Crime PDE

Sean Harnett, Michael Jenkinson

December 17, 2010

Abstract

The evolution in space and time of crime risk and criminal density is modeled by a coupled system of partial differential equations, and solved with a semi-implicit method using finite differences. Under appropriate conditions, isolated stationary hot spots of high criminal activity form.

1 Problem

We intended to model the movement of criminals and the correlated evolution of crime in space and time on an urban grid by examining the model proposed by Short et al. in *A Statistical Model of Criminal Behavior*. The paper establishes a system of two coupled reaction-diffusion partial differential equations which take into account a number of characterstics of urban residential crime following a study of available criminal statistics and psychology. Specifically, and as per the aforementioned model, we sought to model the progression of stationary burglary crime in an urban area.

Such a model fails to take into account a number of characteristics of urban crime such as other types of infractions (both location-specific and mobile) or police response; however, stationary crimes on an organized rectangular grid are much easier to conceptualize and track, and such a study could potentially give a preliminary understanding of the formation of crime hot spots, areas subject to a higher density of repeated crime than their surroundings, and which characteristics of an urban area might have the most pronoucned effect on them. The results, given further study, could be useful in providing the police with more information to use in deciding when and where to allocate finite resources when responding to regions of varied criminal topography.

We chose to implement our own solvers and discretize the system through our own choices, though we took some cues from the authors' strategy in eventually formulating an implicit method. Our attempts and their results are summarized below.

2 Mathematical Model

In *A Statistical Model of Criminal Behavior*, Short et al. derive a model to predict the behavior of individual criminals on a discrete grid of residences and then derive an equivalent continuous model in the limit of a larger domain where houses are irresolvably close together. We focus on discretizing and solving their continuous model, summarized below, which consists of two coupled partial differential equations.

The derivation in the paper relies on a number of assumptions:

- The domain is a square grid, and the only crimes modelled are stationary burglaries.
- Criminals return home after committing a successful burglary, while other criminals appear on the grid simultaneously.
- The continuous model applies to a large domain where houses are irresolvably close together.
- Criminals tend to return to houses previously burglarized or familiar neighborhoods due to previously established knowledge of the area.
- Some areas may be more attractive than others due to image or physical appearance, which themselves may be factors of previous crime.
- Choice of burglary location behaves probabilistically but is a function of attractiveness and criminal density.

We also chose the boundary conditions to be periodic, though others might also be applied.

We let $A(x, y, t)$ be the measure of the attractiveness to burglary for a location on the grid, which in essence serves as a measure of the success of previous burglaries at that location. We have

$$A(x, y, t) = A_0(x, y) + B(x, y, t)$$

where $A_0(x, y)$ represents a time-independent measure of attractiveness and $B(x, y, t)$ represents the component of attractiveness which varies based on the dynamics of the system of equations explained below. We also let $\rho(x, y, t)$ be the measure of criminal density for a location on the grid.

The resulting system of equations is as follows.

$$\frac{\partial B}{\partial t} = \frac{\eta D}{z} \nabla^2 B - \omega B + \epsilon D \rho A \tag{1}$$

$$\frac{\partial \rho}{\partial t} = \frac{D}{z} \vec{\nabla} \cdot [\vec{\nabla} \rho - 2\rho \vec{\nabla} \ln A] - \rho A + \gamma \tag{2}$$

In equation 1, η is a measure between zero and one of the likelihood of crime attractiveness spreading to neighboring areas, designed to represent the negative effect of a crime-heavy area's image on surrounding areas. ω is a measure of time by which a location's attractiveness to crime will decay and is meant to represent the assumption that a crime's impact on local attractiveness will diminish as time passes. D is related to the diffusivity of criminals and is essentially a ratio of the area of the region in question to the time it takes for a criminal to travel a certain distance. ϵ is an arbitrary weighting term on the increase of attractiveness as a direct function of our two uknown variables. Finally, z is the number of adjacent neighbors to a location, which for a square grid we will assume to be 4.

From these features, we can see that the equation for the time evolution of the (time-dependent component of) attractiveness $B(x, y, t)$ consists of three terms:

- A diffusive term $\frac{\eta D}{z}\nabla^2 B$ weighted by the diffusivity of criminals and the extent to which attractiveness will spread to other areas. This factors in the spread of attractiveness to neighboring locations and attempts to model the likelihood that the increase in attractiveness at a location due to a successful crime may disperse to neighboring areas, for example due to negative effects on the region's image.

- A decay term $-\omega B$ weighted by the time by which a crime's positive effect on attractiveness diminishes. This factors in the gradual abatement of the effects of previous crimes on attractiveness. By this assumption, one would not expect the success of a crime in the distant past to encourage new burglaries as much as the success of one which occurred recently. Since law enforcement is not factored into this model, this ignores the consideration of a period of heightened police activity following a burglary.

- A source term $\epsilon D\rho A$ weighted by one arbitrary value, the diffusivity of criminals, and the density of criminals at a location. This factors in the increase of attractiveness due to current criminal density in conjunction with the success of previous burglaries given by the current attractiveness, and serves to model the manner by which successful burglaries at a location directly encourage further ones. In essence, this is the positive effect of successful burglaries on attractiveness.

In equation 2, all parameters remain the same as in the first equation with the addition of γ, which represents the generation rate of criminals per area at a location. We can expand the equation as follows:

$$\frac{\partial \rho}{\partial t} = \frac{D}{z}\nabla^2\rho - 2\vec{\nabla}\cdot[\rho\vec{\nabla}(\ln A)] - \rho A + \gamma \tag{3}$$

This evolution equation in turn consists of the following four terms:

- A diffusive term $\frac{D}{z}\nabla^2\rho$ weighted by the diffusivity of the criminals. This factors in the likelihood that criminals will pass to neighboring locations without committing a burglary at the location in question, and represents the random motion of criminals.

- An advective term $-2\vec{\nabla}\cdot[\rho\vec{\nabla}(\ln A)]$. In analogy to the archetypal advection example, the negative sign is standard on this side of the equality, ρ is the variable being advected, and the gradient of the natural logarithm of A serves as the velocity field. As such, this term produces an effect where the criminal density is advected up the steepest gradient in $\ln A$. Because of the nature of the logarithm, the steepest gradient will be less pronounced for larger values of A, resulting in slower movement to adjacent locations. This suggests that higher attractiveness at one cell relative to others slows movement to the other cells, as is to be expected. As such, here we are modelling the effect of the attractiveness on the movement of criminals from one location to another, given that they don't commit a burglary at that location in question.

- A decay term $-\rho A$ which reduces the criminal density in a manner directly proportional to attractiveness. This implies that greater attractiveness increases the likelihood that a criminal will commit a crime at a location rather than moving to another, and is based on the assumption that criminals leave the grid for a period of time after committing a successful robbery.

- A constant (though possibly spatially varying) source term γ. This represents the rate of introduction of new criminals to the grid, and is based on the assumption that while some criminals are leaving the grid due to a successful robbery, others will return to it at a later time.

3 Numerical Methods

We used MATLAB to implement all of the following methods.

3.1 Forward-Time, Centered-Space Scheme

We first attempted a forward-time, centered-space finite difference method including the use of the five-point Laplacian:

$$\frac{B_{i,j}^{n+1} - B_{i,j}^{n}}{k} = \frac{\eta D}{z}\nabla_5^2 B - \omega B_{i,j}^{n} + \epsilon D \rho_{i,j}^{n} A_{i,j}^{n} \tag{4}$$

$$\frac{\rho_{i,j}^{n+1} - \rho_{i,j}^{n}}{k} = \frac{D}{z}\nabla_5^2 \rho - \rho_{i,j}^{n} A_{i,j}^{n} + \gamma_{i,j}^{n} - \frac{D}{z}\left(\vec{\nabla}\rho \cdot \vec{\nabla} \mathrm{ln} A\right) - \frac{D}{z}\rho_{i,j}^{n}\nabla_5^2 \mathrm{ln} A \tag{5}$$

where for some vector v, the five-point Laplacian is given by

$$\nabla_5^2 v = \frac{1}{h^2}\left(v_{i-1,j}^{n} + v_{i+1,j}^{n} - 4v_{i,j}^{n} + v_{i,j-1}^{n} + v_{i,j+1}^{n}\right) \tag{6}$$

and the gradient is given by

$$\vec{\nabla} v = \frac{1}{2h}\left(\left[v_{i-1,j}^{n} + v_{i+1,j}^{n}\right]\vec{i} + \left[v_{i,j-1}^{n} + v_{i,j+1}^{n}\right]\vec{j}\right) \tag{7}$$

according to centered-space finite differencing with time step k and spatial step h. We then stepped explicitly in time, calculating both B^{n+1} and ρ^{n+1} from B^n and ρ^n using equations 4 and 5.

This method was our first quick attempt at a functional solver before more sophisticated attempts, and we quickly abandoned it for the next method.

3.2 Fully-Coupled Method of Lines

Our next scheme utilized the adaptive Runge-Kutte solver ODE45 in Matlab after using centered-spatial finite differencing as described in the previous section. That is, we solved the following system of ordinary differential equations using ODE45 in order to utilize the solver's built in adaptive time-stepping:

$$\frac{\partial B}{\partial t} = \frac{\eta D}{z}\nabla_5^2 B - \omega B_{i,j}^n + \epsilon D \rho_{i,j}^n A_{i,j}^n \equiv f\ (B, \rho) \tag{8}$$

$$\frac{\partial \rho}{\partial t} = \frac{D}{z}\nabla_5^2 \rho - \rho_{i,j}^n A_{i,j}^n + \gamma_{i,j}^n - \frac{D}{z}\left(\vec{\nabla}\rho \cdot \vec{\nabla}\ln A\right) - \frac{D}{z}\rho_{i,j}^n \nabla_5^2 \ln A \equiv g(B, \rho) \tag{9}$$

with ∇_5^2 and the centered-space gradient $\vec{\nabla}$ defined as above in equations 6 and 7.

The built-in adaptive time-stepping proved very helpful, as the FTCS scheme required careful management of the step size to prevent instability.

3.3 Finite Difference in Space, Semi-Implicit Scheme

After observing the tendency for the FTCS scheme to go unstable and the tiny timesteps taken by the ODE45 method of lines, we tried an implicit method. We used an implicit-explicit method as described in 11.5 of LeVeque's *Finite Difference Methods for Ordinary and Partial Differential Equations.* This general idea was also used by Short et al., though theirs was more complicated.

Essentially, we made the linear parts of the equation implicit and allowed the non-linear terms to remain explicit. The diffusion operator is stiff and was likely causing the aforementioned stability issues; however, it's linearity allowed us to devise such a method. We hoped this scheme would allow us to take much larger time steps, and consider much finer meshes. Our method was as follows:

$$\left([1 + \omega k]I - \frac{\eta D}{z} k \nabla_5^2\right) B^{n+1} = B^n + \epsilon D k \rho^n A^n \tag{10}$$

$$\left(I - \frac{D}{z} k \nabla_5^2\right) \rho^{n+1} = \left(I - k\left[\frac{2D}{z}\nabla_5^2 \ln A + A\right]\right)\ \rho^n - \frac{2D}{z} k\ \left(\vec{\nabla}\rho \cdot \vec{\nabla}\ln A^n\right) + k\gamma \tag{11}$$

where again ∇_5^2 and $\vec{\nabla}$ are as defined above in equations 6 and 7.

For the linear solves, we computed the constant Cholesky factor once and reused it to solve at each time step. This proved to be around fifteen times faster than simply using the backslash operator. The matrices are very sparse and positive definite, which suggests that the conjugate gradient method would be a good choice. We found this to be only slightly faster than plain backslash, however.

4 Results

We solved Eqs. 10-11 on a 512x512 numerical grid with constant initial conditions, with the exception of a few grid points with a slightly higher B value. Specifically, we used $A^0 = 1/30, \kappa = 2.5, \omega = 1/30, D = 100, \gamma = .02, B_0 = \kappa\gamma/\omega, p_0 = \gamma/(A^0 + B_0)$. We used $\eta = .01, .04$ to get the results seen in Figs. 1-2. For a sweet animation, see this video (link to 70MB avi file).

For η large, the attractiveness of a particular site diffuses too quickly to form into a hot spot. For η small, hot spots begin to emerge. Smaller values of η lead to more tightly packed hot spots, and for

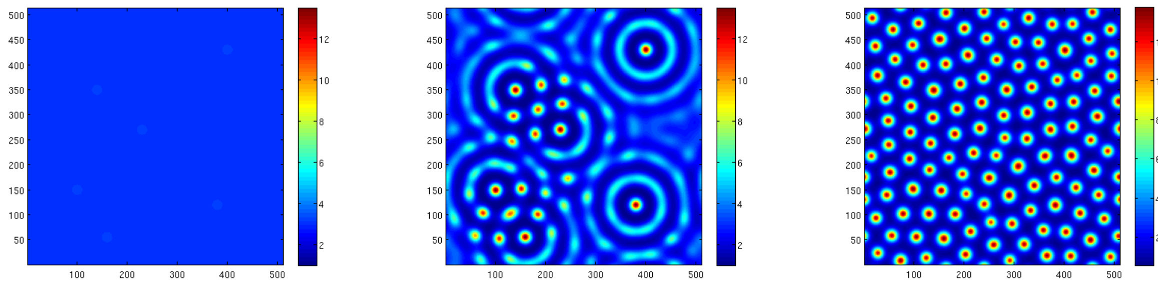

Figure 1: Formation of hot spots for $\eta = .01$ at times t=225, 425, 1600 days.

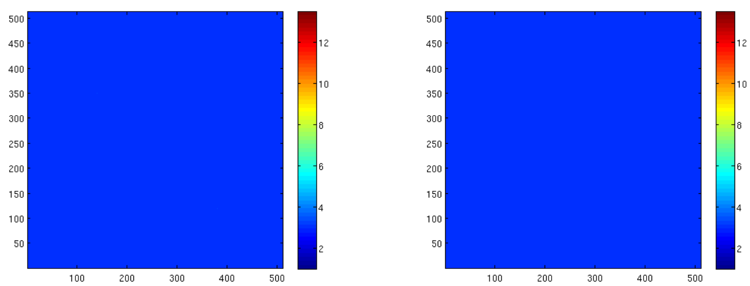

Figure 2: No hot spots for $\eta = .04$, $t = 0, 400$ days. At full resolution you can see the tiny specks for the initial condition at $t = 0$.

η too small, stability becomes an issue. We were unable to produce the larger hot spots as seen in the Short paper. This was due to the very long run times needed to produce the figures; with more time to tinker we think we could have achieved this.

Even with the semi-implicit method, we found the problem to be sensitive to stability issues. Smaller meshes seemed to be more vulnerable to this problem; hot spots would form, but would continue to grow without bound. Using a larger grid mitigated this somewhat, but at much greater computational cost. This made testing quite difficult.

We spent a lot of time testing out different parameters. Short et al. produced a scaled version of Eqs. 1-2 that eliminated many of the parameters, but they reported their results with the full list. Since we were trying to duplicate these results, it made sense to use the unscaled version of the equations.

5 Discussion

We left a number of ideas unexplored. Here are the main ones:

- higher order implicit-explicit methods, as described in Leveque
- a better way of computing the $\frac{D}{z}\vec{\nabla}\cdot[\vec{\nabla}\rho - 2\rho\vec{\nabla}\ln A]$ term (as described in office hours and MMO)
- multigrid as linear solver
- exponential time differencing methods, perhaps in combination with
- spectral methods

If we were to continue the project, the next thing we'd try would be some sort of spectral method. The simplest thing would be to Fourier transform Eqs. 10-11 and solve them in the frequency domain, then transform back. The periodic boundary conditions should make that pretty straightforward. We would also try computing all the space derivatives on the right hand side of the equations with the pseudospectral trick done in lecture, as opposed to finite differences. This would hopefully give us a lot more spatial accuracy at each time step.

The extra accuracy in space could be combined with a better time-stepping method. What we used is essentially a mix of forward and backward Euler; the two-step implicit-explicit method in Leveque (trapezoidal plus Adams-Bashforth) could perhaps give us something higher than first order in time.

When we used plain backslash, by far the most expensive part of each time step was the linear solves. It was frustrating trying to find a good preconditioner to get PCG to substantially beat sparse direct. Ultimately we abandoned conjugate gradient altogether, as simply reusing the complete Cholesky factor in a direct solve was faster. Suspecting that MATLAB's built in PCG algorithm is broken, we briefly looked for a way to interface with an external solver, to no avail. Based on the lecture and office hours, it seems that multigrid might be the best choice.

6 References

1. Short, M.B., D'Orsogna, M.R., Pasour, V.B., Tita, G.E., Brantingham, P.J., Bertozzi, A.L., & Chayes, L.B. (2008). A statistical model of criminal behavior. *Mathematical Models and Methods in Applied Sciences*, 18(Suppl.), 1249–67.

2. LeVeque, R.J. (2007). *Finite difference methods for ordinary and partial differential equations.* Philadelphia: Society for Industrial and Applied Mathematics.