Implementing a Bayesian Approach for the Geographic Profiling Problem

Mike O'Leary

Towson University

The First International Conference on Computing for Geospatial Research & Applications

Washington, DC
June 22, 2010

Acknowledgments

- This work has been supported by the National Institute of Justice, through grant 2009-SQ-B9-K014
- We thank the Baltimore County Police Department for providing the data used in this study
- We thank Phil Canter from the Baltimore County Police Department for his assistance.

The Geographic Profiling Problem

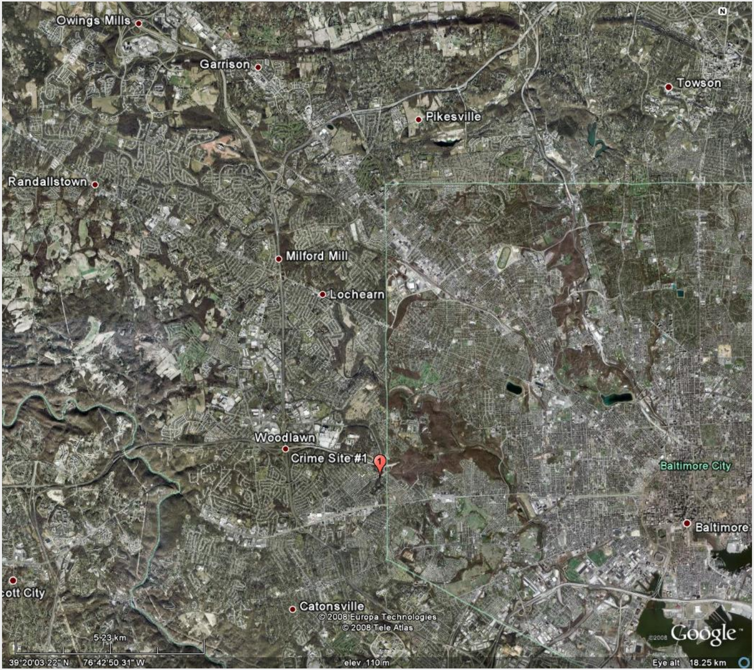
The Geographic Profiling Problem

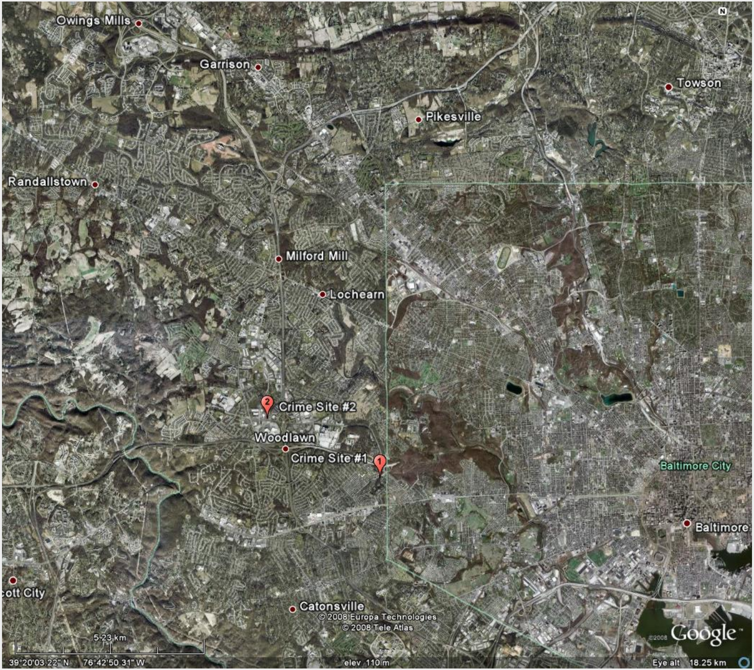
How can we estimate for the location of the anchor point of a serial offender from knowledge of the locations of the offender's crime sites?

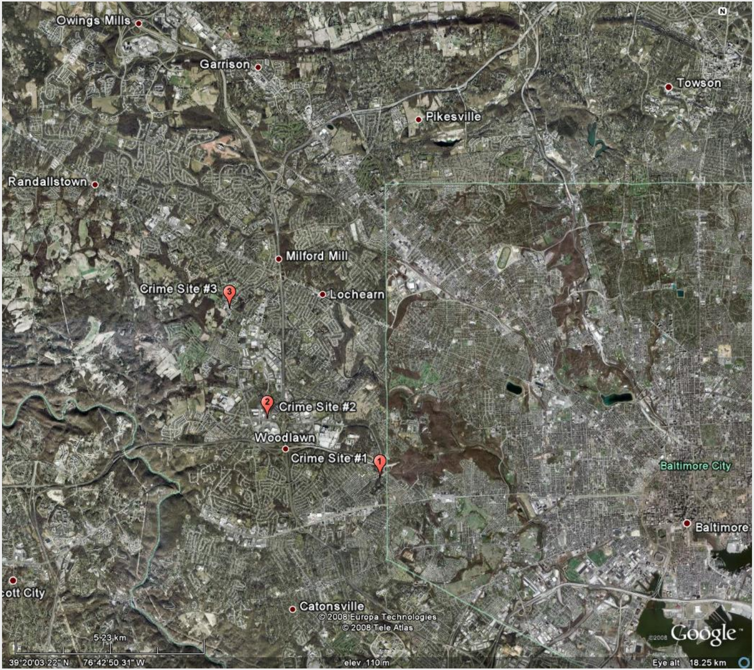
- The anchor point can be the offender's place of residence, place of work, or some other location important to the offender.

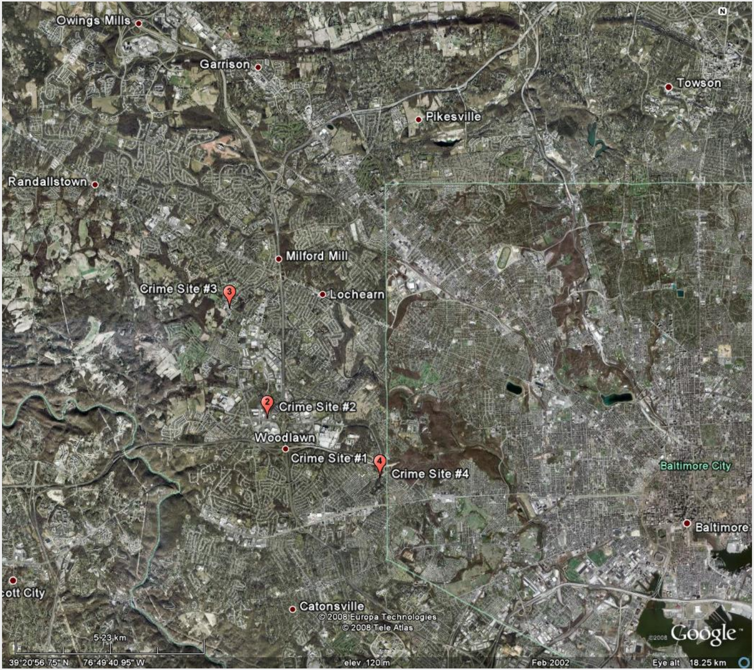
Example- Convenience Store Robberies

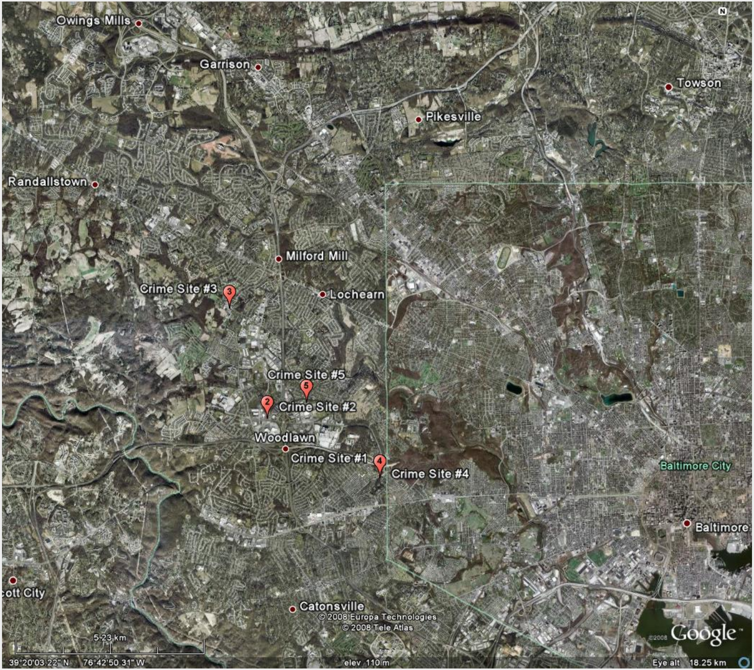
Date	Time	Location		Target
		Latitude	Longitude	
March 8	12:30 pm	-76.71350	39.29850	Speedy Mart
March 19	4:30 pm	-76.74986	39.31342	Exxon
March 21	4:00 pm	-76.76204	39.34100	Exxon
March 27	2:30 pm	-76.71350	39.29850	Speedy Mart
April 15	4:00 pm	-76.73719	39.31742	Citgo
April 28	5:00 pm	-76.71350	39.29850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

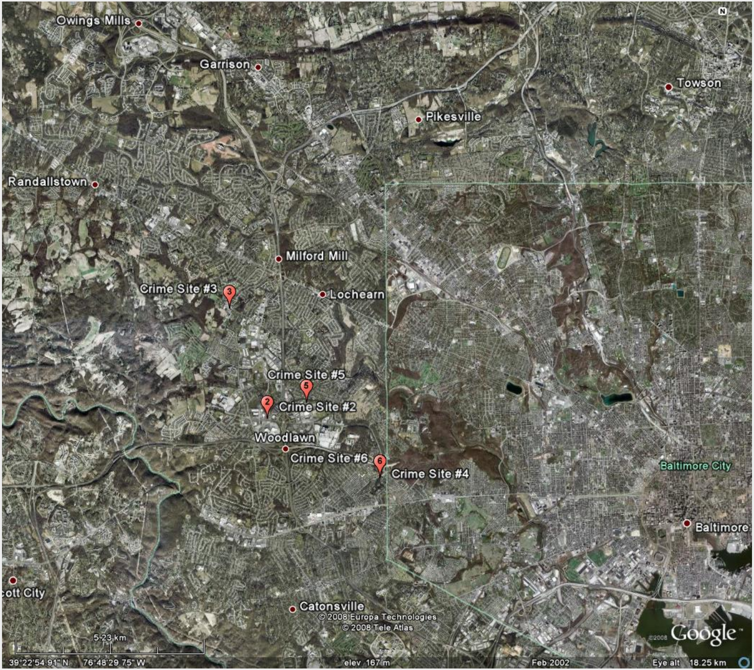

Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Lochearn
Woodlawn
Crime Site #1
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°20'03.22" N 76°42'50.31" W
elev 110 m
Eye alt 18.25 km
Google


Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Lochearn
Crime Site #2
Woodlawn
Crime Site #1
Baltimore City
Baltimore
cott 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
Google


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


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

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 $d(\mathbf{x}, \mathbf{y})$.

Our 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 $\mathrm{P}(\mathbf{x}|\mathbf{z}, \alpha)$

A New Mathematical Approach

- Our mathematical problem then becomes the following:
  - Given a sample $x_1, x_2, \ldots, x_n$ (the crime sites) from a probability distribution $P(\mathbf{x}|\mathbf{z}, \alpha)$, estimate the parameter $\mathbf{z}$ (the anchor point).
- This is a well-studied mathematical problem
- One approach is the theory of *maximum likelihood*.
  - Construct the likelihood function

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

  - 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

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

- The primary disadvantage is that this produces a point estimate.

- 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
$$P(\mathbf{z}, \alpha | \mathbf{x}_1, \dots, \mathbf{x}_n) = \frac{P(\mathbf{x}_1, \dots, \mathbf{x}_n | \mathbf{z}, \alpha)\pi(\mathbf{z}, \alpha)}{P(\mathbf{x}_1, \dots, \mathbf{x}_n)}.$$
- If we assume that the offender's choice of crime sites are mutually independent, then
$$P(\mathbf{x}_1, \dots, \mathbf{x}_n | \mathbf{z}, \alpha) = P(\mathbf{x}_1 | \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n | \mathbf{z}, \alpha)$$
giving us the relationship
$$P(\mathbf{z}, \alpha | \mathbf{x}_1, \dots, \mathbf{x}_n) \propto P(\mathbf{x}_1 | \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n | \mathbf{z}, \alpha) H(\mathbf{z}) M(\alpha).$$
- 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 $P(\mathbf{x}|\mathbf{z}, \alpha)$ 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

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

Models of Offender Behavior

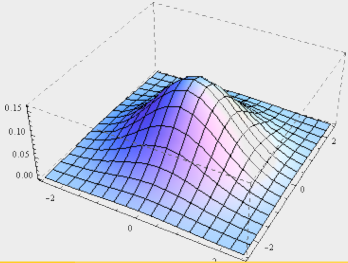
- Suppose that we assume that offenders choose offense sites according to a normal distribution, so that

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

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

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

  - 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

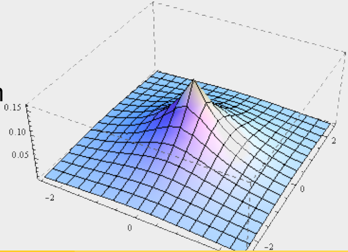
- Suppose that we assume that offenders choose offense sites according to a negative exponential distribution, so that

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

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

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

    - 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

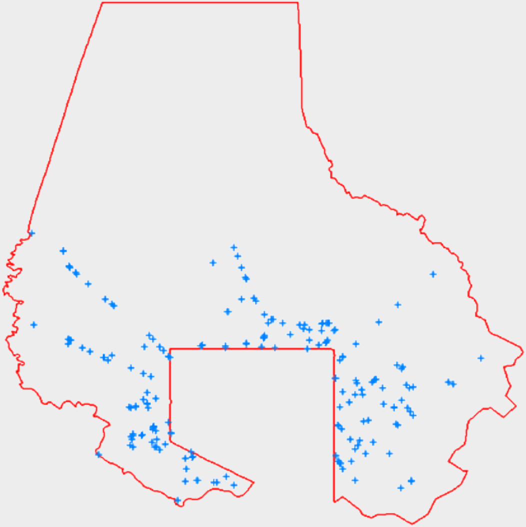
$$P(\mathbf{x}|\mathbf{z},\alpha) = D(d(\mathbf{x},\mathbf{z}),\alpha)\cdot G(\mathbf{x})\cdot N(\mathbf{z})$$

  - 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 $\mathrm{G}(\mathbf{x})$
    - 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 $K(\mathbf{x}|\lambda)$ with bandwidth λ given by

$$K(\mathbf{x}|\lambda) = \begin{cases} \dfrac{3}{\pi\lambda^6}(|\mathbf{x}|^2 - \lambda^2)^2 & \text{if } |\mathbf{x}| \leqslant \lambda, \\ 0 & \text{if } |\mathbf{x}| \geqslant \lambda. \end{cases}$$

- Construct an approximation of our target attractiveness by calculating

$$G(\mathbf{x}) = \sum_{i=1}^{N} K(\mathbf{x} - \mathbf{c}_i|\lambda)$$

- The question of the optimal selection of the bandwidth parameter λ remains open.
  - One reasonable choice is the mean nearest neighbor distance

Distance Decay

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

$$\mathrm{D}_{\text{one-dim}}(\mathrm{d}|\alpha) = 2\pi\mathrm{d} \cdot \mathrm{D}(\mathrm{d}|\alpha)$$

- In particular, $\mathrm{D}_{\text{one-dim}}(\mathrm{d}|\alpha) \to 0$ as $\mathrm{d} \to 0$, regardless of the particular choice of $\mathrm{D}(\mathrm{d}|\alpha)$, provided $\mathrm{D}(0|\alpha) < \infty$.
- In what follows, we use a normal form for the two dimensional distance decay

$$\mathrm{D}(\mathrm{d}(\mathbf{x},\mathbf{z}),\alpha) = \frac{1}{4\alpha^2}\exp\left(-\frac{\pi}{4\alpha^2}\,\mathrm{d}(\mathbf{x},\mathbf{z})^2\right)$$

so that the distribution of distances is Rayleigh

$$\mathrm{D}_{\text{one-dim}}(\mathrm{d}|\alpha) = \frac{\pi\mathrm{d}}{2\alpha^2}\exp\left(-\frac{\pi\mathrm{d}^2}{4\alpha^2}\right)$$

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 $A(d)$.
- 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 $D_{\text{one-dim}}(d|\alpha)$
- Then
$$A(d) = \int_0^\infty D_{\text{one-dim}}(d|\alpha)M(\alpha)\,d\alpha$$
- Since $A(d)$ can be measured and $D_{\text{one-dim}}(d|\alpha)$ modeled, we can solve this equation for the prior mean offense travel distance $M(\alpha)$

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

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

  - $G = G_{jk} = \dfrac{\pi S \delta}{2} \dfrac{(j - \frac{1}{2})}{(k - \frac{1}{2})^2} \exp\left(-\dfrac{\pi}{4} \dfrac{(j - \frac{1}{2})^2}{(k - \frac{1}{2})^2} \right)$
  - $\mathbf{a} = (a_1, a_2, \ldots, a_N)$
  - $\mathbf{m} = (M(\alpha^*_1), M(\alpha^*_2), \ldots, M(\alpha^*_N))$

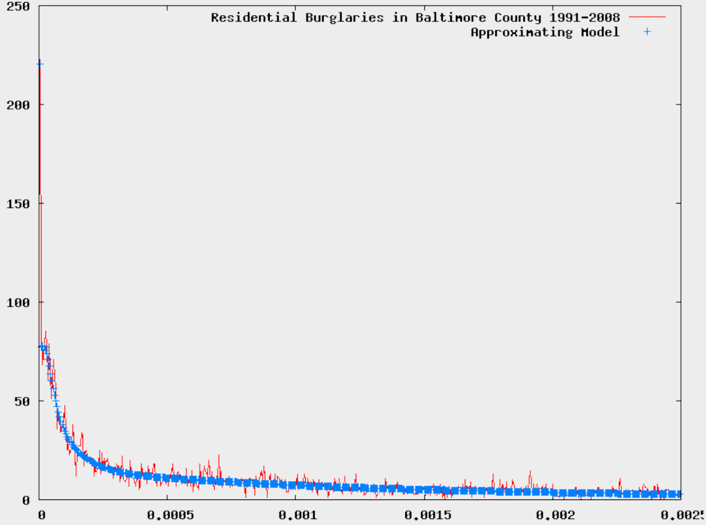
Distance Decay: Solving the Integral Equation

- 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 $S = \operatorname{diag}(s_1, s_2, \dots, s_N)$, then $s_j \to 0$ with no appreciable gaps.
    - G has ill-defined numerical rank.
- We can apply Tikhonov regularization; i.e. replace $S^\dagger$ with

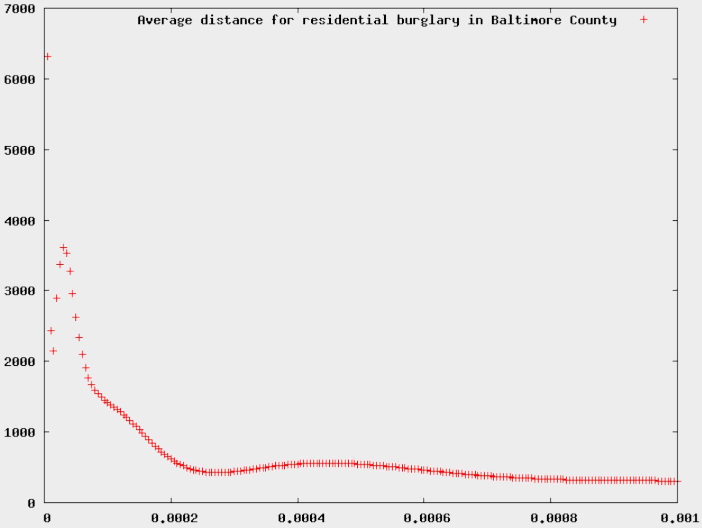
$$S_\lambda^\dagger = \operatorname{diag}\left(\frac{s_1}{s_1^2 + \lambda^2}, \frac{s_2}{s_2^2 + \lambda^2}, \dots, \frac{s_N}{s_N^2 + \lambda^2} \right)$$

then $\mathbf{m} = \mathrm{G}_\lambda^\dagger \mathbf{a}$ can be calculated.

Distance Decay: Residential Burglaries in Baltimore County- Model Fit



Distance Decay: Residential Burglaries in Baltimore County- Prior Distribution



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 $H(\mathbf{z}) = \sum_{i=1}^{N_{\text{blocks}}} = p_i K(\mathbf{z} - \mathbf{q}_i | \sqrt{A_i})$
    - 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 $G(\mathbf{x})$

Geographic Triangulation

- On the geographic region under study, we construct a pair of meshes of equilateral triangles
  - A coarse mesh
  - A fine mesh
- This approach lets us precompute the values of the functions G and H at the centroids of each element in the mesh.
  - Kernel density calculations of the form $G(\mathbf{x}) = \sum_{i=1}^{N} K(\mathbf{x} - \mathbf{c}_i | \lambda)$ are computationally expensive for large N.

Normalization

- To evaluate the normalization function

$$N(\mathbf{z}, \alpha) = \left[\iint D(d(\mathbf{x}, \mathbf{z}), \alpha) G(\mathbf{x}) dx^{(1)} dx^{(2)} \right]^{-1}$$

define

$$I(\mathbf{z}, \alpha) = \iint D(d(\mathbf{x}, \mathbf{z}), \alpha) G(\mathbf{x}) dx^{(1)} dx^{(2)}$$

- Then for any mesh of triangles Δ, the midpoint rule gives

$$\begin{aligned} I(\mathbf{z}, \alpha) &= \iint D(d(\mathbf{x}, \mathbf{z}), \alpha) G(\mathbf{x})\, dx^{(1)}\, dx^{(2)} \\ &\approx \frac{3\sqrt{3}}{16} \frac{1}{\alpha^2} \sum_{T \in \Delta} R_T^2 \exp\left(-\frac{\pi}{4\alpha^2} d(\mathbf{x}_T, \mathbf{z})^2 \right) G(\mathbf{x}_T) \end{aligned}$$

where R_T is the circumradius and x_T is the centroid of the triangle T.

Normalization

- These are computationally expensive to calculate when the number of triangles in the mesh Δ is large.
- To reduce the computational cost, start with the coarse mesh and note that
  - If $\mathsf{T} \in \Delta$ satisfies $\mathrm{d}(\mathbf{x}_\mathsf{T}, \mathbf{z}) > 3\alpha$, then
    $$\exp\left(-\frac{\pi}{4\alpha^2}\mathrm{d}(\mathbf{x}_\mathsf{T}, \mathbf{z})^2\right) < e^{-9\pi/4} \approx 0.000851;$$
    - We ignore such triangles in our approximation
  - If $2\alpha < \mathrm{d}(\mathbf{x}_\mathsf{T}, \mathbf{z}) \leqslant 3\alpha$, then
    $$\exp\left(-\frac{\pi}{4\alpha^2}\mathrm{d}(\mathbf{x}_\mathsf{T}, \mathbf{z})^2\right) < e^{-\pi} \approx 0.0432$$
    - All such triangles are retained
  - If $\mathrm{d}(\mathbf{x}_\mathsf{T}, \mathbf{z}) \leqslant 2\alpha$, then subdivide the coarse triangle into its fine subtriangles.

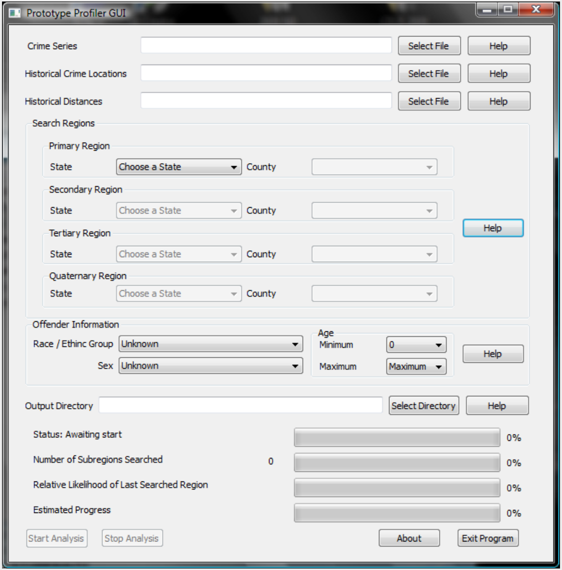
Normalization

- This is still computationally very expensive!
  - We want to estimate $\mathrm{P}(\mathbf{z})$ on roughly a block scale across one or more counties
  - Each calculation of P requires the evaluation of the double integral for I across one or more counties.
- Though $\mathrm{P}(\mathbf{z})$ may vary rapidly, the function $\mathrm{I}(\mathbf{z}, \alpha)$ does not; instead it varies on a size scale of roughly α
  - Our approach is:
    1. Calculate $\mathrm{I}(\mathbf{z}, \alpha)$ for each α and each $\mathbf{z}$ in the coarse mesh.
    2. Interpolate these values into the fine mesh for large values of α.
    3. Use Hermite approximation on the interpolants to craft an interpolant for small α.
  - Hermite interpolation is reasonable because both $\mathrm{I}(\mathbf{z}, \alpha = 0)$ and $\frac{\partial \mathrm{I}}{\partial \alpha}(\mathbf{z}, \alpha = 0)$ can be evaluated analytically, while we know that interpolation gives good results for large α.

The Geographic Profiling Problem

- We have developed software that implements these methods and released it to police agencies for evaluation.
  - It is free for download and use, and is entirely open source.
  - It is still in the prototype stage.
  - The tool is designed to be simple and easy to use.

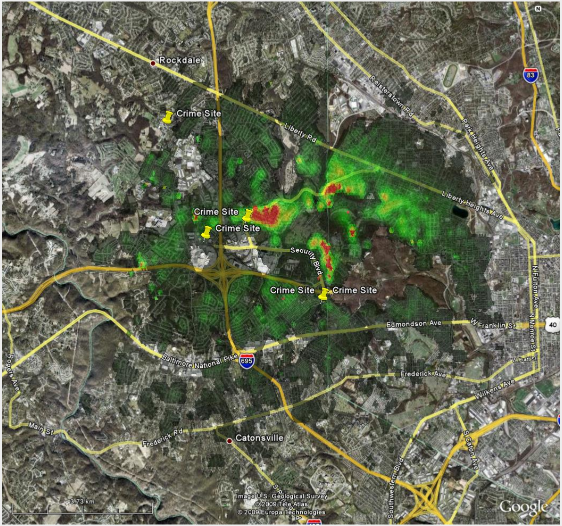
The Tool



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

Sample Results

- When the program runs, it produces an estimate for the offender's anchor point



Questions?

Mike O'Leary
Department of Mathematics
Towson University
moleary@towson.edu

`http://pages.towson.edu/moleary/Profiler.html`