

Centroids, Clusters, and Crime: Anchoring the Geographic Profiles of Serial Criminals

(Catchy, but informative title)

Team # 12345

(Dated: January 23, 2014)

A particularly challenging problem in crime prediction is modeling the behavior of a serial killer. Since finding associations between the victims is difficult, we predict where the criminal will strike next, instead of whom. Such predicting of a criminal's spatial patterns is called geographic profiling.

Research shows that most violent serial criminals tend to commit crimes in a radial band around a central point: home, workplace, or other area of significance to the criminal's activities (for example, a part of town where prostitutes abound). These "anchor points" provide the basis for our model.

We assume that the entire domain of analysis is a potential crime spot, movement of the criminal is uninhibited, and the area in question is large enough to contain all possible strike points. We consider the domain a metric space on which predictive algorithms create spatial likelihoods. Additionally, we assume that the offender is a violent serial criminal, since research suggests that serial burglars and arsonists are less likely to follow spatial patterns.

There are substantial differences between one anchor point and several. We treat the single-anchor-point case first, taking the spatial coordinates of the criminal's last strikes and the sequence of the crimes as inputs. Estimating the point to be the centroid of the previous crimes, we generate a "likelihood crater," where height corresponds to the likelihood of a future crime at that location. For the multiple-anchor-point case, we use a cluster-finding and sorting method: We identify groupings in the data and build a likelihood crater around the centroid of each. Each cluster is given weight according to recency and number of points. We test single point vs. multiple points by using the previous crimes to predict the most recent one and comparing with its actual location. We extract seven datasets from published research. We use four of the datasets in developing our model and examining its response to changes in sequence, geographic concentration, and total number of points. Then we evaluate our models by running blind on the remaining three datasets. The results show a clear superiority for multiple anchor points.

This is the most important section of the paper. Make it beautiful. The above was the summary page of a winning entry by Damle, West, and Benzal at CU-Boulder in 2010.

I. INTRODUCTION

II. BACKGROUND

Either in the introduction or in the background section should be your first figure where you bring the problem to life. Nobody wants to read three pages about leaves. Just show a picture of a leaf, be sure to talk about anything you plot, and say Fig. 1 shows a cross-section of pro vascular strands of tertiary and quaternary veins in a developing leaf, and you make the reader's day.

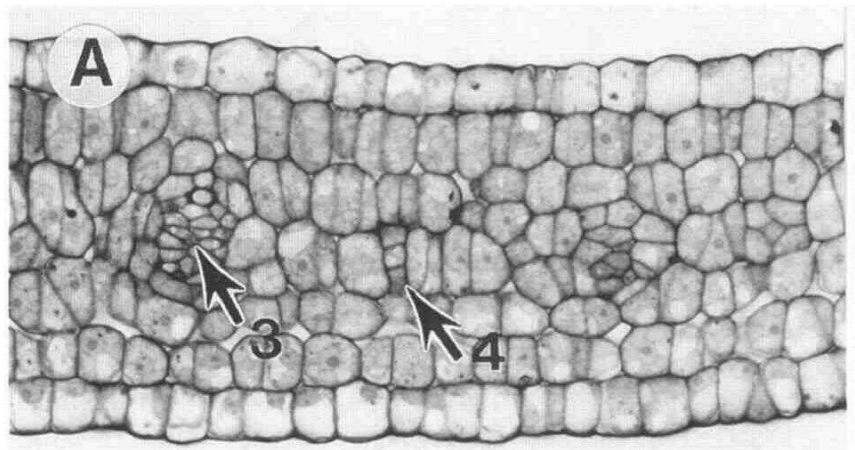


FIG. 1. Cross-section of pro vascular strands of tertiary (3) and quaternary (4) veins in a developing leaf. From Ref. [1]. For citations, use Google Scholar's cite function.

Notice that we have labeled the figure, so that later on we can always refer to Fig. 1 (see the latex). Do not write “Fig. 1” in the main document. (Quotation marks need two of these, ‘ and two of these ’). I promise you that you’ll add a new figure with two hours left to go, and now all your numbering will be a disaster.

Cite, cite, cite your references. Winning papers have 10-15 references at least.

III. ASSUMPTIONS

Bullet points can improve readability dramatically. To summarize the above, our assumptions are

- The earth is spherical.
- The earth's rotation is not important on the time scale relevant to the dynamics.
- The human population on the earth is uniformly distributed.

Or, perhaps you like to number your assumptions to be able to refer to them, later,

1. One,

2. Two,
3. Three.

The paper is organized as follows. The equations of motion are presented and written in terms of a helical coordinate system in §IV. The numerical method used to study the dynamics of an arbitrary (e.g. large amplitude) body shape is described in §V, which provides a basis for comparison for the asymptotic results. In §VI, the predictions of the mathematical model are compared to the results of our numerical method, where we find that things are great. Sometimes. An improved model is described in §VII. We conclude with a discussion in §VIII, where we discuss strengths and weaknesses of the model, and suggest future directions and stuff.

IV. THE MATHEMATICAL MODEL

About now, you'll be wanting to write some equations, either in line, $U(\mathbf{x}, t) = (\sqrt{4\pi}/M)\mathbf{x}e^t$ (note the use of the latex shortcut defined in the preamble for bold lettering). But for clarity, sometimes you will want to write something bigger, or something small that you want to stand out. After some algebraic manipulation of the expression above, we find

$$\mathcal{F}(x; \phi) = \int_0^{x^2} \frac{\xi}{\xi + \sin(\phi)} d\xi. \quad (1)$$

Note that we have labeled the equation, so that in our prose we can once again say things many pages later like “recall the expression for $\mathcal{F}(x)$ from Eq. (1),” so that later if we add a new equation we don't need to go and renumber everything. Another thing to do, is refer to whole sections in this manner. According to the assumptions described in §III, then something. Oh, that brings up another important point.

V. NUMERICAL METHOD

Describe exactly what you're going to compute, and how.

VI. RESULTS AND ANALYSIS

Results will come in the form of graphs and discussion, and sometimes tables. Table VI shows a nice template. Again, reference the label, not the number.

VII. SOMETHING MORE

A new and improved model? Measuring the robustness of the model? Something?

VIII. DISCUSSION

Summarize the above. Put a bow around your wonderful story. You should have had a beginning and middle, and here is the end.

M	U	L	N_A/N_W	U	L	N_W	U	L
16	0.081733	19.60312	12	0.078402	19.59587	6	0.082795	19.59965
32	0.081090	19.57638	24	0.081149	19.61341	12	0.084734	19.58195
64	0.081018	19.57598	48	0.080930	19.60641	24	0.085422	19.57128
128	0.081019	19.57599	96	0.080938	19.60708	48	0.085432	19.56992

TABLE I. Three convergence studies: (1) varying M , with $N_W = 5$ and $N_A/N_W = 16$ fixed; (2) varying N_A , with $N_W = 5$ and $M = 16$ fixed; (3) varying N_W , with $N_A/N_W = 16$ and $M = 16$ fixed.

A. Strengths of the model

Maybe some bullet points.

B. Weaknesses of the model

C. Future directions

And conclude with a final summary.

Appendix A: Details for the derivation of $\mathcal{F}(x; \phi)$ in Eq. (1)

A slightly messy calculation that was still small enough to include, but long enough not to want to include up top.

[1] Nelson, Timothy, and Dengler, Nancy. “Leaf vascular pattern formation.” The Plant Cell 9.7 (1997): 1121.