Applying Mathematics....

... to catch criminals

Mike O'Leary

Department of Mathematics
Towson University

American Statistical Association
February 5, 2015

Acknowledgments

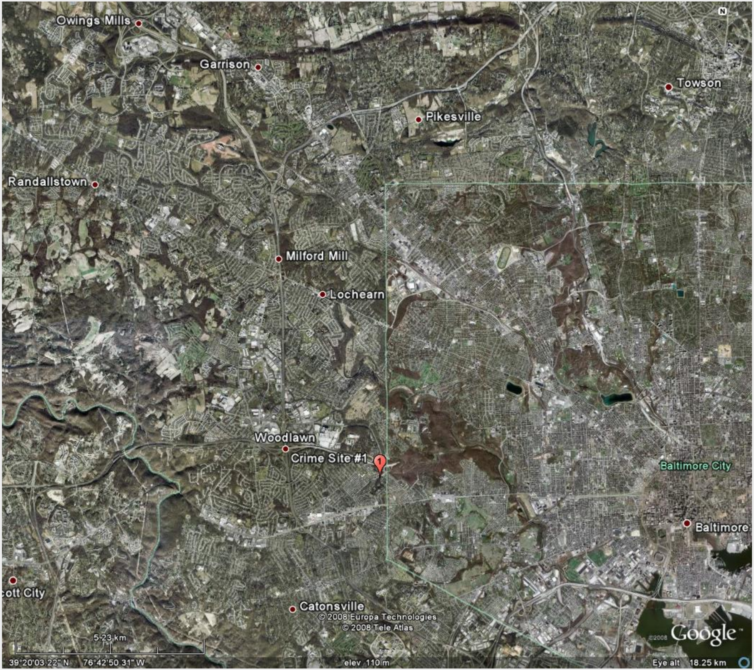
- This began as a project of the Towson University Applied Mathematics Laboratory
  - Undergraduate research projects in mathematics sponsored by companies and government agencies
  - Student participants in 2005-06, 2006-07, and 2007-08:
    - Lauren Amrhine, Brooke Belcher, Brandie Biddy, Colleen Carrion, Chris Castillo, Paul Corbitt, Gregory Emerson, Adam Fojtik, Yu Fu, Natasha Gikunju, Laurel Mount, Kristopher Seets, Jonathan Vanderkolk, Grant Warble, Ruozhen Yao, Melissa Zimmerman
- 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

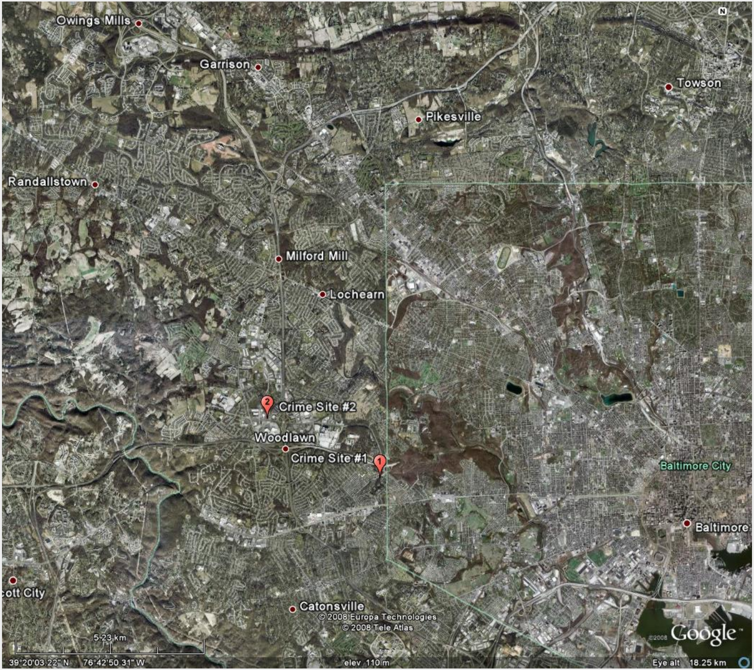
- The geographic profiling problem is to estimate the location of the home base of a serial criminal from the known locations of the elements of the offender's crimes.
  - The home base is also called the anchor point of the offender. It may be the offenders home, the home of a relative, a place of work, or even a favorite bar.
- This is an *operational* problem.
- We have developed a new tool for the geographic profiling problem.
  - It is free for download and use, and is entirely open source.
    - `http://pages.towson.edu/moleary/Profiler.html`
  - It is still in the prototype stage and is being evaluated by different police agencies across the country.

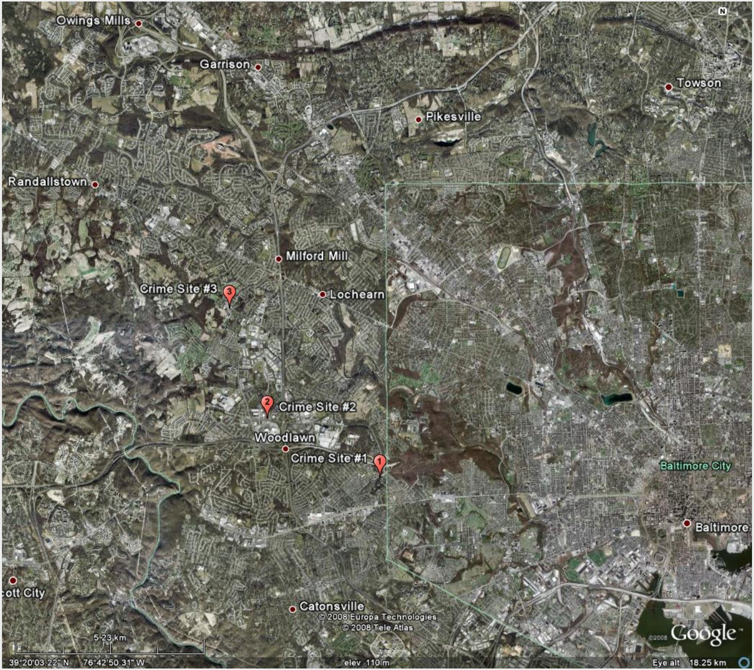
Example- Convenience Store Robberies

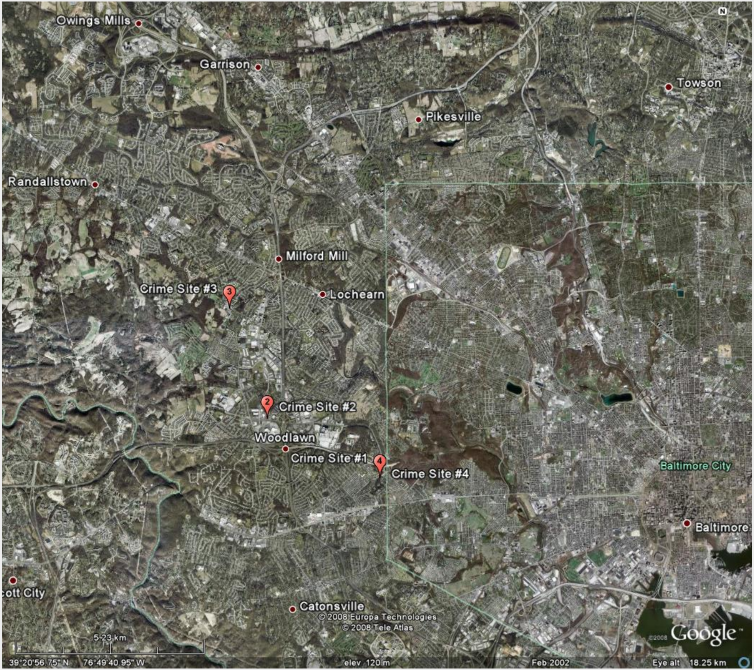
Date	Time	Location		Target
		Latitude	Longitude	
March 8	12:30 pm	-76.71350	39.29850	Speedy Mart
March 19	4:30 pm	-76.74986	39.31342	Exxon
March 21	4:00 pm	-76.76204	39.34100	Exxon
March 27	2:30 pm	-76.71350	39.29850	Speedy Mart
April 15	4:00 pm	-76.73719	39.31742	Citgo
April 28	5:00 pm	-76.71350	39.29850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

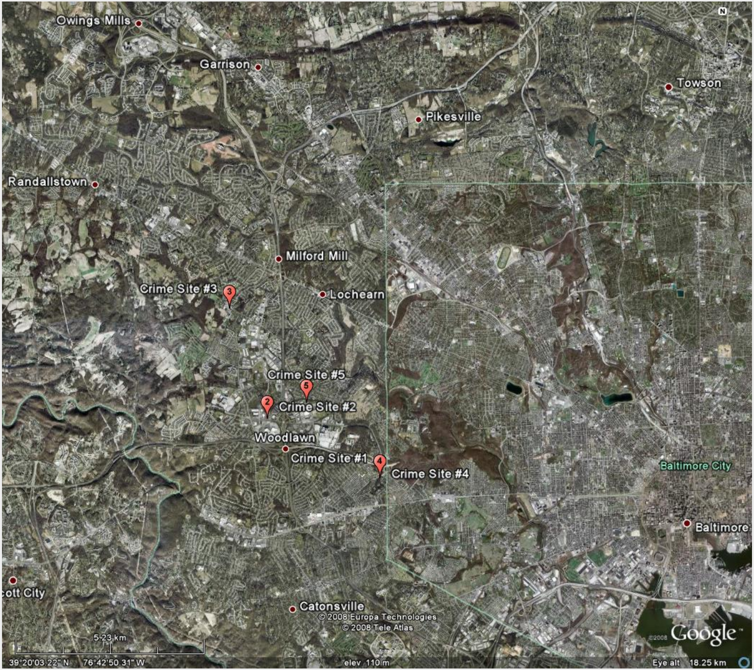


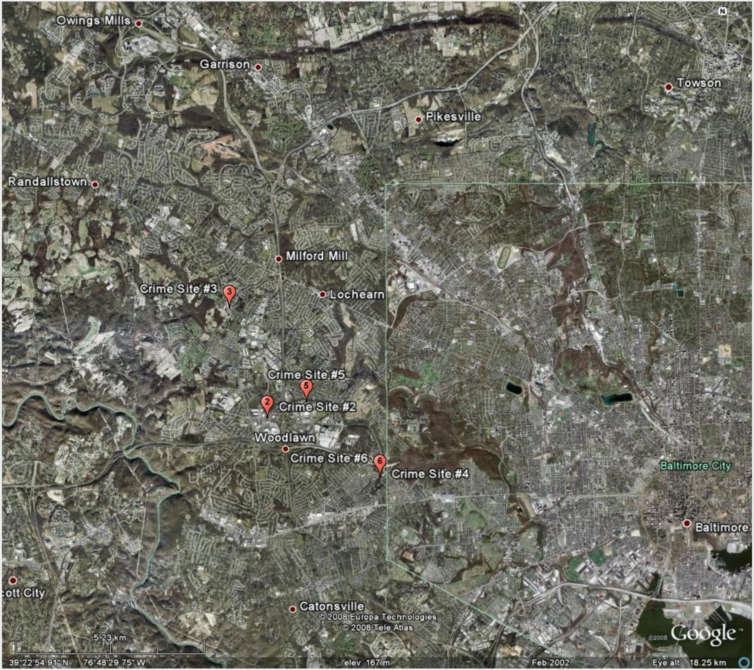
Owings Mills
Garrison
Towson
Pikesville
Randallstown
Milford Mill
Lochearn
Woodlawn
Crime Site #1
Baltimore City
Baltimore
ott City
Catonsville
© 2008 Europa Technologies
© 2008 Tele Atlas
5.23 km
39°20'03.22" N 76°42'50.31" W
elev 110 m
Eye alt 18.25 km
Google


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
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
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

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

How should we measure distances?

- The Euclidean distance $d_2(\mathbf{x}, \mathbf{y}) = \sqrt{(x^{(1)} - y^{(1)})^2 + (x^{(2)} - y^{(2)})^2}$
- The Manhattan distance $d_1(\mathbf{x}, \mathbf{y}) = |x^{(1)} - y^{(1)}| + |x^{(2)} - y^{(2)}|$

How should we measure distances?

- The Euclidean distance $d_2(\mathbf{x}, \mathbf{y}) = \sqrt{(x^{(1)} - y^{(1)})^2 + (x^{(2)} - y^{(2)})^2}$
- The Manhattan distance $d_1(\mathbf{x}, \mathbf{y}) = |x^{(1)} - y^{(1)}| + |x^{(2)} - y^{(2)}|$
- The spherical distance

$$
\begin{aligned}
&d_s((\lambda_1, \phi_1), (\lambda_2, \phi_2)) \\
&\qquad = 2R \arcsin \sqrt{\sin^2\left(\tfrac{1}{2}\Delta\lambda\right) + \cos\lambda_1 \cos\lambda_2 \sin^2\left(\tfrac{1}{2}\Delta\phi\right)}
\end{aligned}
$$

Here ϕ is longitude and λ is latitude.

How should we measure distances?

- The Euclidean distance $d_2(\mathbf{x}, \mathbf{y}) = \sqrt{(x^{(1)} - y^{(1)})^2 + (x^{(2)} - y^{(2)})^2}$
- The Manhattan distance $d_1(\mathbf{x}, \mathbf{y}) = |x^{(1)} - y^{(1)}| + |x^{(2)} - y^{(2)}|$
- The spherical distance

$$\begin{aligned} d_s((\lambda_1, \phi_1), (\lambda_2, \phi_2)) \\ = 2R \arcsin \sqrt{\sin^2\left(\tfrac{1}{2}\Delta\lambda\right) + \cos\lambda_1 \cos\lambda_2 \sin^2\left(\tfrac{1}{2}\Delta\phi\right)} \end{aligned}$$

  Here ϕ is longitude and λ is latitude.
- The highway distance?
- The street distance?

A Mathematical Model

- 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 Mathematical Model

- 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 Mathematical Model

- Our mathematical problem then becomes the following:
  - Given a sample $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ (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

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

Bayesian Analysis

- Suppose that there is only one crime site $\mathbf{x}$. Then Bayes' Theorem implies that
$$P(\mathbf{z}, \alpha \mid \mathbf{x}) = \frac{P(\mathbf{x} \mid \mathbf{z}, \alpha)\pi(\mathbf{z}, \alpha)}{P(\mathbf{x})}$$
    - $P(\mathbf{z}, \alpha \mid \mathbf{x})$ 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
$$\pi(\mathbf{z}, \alpha) = H(\mathbf{z})\pi(\alpha)$$
where $H(\mathbf{z})$ is the prior distribution of anchor points, and $\pi(\alpha)$ is the prior distribution of average offense distances
    - $P(\mathbf{x})$ 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

$$P(\mathbf{z}, \alpha \,|\, \mathbf{x}_1, \dots, \mathbf{x}_n) \propto P(\mathbf{x}_1 \,|\, \mathbf{z}, \alpha) \cdots P(\mathbf{x}_n \,|\, \mathbf{z}, \alpha) H(\mathbf{z}) \pi(\alpha).$$

- 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 Theorem of Geographic Profiling

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 average offense distance is $\pi(\alpha)$ and these are independent of one another.

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

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

Remarks

1. The framework is independent of the significance of the anchor point $\mathbf{z}$
2. This framework holds for any model of offender behavior $P(\mathbf{x} \mid \mathbf{z}, \alpha)$
3. This framework holds for any choice of prior distributions $H(\mathbf{z})$ and $\pi(\alpha)$
4. The framework is independent of the choice of distance metric
5. Geographic features that affect crime selection can be incorporated into the form of $P(\mathbf{x} \mid \mathbf{z}, \alpha)$
6. Geographic features that affect the selection of anchor points are incorporated into the form of $H(\mathbf{z})$
7. The framework provides a prioritized search area; the framework estimates $P(\mathbf{z} \mid \mathbf{x}_1, \dots, \mathbf{x}_n)$ which is the probability density for the offender's anchor point; by definition locations where $P(\mathbf{z} \mid \mathbf{x}_1, \dots, \mathbf{x}_n)$ are larger are more likely to contain the anchor point than regions where it is smaller.

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

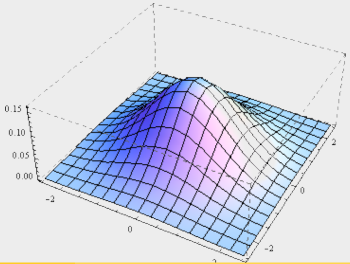
- Suppose that we assume that offenders choose offense sites according to a normal distribution, so that

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

- If we also assume that all offenders have the same average offense distance α, and that all anchor points are equally likely, then

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

  - 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

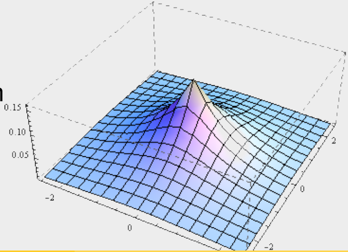
- Suppose that we assume that offenders choose offense sites according to a negative exponential distribution, so that

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

- If we also assume that all offenders have the same average offense distance α, and that all anchor points are equally likely, then

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

  - 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.



Determining a Model for Offender Behavior

- How can we determine a good model for offender behavior?

Determining a Model for Offender Behavior

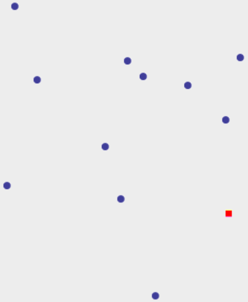
- How can we determine a good model for offender behavior?
- We must look at criminology and at data!

Circle Theory

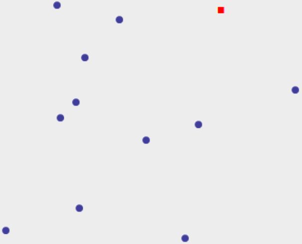
- Canter's Circle hypotheses[1]: Given a series of crimes, construct the circle whose diameter is the segment connecting the two crimes that are farthest apart.
  - If the offender is a *marauder*, then their anchor point will lie in this circle.
  - If the offender is a *commuter*, then their anchor point will lie outside this circle.
    - Note that all of the crimes are not necessarily within the circle.
- For crimes like rape and arson, there is evidence that most offenders are marauders; for crimes like residential burglary the evidence shows a mixture of marauders and commuters.
- This is a binary approach- either someone is a commuter or they are a marauder.
  - This binary approach may not be suitable in many cases.

[1] Canter D. & Larkin, P. (1993). The environmental range of serial rapists. Journal of Environmental Psychology, 13, 63-69.

Which is the Commuter?



or



- Here the crime locations are blue points, and the offender's anchor point is a red square.

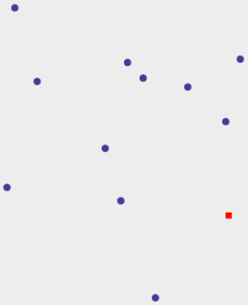
Commuters & Marauders

- We have created a different way to differentiate between commuters and marauders.
- Suppose that:
  - The crimes are at $x_1, x_2, \ldots, x_n$;
  - The offender's anchor point is z.
- For $1 \leqslant p < \infty$ define

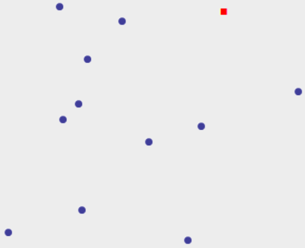
$$\mu_p = \min_{y} \left[\frac{\sum_{i=1}^{n} d(x_i, y)^p}{\sum_{i=1}^{n} d(x_i, z)^p} \right]^{1/p}$$

- Note that $0 \leqslant \mu_p \leqslant 1$.
- Offenders with small μ_p correspond to μ_p-commuters, while offenders with large μ_p correspond to μ_p-marauders.

Which is the Commuter?



or



$\mu_2 = 0.58$
(Canter Marauder)

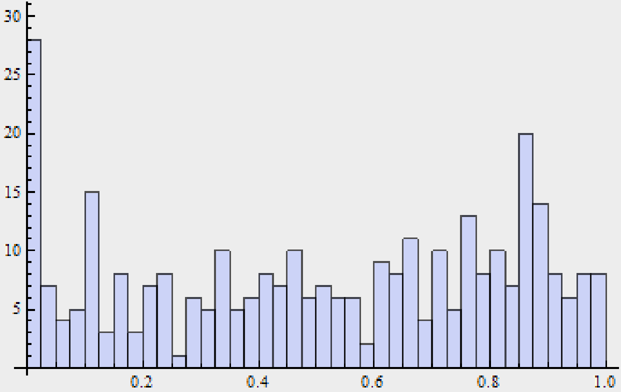
$\mu_2 = 0.56$
(Canter Commuter)

Data

- We have data for residential burglaries in Baltimore County
  - 5863 solved offenses from 1990-2008
  - We have 324 crime series with at least four crimes
    - A series is a set of crimes for which the Age, Sex, Race, DOB and home location of the offender agree.
  - The average number of elements in a series is 8.1, the largest series has 54 elements.
- We have data for non-residential burglaries in Baltimore County
  - 2643 solved offenses from 1990-2008
  - We have 167 crime series with at least three crimes.
  - The average number of elements in a series is 7.87, the largest series has 111 elements.
- We have data for bank robberies in Baltimore County
  - 602 solved offenses from 1993-2009.
  - We have 70 crime series with at least three crimes.
  - The average number of elements in a series is 4.51, the largest series has 15 elements.

Commuters & Marauders

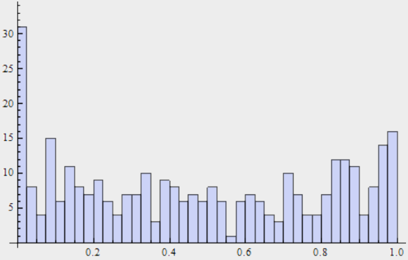
- What is the distribution of μ_2 commuters and marauders for residential burglary?



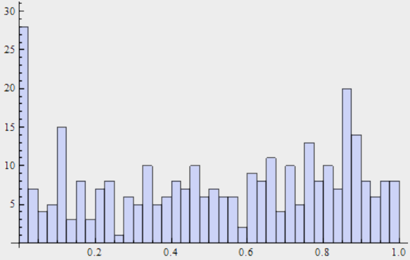
- There does not appear to be a sharp distinction between commuters and marauders in this data

Commuters & Marauders

- There is little difference between μ_1 and μ_2 for residential burglary



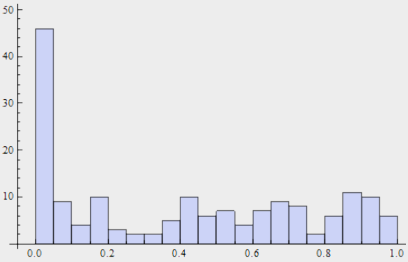
μ_1



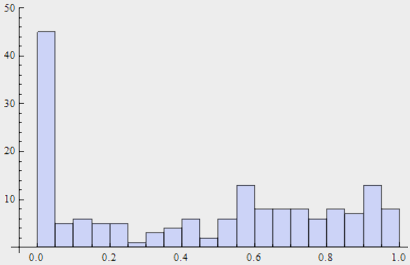
μ_2

Commuters & Marauders

- Non-residential burglary shows a decided preference for commuters.



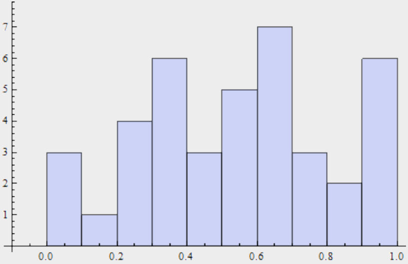
μ_1



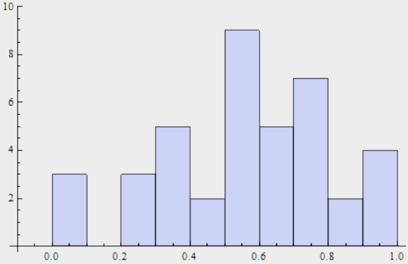
μ_2

Commuters & Marauders

- Bank robberies show a slight preference towards marauders.



μ_1



μ_2

Distance Decay

- The distance decay patterns of offenders are of fundamental importance in the geographic profiling problem.
- Though we have data for the distance from the offenders home to the offense site for a large number of solved crimes, we cannot directly use this information to draw inferences about the behavior of any individual offender.
  - To do so is to commit the *ecological fallacy*.
- There are two sources of variation- the variation within each individual, and the variation between individuals.
  - If all of the individuals behaved in the same fashion, then the aggregate data can be used to draw inference about the (common) underlying behavior.

Distance Decay

- If the only quantity that varies between offenders is the average offense distance, then the resulting scaled distances should exhibit the same behavior regardless of the offender.
  - In particular, this will allow us to aggregate the data across offenders and draw valid inference about the (assumed) universal behavior.
- For each serial offender with crime sites $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ and home $\mathbf{z}$, estimate the average offense distance α by

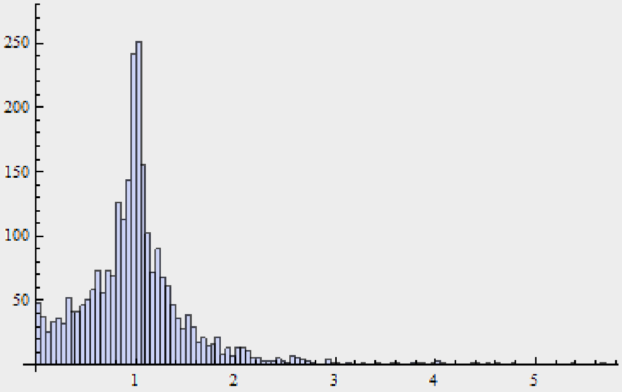
$$\hat{\alpha} = \frac{1}{n} \sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{z})$$

and now consider the set of scaled distances

$$\rho_i = \frac{d(\mathbf{x}_i, \mathbf{z})}{\hat{\alpha}}$$

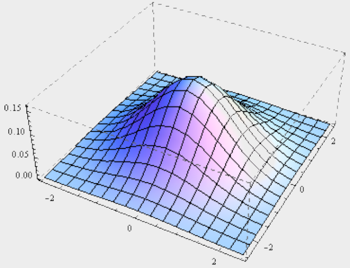
Distance Decay

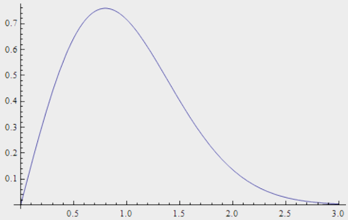
- What do we obtain when we graph not offense distance, but scaled distance?



Distance Decay

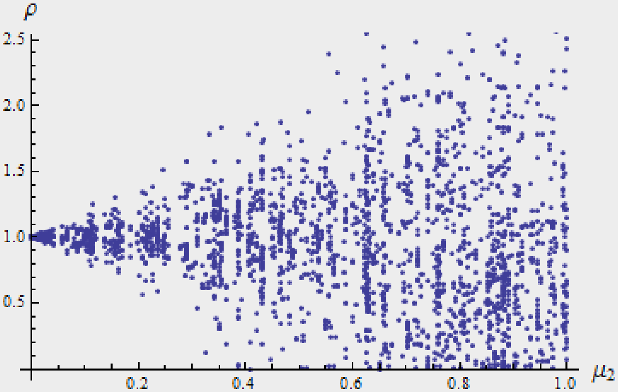
- When considering distance, it is important to realize that it is a *derived* quantity.
  - Offenders do not select a distance- they select a target.
- For example, if the offender selects a target from a two-dimensional normal distribution; then the distribution of distances is a Rayleigh distribution.





Distance Decay

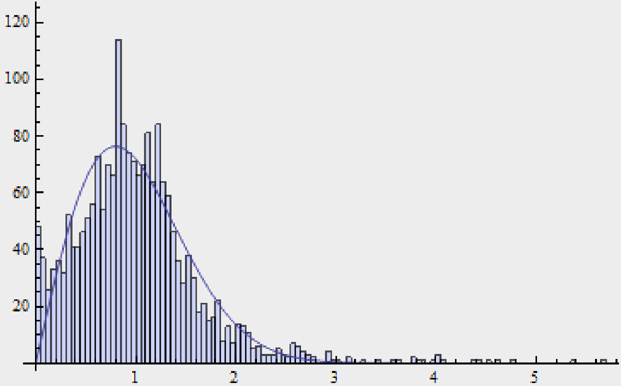
- It is useful to look at the dependence of μ_2 versus ρ for our residential burglars



- Commuters ($\mu_2 \approx 0$) exhibit very different behavior than marauders ($\mu_2 \approx 1$).
- Focus our attention only on non-commuters- say $\mu_2 \geqslant 0.25$.

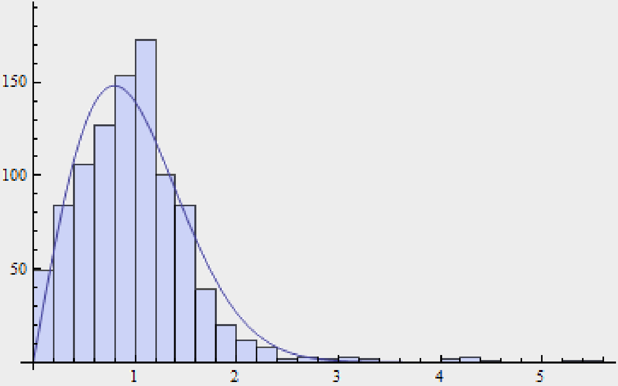
Distance Decay

- If we assume that each offender chooses targets from a two-dimensional normal distribution with their own average offense distance, then the distribution of scaled distances should follow a Rayleigh distribution with mean 1:



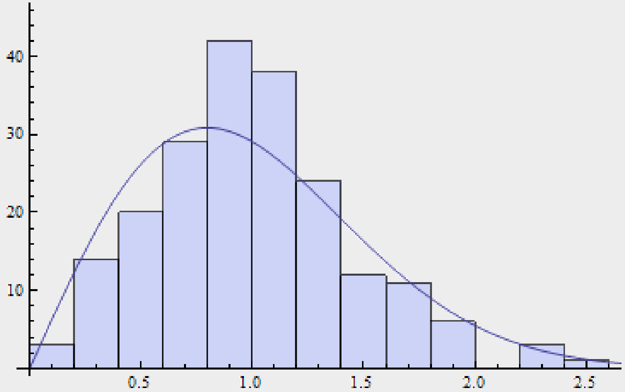
Distance Decay

- The agreement with the Rayleigh distribution does not appear to be happenstance. Here is what occurs for non-residential burglaries with $\mu_2 \geqslant 0.25$



Distance Decay

- Here is what occurs for bank robberies with $\mu_2 \geqslant 0.25$



Distance Decay

- It is possible that these fits are caused by something peculiar to the geography of Baltimore County.
- However, we are not the first to examine scaled distances.
  - Warren, Reboussin, Hazelwood, Cummings, Gibbs, and Trumbetta (1998). *Crime Scene and Distance Correlates of Serial Rape*, Journal of Quantitative Criminology **14** (1998), no. 1, 3559.
  - In that paper, they graphed scaled distances for serial rape:

Distance Decay

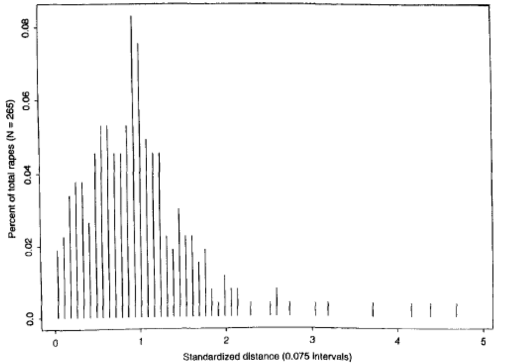
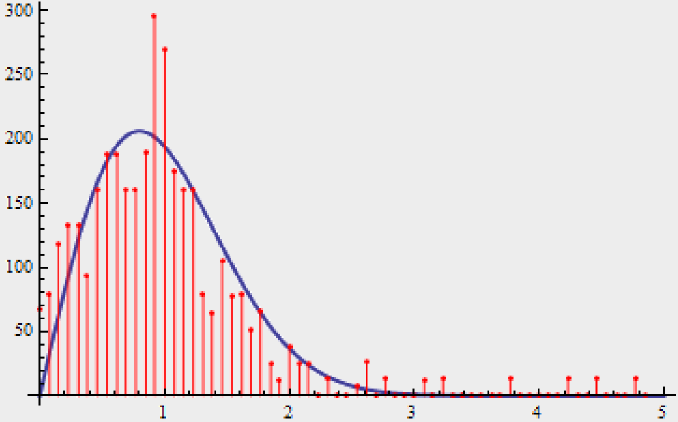


Fig. 2. Proportion of rapes by standardized distance from residence to rape location. Cases with five or more rapes.

two reasons. First, the nonrepresentative nature of the data diminishes the meaningfulness of significance levels. Second, the applied purpose of the paper heightened the need to present the data in a visually clear and practically interpretable form. Distance was found to vary with the demographic characteristics of the offender as well as certain "signature" and "modus

Distance Decay

- Our Rayleigh distribution with mean 1 appears to fit this data as well:

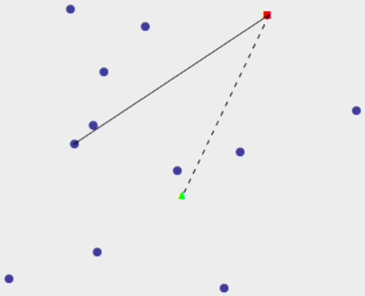


Distance Decay- Caveats

- It is important to note that, though compelling, these graphs do not provide justification that offenders follow a bivariate normal distribution.
    - Agreement is necessary, but not sufficient for this conclusion.
    - There are other two dimensional distributions whose distribution of distances also is Rayleigh.
- We still do not understand the situation yet with commuters.
    - The Warren *et. al.* data is for serial rape, which is known to be well approximated by circle theory- suggesting that this data set may be weighted away from commuters, which our theory does not yet handle.
- Though not presented, we obtained a similar degree of fit using μ_1 instead of μ_2 to characterize commuters and marauders.

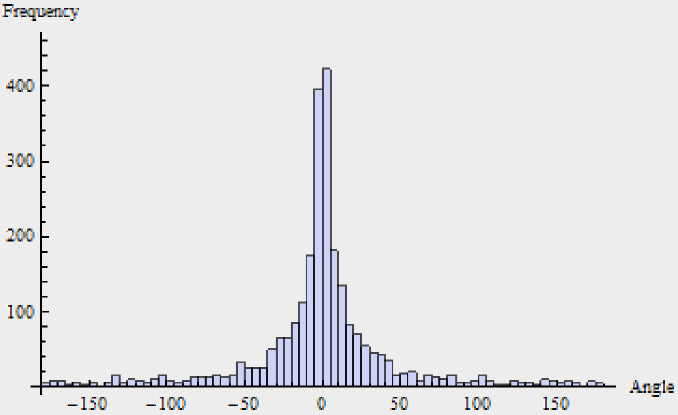
Angular Dependence

- If our idea that the underlying distribution is bivariate normal is correct, then there should be no angular dependence in the results.
- To measure angles, let the blue dots represent crime locations, the red square the anchor point, and the green triangle the centroid of the crime series.
- Then measure the angle between the ray from the anchor point to the crime site and the ray from the anchor point to the centroid.



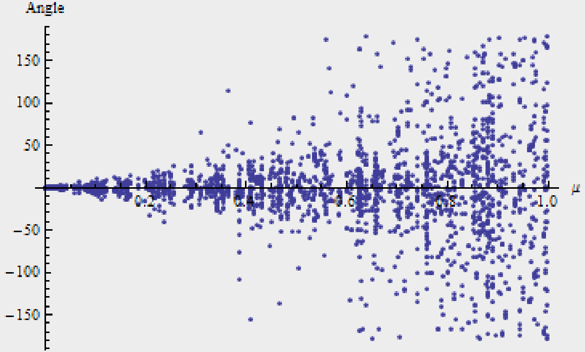
Angular Dependence

- The residential burglary data shows a striking relationship- nearly all of the crime sites lie in the same direction as the centroid.



Angular Dependence

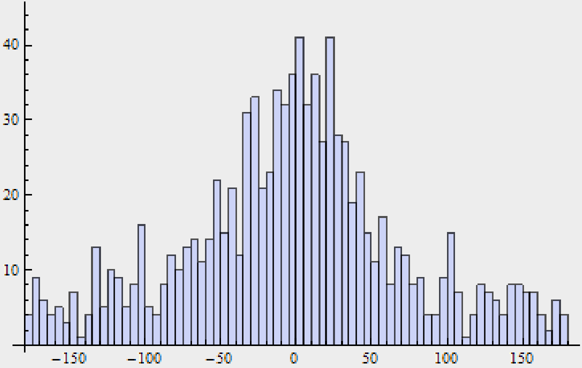
- We can again examine the angular variation as μ_2 varies.



- Even for relatively large values of μ_2, the data is clustered near the zero angle.

Angular Dependence

- The strong central peak remains, even if we restrict our attention to series with $\mu_2 > 0.7$:



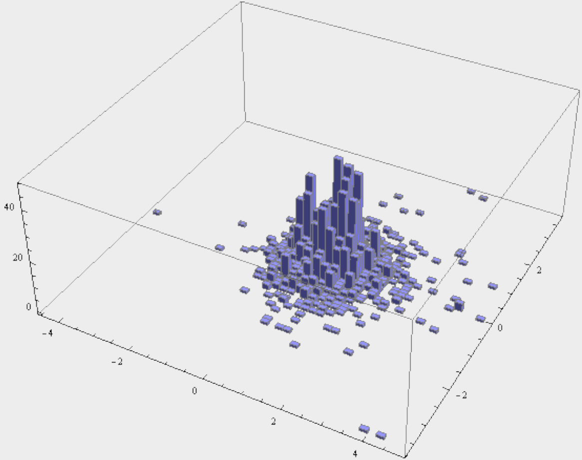
- Note the dramatic changes in the vertical scale between these images!

Angular Dependence

- Clearly there is a strong relationship between the directions the offender took to the different crime sites.
- Moreover, this relationship appears to be strong whether the offender is a commuter or a marauder.
- This suggests that weak information about direction would be more valuable than strong information about distance if one wanted to reduce the area necessary to search for the offender.

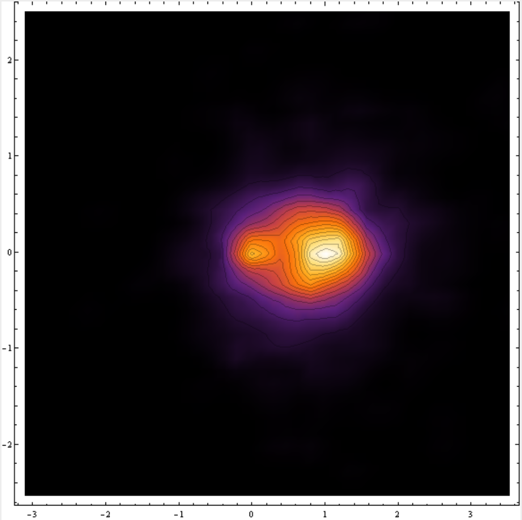
Two-dimensional Distribution

- Plot the histogram of the scaled two dimensional data set; here the offender's home is at the origin, and the centroid of the crime series is at $(x, y) = (1, 0)$.



Two-dimensional Distribution

- Here is another view as a two-dimensional density; note that it is not centered at the origin.



Alternative Hypotheses

- One one hand, the distribution of distances from the offender's home to crime site appears to follow a Rayleigh distribution- at least for marauders.
- On the other hand, it is just as clear that the bivariate distribution is not bivariate normal.
- Indeed, it is clear that there are significant correlations between the locations of the different crime site locations.
  - As evidence, we have the fact that the scaled bivariate distribution clusters not around the offender's home, but around the centroid of the crime series.
- Perhaps we should consider a two stage mixture model:
  - Offenders select a target area
  - Within that target area, offenders select a target.
- Can we model these processes separately?

Models of Offender Behavior

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

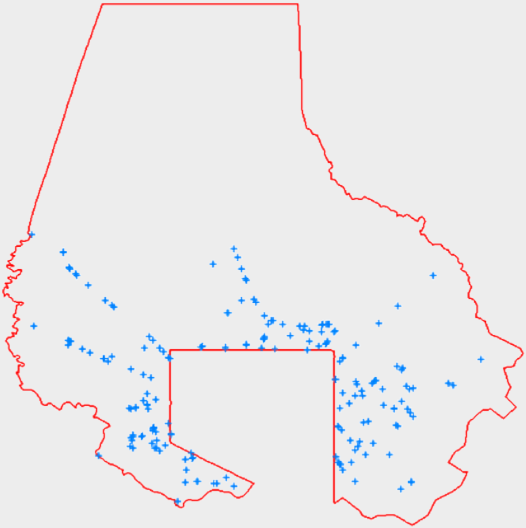
$$P(\mathbf{x} \mid \mathbf{z}, \alpha) = D(d(\mathbf{x}, \mathbf{z}), \alpha) \cdot G(\mathbf{x}) \cdot N(\mathbf{z})$$

  - D models the effect of distance decay using the distance metric $d(\mathbf{x}, \mathbf{z})$
    - Based on the previous work, we use a bivariate normal distribution.
  - 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

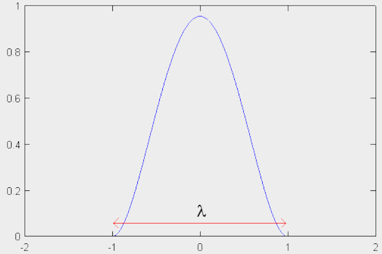
- G models the geographic features that influence crime site selection, with high values indicating the location was more likely to be targeted by an offender.
- How can we calculate G?
  - Use available geographic and demographic data and the correlations between crime rates and these variables that 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



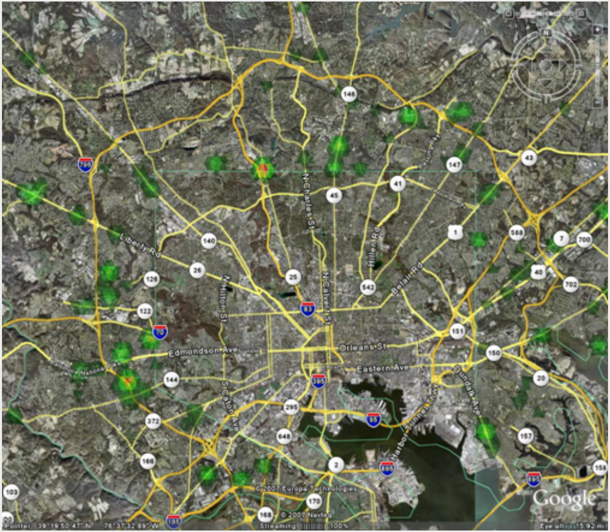
Geographic Features that Influence Crime Selection

- Suppose that historical crimes have occurred at the locations $\mathbf{c}_1, \mathbf{c}_2, \ldots, \mathbf{c}_N$.
- 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

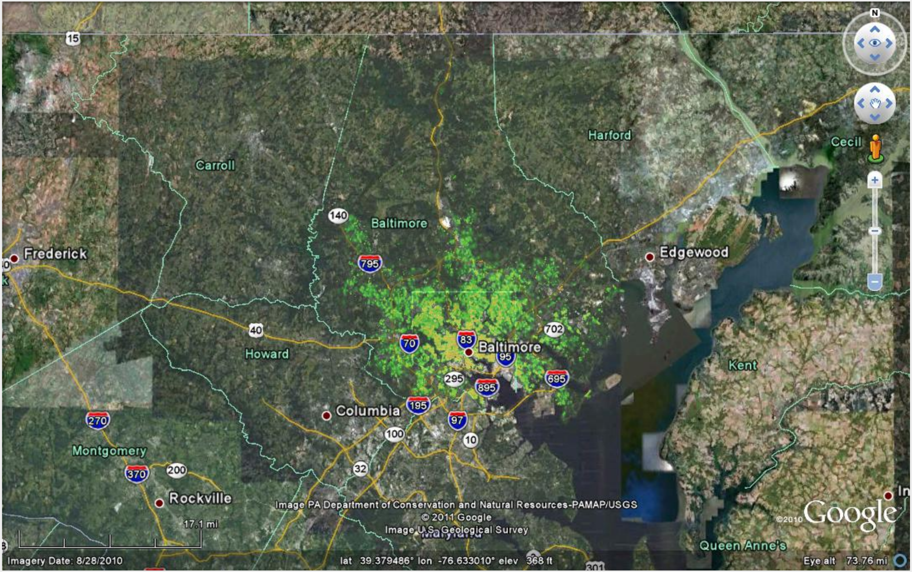
Convenience Store Robberies, Baltimore County



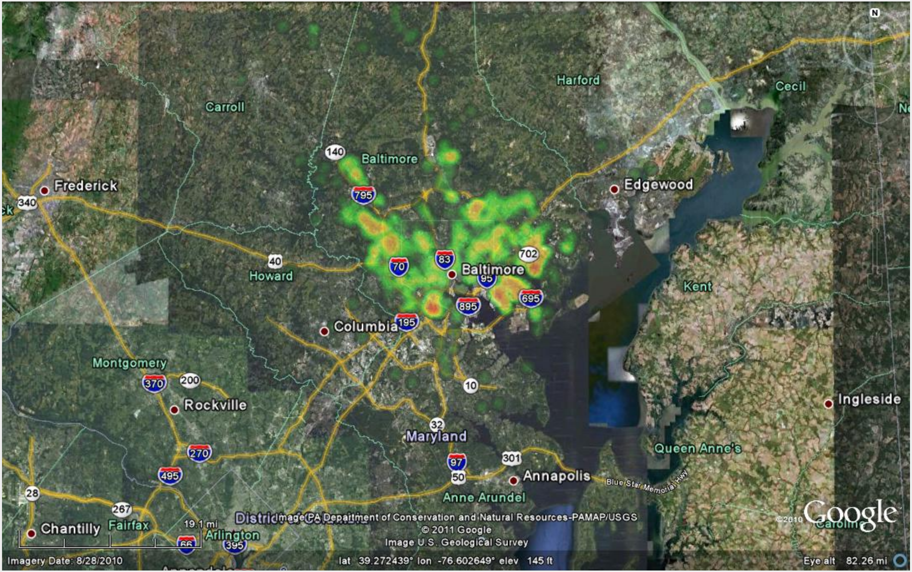
Anchor Points

- 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 $G(\mathbf{x})$

Population Density for Baltimore City & County



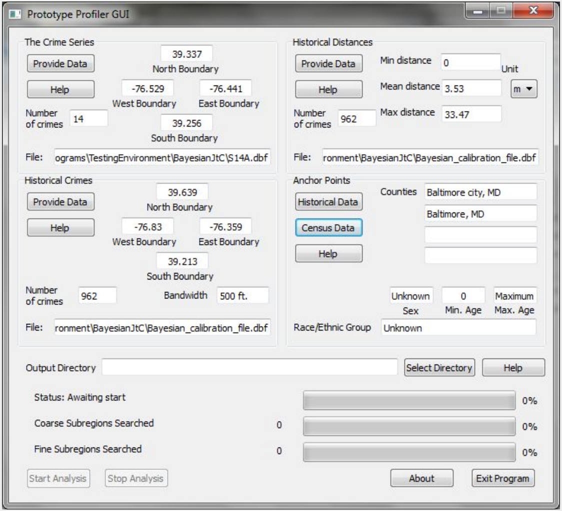
Baltimore County Residential Burglary Offenders



The Tool

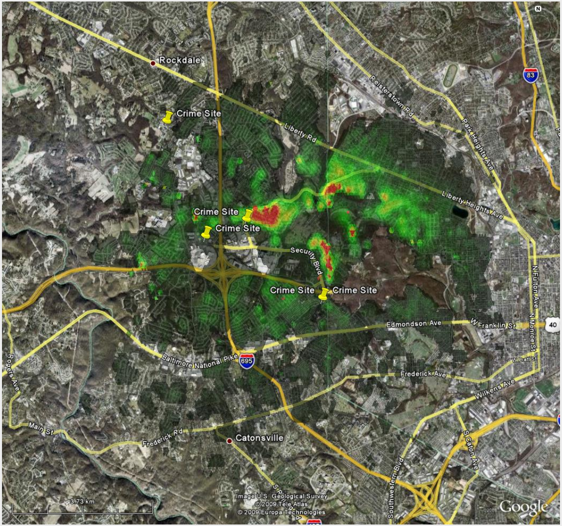
- We have developed a new tool for the geographic profiling problem.
  - It is free for download and use, and is entirely open source.
    - `http://pages.towson.edu/moleary/Profiler.html`
  - It is still in the prototype stage and is being evaluated by different police agencies across the country.

The Tool


Prototype Profiler GUI
The Crime Series
Provide Data
39.337
North Boundary
Help
-76.529
West Boundary
-76.441
East Boundary
Number of crimes
14
39.256
South Boundary
File: ograms\TestingEnvironment\BayesianJtC\S14A.dbf
Historical Distances
Provide Data
Min distance
0
Unit
Help
Mean distance
3.53
m
Number of crimes
962
Max distance
33.47
File: ronment\BayesianJtC\Bayesian_calibration_file.dbf
Historical Crimes
Provide Data
39.639
North Boundary
Help
-76.83
West Boundary
-76.359
East Boundary
39.213
South Boundary
Number of crimes
962
Bandwidth
500 ft.
File: ronment\BayesianJtC\Bayesian_calibration_file.dbf
Anchor Points
Historical Data
Counties
Baltimore city, MD
Baltimore, MD
Census Data
Help
Unknown
Sex
0
Min. Age
Maximum
Max. Age
Race/Ethnic Group
Unknown
Output Directory
Select Directory
Help
Status: Awaiting start
0%
Coarse Subregions Searched
0
0%
Fine Subregions Searched
0
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`

Scaling

- If we want to model the selection of crime sites within a hunting area, we should not use as the length scale the distance from the offender's home to the crime site.
- A reasonable choice for the length scale is the distance from the individual crime sites to the centroid of the crime series.
- For each serial offender with crime sites $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, let

$$\mathbf{c} = \frac{1}{n} \sum_{i=1}^{n} \mathbf{x}_i$$

be the centroid, select the length scale

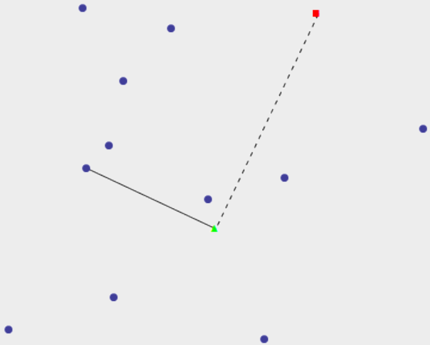
$$\hat{\alpha}_c = \frac{1}{n} \sum_{i=1}^{n} d(\mathbf{x}_i, \mathbf{c})$$

and consider the set of scaled distances

$$\rho_i = \frac{d(\mathbf{x}_i, \mathbf{c})}{\hat{\alpha}_c}$$

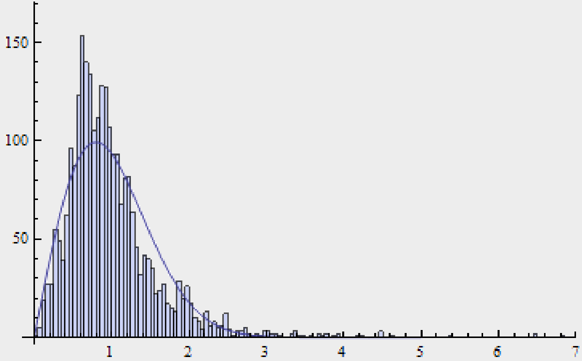
Site Distribution

- The center of our coordinate system is on the centroid of the crime series, angles are measured from the ray from the centroid (green triangle) to the offender's home base (red square).



Site Distribution

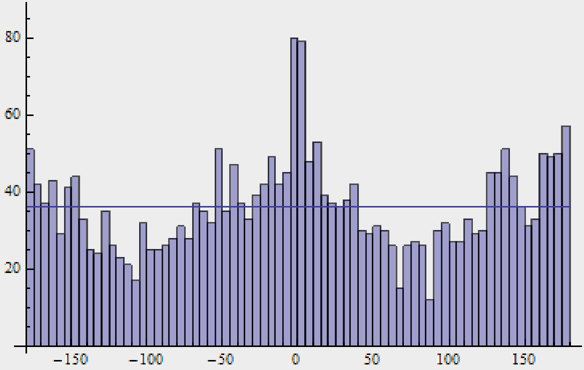
- If we plot the scaled distances from the crime site to the centroid, we again obtain a good match with the Rayleigh distribution.



- This includes all offenders- commuters and marauders.

Site Distribution

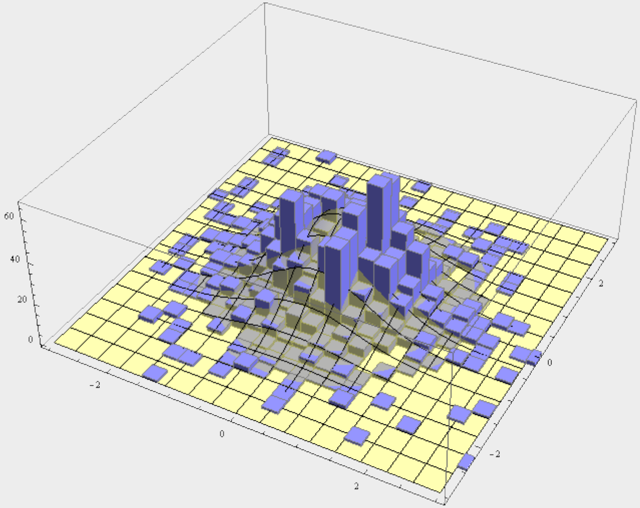
- Now if we plot the angles, we see that the uniform distribution is a much better fit.



- Clearly some anisotropy remains.

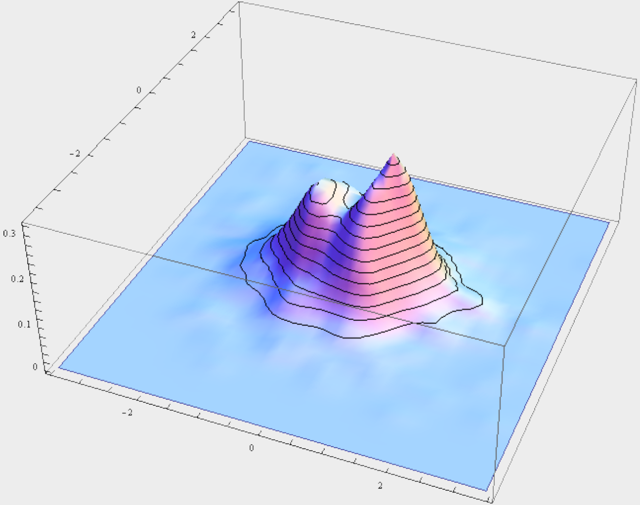
Site Distribution

- We can now directly compare the bivariate normal to the scaled two-dimensional distribution, to see a reasonable fit.



Site Distribution

- The deviation of the scaled distribution from a bivariate normal is more obvious when we smooth the histogram.



Site Distribution: Conclusions

- We have evidence that the distribution of crime site locations is roughly bivariate normal, and centered around the centroid of the crime series.
- No distinction needs to be drawn between commuters and marauders.
- There are noticeable deviations from normality:
  - Directions in line with the offender's home address are preferred to perpendicular directions.
  - There is a preference for crime sites closer to the offender's home address than locations farther away.
  - There appears to be weak evidence for the existence of a buffer zone around the offender's home.
- These hypotheses have only been tested on residential burglaries in Baltimore County.