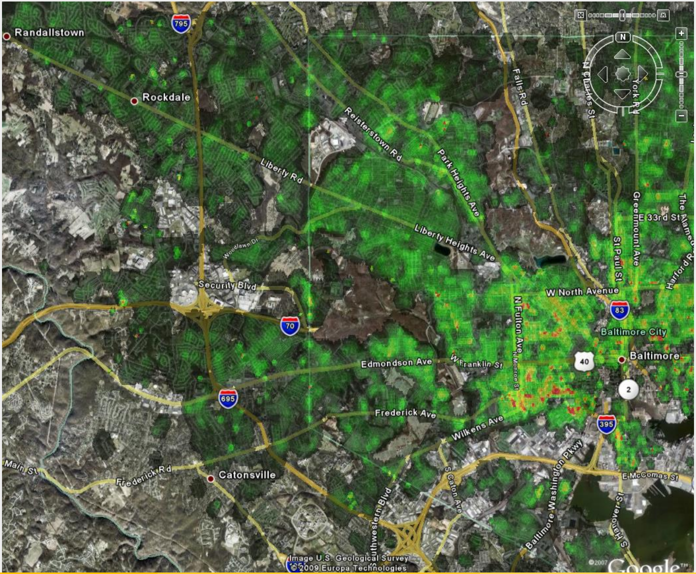
Offender information

- Population density for asian 18-34 year old men



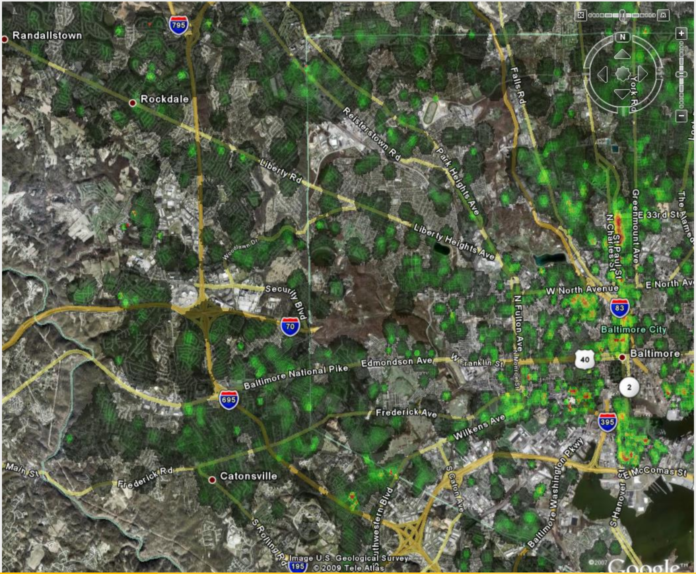
Offender information

- Population density for black 18-34 year old men



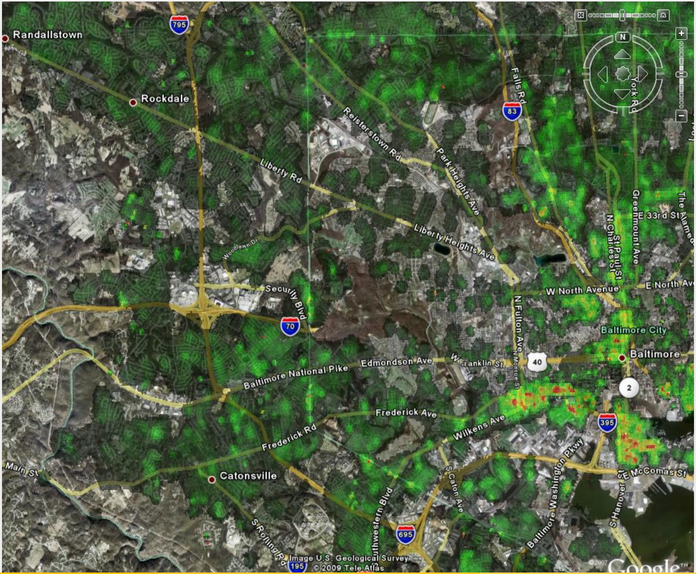
Offender information

- Population density for hispanic 18-34 year old men



Offender information

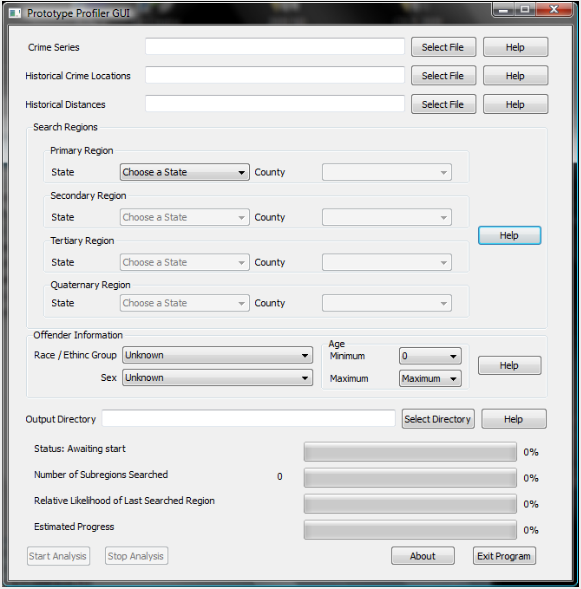
- Population density for white 18-34 year old men



Software

- We have developed a new tool for the geographic profiling problem.
  - It is free for download and use, and is entirely open source.
  - It is still in the prototype stage.
- Required Input:
  - Crime series locations
  - Representative selection of the locations of historically similar crimes, (as determined by the analyst) to estimate target attractiveness
  - Geographic boundaries of the jurisdiction(s) reporting the crime series and historical crimes
  - Available demographic information about the offender, if any
  - Locations of both anchor points and crime sites of historically similar crimes (as determines by the analyst) to estimate the distribution of average offense distances
- The code will then automatically
  - Calculate an estimate of the target attractiveness distribution
  - Estimate the prior distribution of anchor points, assuming anchor point density is proportional to population density
  - Estimate the prior distribution of average offense distances

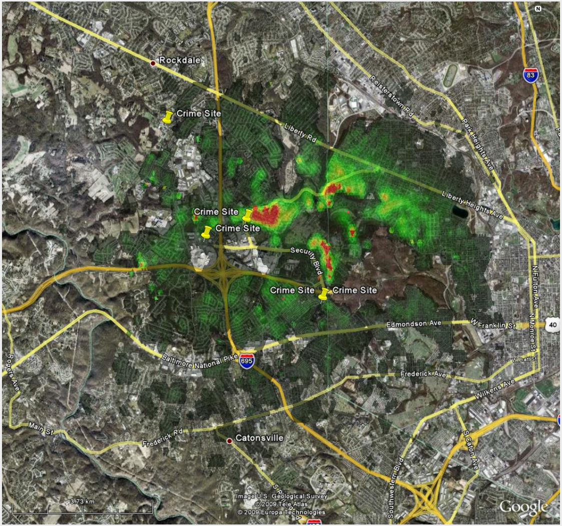
Software



Prototype Profiler GUI
Crime Series
Select File
Help
Historical Crime Locations
Select File
Help
Historical Distances
Select File
Help
Search Regions
Primary Region
State
Choose a State
County
Secondary Region
State
Choose a State
County
Tertiary Region
State
Choose a State
County
Quaternary Region
State
Choose a State
County
Help
Offender Information
Race / Ethinc Group
Unknown
Sex
Unknown
Age
Minimum
0
Maximum
Maximum
Help
Output Directory
Select Directory
Help
Status: Awaiting start
0%
Number of Subregions Searched
0
0%
Relative Likelihood of Last Searched Region
0%
Estimated Progress
0%
Start Analysis
Stop Analysis
About
Exit Program

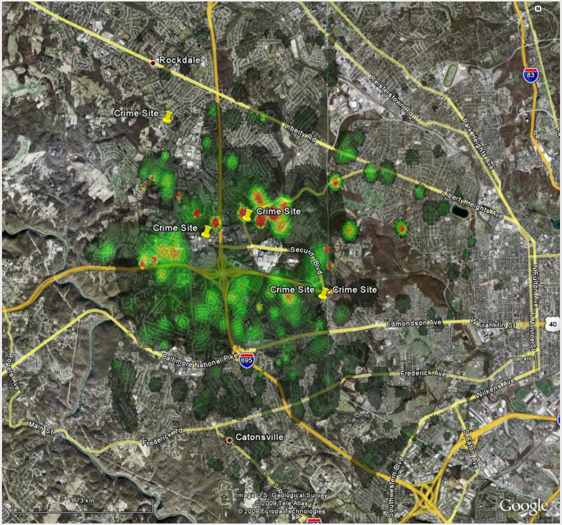
Final Results

- Here is the proposed final search area for the convenience store series of our example



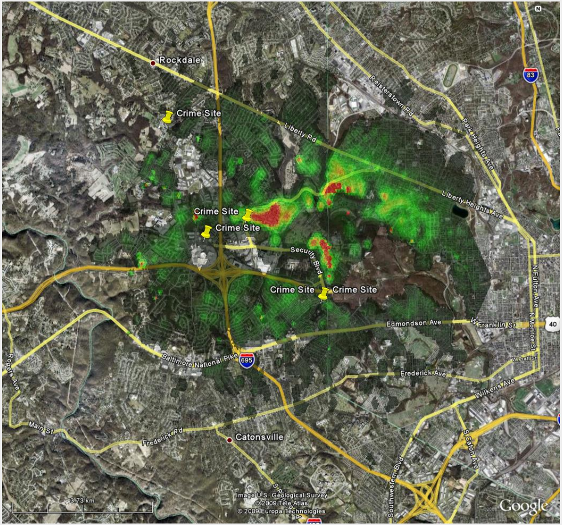
Final Results

- Here is the proposed final search area is we assume our offender is an asian male aged 18-34



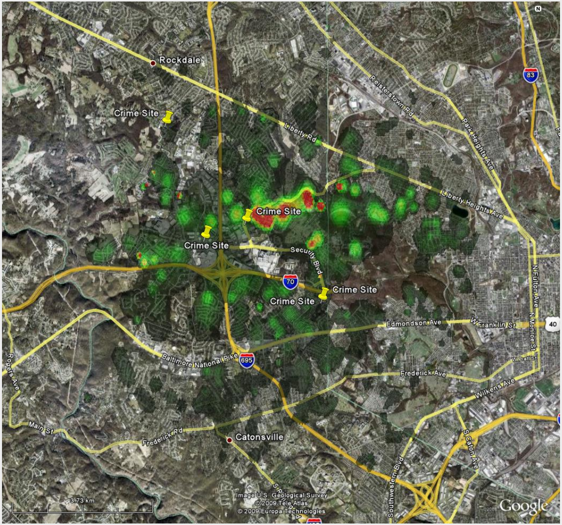
Final Results

- Here is the proposed final search area is we assume our offender is a black male aged 18-34



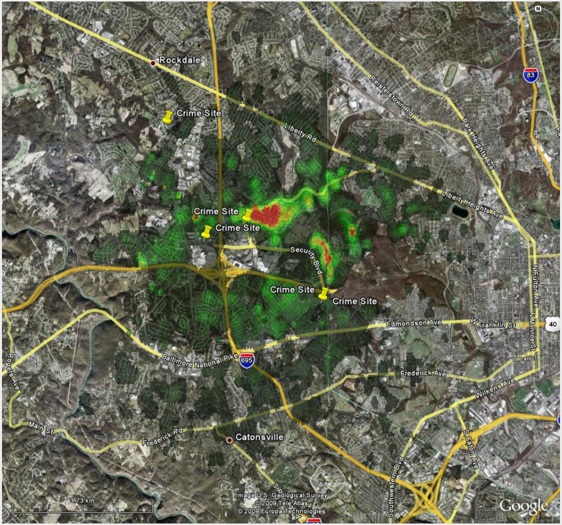
Final Results

- Here is the proposed final search area is we assume our offender is a hispanic male aged 18-34



Final Results

- Here is the proposed final search area is we assume our offender is a white male aged 18-34



The Software

- The current software prototype has been provided to a number of police departments, and is currently being evaluated for effectiveness

Open Questions

- How far are offenders willing to travel?
  - What is the best estimate for their distance decay function?
  - Does the distance decay depend on the offense type? If so, how?
  - Does the distance travelled vary with the home location of the offender? If so, how?
- How do offenders select their targets?
  - Are they chosen independently? If not, what is the best form of the relationship?
  - Does the choice of targets vary over time? If so, how?
  - Is the selection of a target influenced by the behavior of other offenders? How?
- How do we incorporate geography?
  - How do we incorporate the structure of the local road network? What is the best mathematical approach? Is it computationally feasible? What are the best data sources?
  - At the address level, geography is discrete rather than continuous, and has a vertical dimension. How should this be modeled?

Questions?

Mike O'Leary
Department of Mathematics
Towson University

moleary@towson.edu