siam news

Predictive Policing

By Dana Mackenzie

Philip K. Dick got it wrong.

In 1956, in a story called "Minority Report" (later made into a movie), the science fiction writer envisioned a world in which police could identify crimes before they happened. In the story, a panel of three "precogs" used their powers of extrasensory perception to foresee these future crimes. The police would then prevent the crimes by arresting the perpetrators-to-be.

More than half a century later, Dick's vision of predictive policing is coming true, at least in part. This summer, police in Santa Cruz, California, began using computer models to identify "hotspots" where crimes are more likely to be committed on a particular day.

Dick got the story wrong in two important respects, however. First, the new predictive policing is based on places, not people. "The models say nothing about which individuals are committing the crimes," says Jeffrey Brantingham, an anthropologist at the University of California at Los Angeles. Thus, the issues of free will and civil liberty that permeated "Minority Report" are not really relevant.

Second, the 21st-century version of predictive policing does not rely on ESP. In reality, no clairvoyance is required—only mathematical modeling.

Crime Begets Crime

Around five years ago, Brantingham began talking with UCLA mathematicians Andrea Bertozzi and Lincoln Chayes about the possibility of using mathematical models to understand criminal behavior. The conversations led to a National Science Foundation grant and a burgeoning research program that has grown to include undergraduates, graduate students, and postdocs, working in close cooperation with the Los Angeles Police Department.

The fundamental idea behind predictive policing is that crimes beget other crimes. For example, when a burglary is committed at a certain location, the likelihood of another burglary at the same location in the next few days is greatly enhanced. Criminals really do return to the scene of the crime. The probability of burglaries near the original crime scene also increases, as either the original burglars or copycats seek other targets.

At the same time, many burglaries are not related to previous burglaries. They are part of a background crime rate that

In addition to the accompanying article, Dana Mackenzie has written recently for SIAM News *about isogeometric analysis and the mathematics of imaging the core–mantle boundary region deep within the earth. For such articles, combined with extensive writing for more general audiences, he received the 2012 Communications Award of the Joint Policy Board for Mathematics. Commending him for "a remarkably broad and deep body of writing for experts and nonexperts alike," the award committee pointed out that Mackenzie's "work focuses largely on mathematics itself, but also touches geology, climate change, astronomy, academic mathematics as a profession, and even the game of chess." He is the author of the forthcoming book* The Universe in Zero Words: The Story of Mathematics as Told Through Equations *(Princeton University Press, June 2012).*

is stationary in time but variable in place. This assumption is very natural; all of us

Policing

know of "bad neighborhoods." Curiously, though, previous efforts to computerize the collection of crime data tended to underestimate the importance of the background crime rate and overemphasize the recent trends.

Several big-city police departments currently use a method called CompStat (short for Computer Statistics or Comparative Statistics): All recent crime complaints are plotted on a map and compared with statistics for previous years; "hotspots" of unusual activity are identified. "I've been to the CompStat meetings," says Bertozzi. "The police get together in a gigantic room in downtown Los Angeles, and go over the data division by division on a map and discuss what they will do in the next few weeks to respond to it."

CompStat has been credited with significantly reducing crime rates in New York, where it was first introduced; it brought hard data to an enterprise that had previously depended on anecdotes and intuition. Nevertheless, CompStat has some limitations. When the meetings are held weekly, as in Los Angeles, the police cannot respond to trends on shorter time scales. It is also a purely reactive approach. Police are sent to patrol places where crimes have already occurred. There is no attempt to construct mathematical models of the spread of crime to nearby locations.

The First Models

Bertozzi, Chayes, Brantingham, and Martin Short (at that time a postdoc) started with a probabilistic, agent-based model in which individual burglars move randomly on a grid and choose targets based on their perceived attractiveness. Computer simulations showed the emergence of the hotspot effect.

In a subsequent paper, Bertozzi and Short studied a simpler, deterministic model, with functions representing the density of criminal agents and the risk of crime at any given place or time. These functions satisfy a system of two reaction–diffusion equations. The researchers proved, for instance, that hotspots form when randomly occurring background crimes are close enough that the "areas of influence" of individual crimes overlap.

At this stage, the mathematical models were highly theoretical. The agent-based model would require information about how individual criminals move, which could be obtained only from tracking devices placed on every criminal or potential criminal. The deterministic model involves parameters like the diffusion rate or the distance scale at which individual crimes start forming clusters, whose values are unknown. "The model I've just described is too myopic," Bertozzi says. "We can tell you everything about an idealized society, but the real world is not that way."

Crimes and Earthquakes

The breakthrough toward a model with real-world applicability came when another postdoc, George Mohler, joined the group. Mohler had heard from a colleague, statistician Rick Schoenberg, about a method that seismologists use to model earthquakes, called "stochastic declustering."

Earthquakes, it turns out, have a lot in common with crimes. Every region has a certain unchanging background risk of earthquakes. In addition, there is a time-dependent risk of aftershocks, particularly after a major quake. Stochastic declustering is a statistical method for parsing earthquake data into "background" events

> *The breakthrough toward a model with real-world applicability came in the form of a method seismologists use to model earthquakes, called "stochastic declustering."*

and events that are aftershocks (or precursors) of other quakes. Obviously, earthquakes that are closer in space and time are more likely to be related. The same probability distributions, or "kernels," that are used retrospectively to determine the relatedness of past earthquakes can also be used prospectively to estimate the likelihood of future ones.

Mohler applied the same type of analysis to one year of burglary data from the Los Angeles San Fernando Valley. He flagged neighborhoods in which the model indicated that burglaries were more likely to occur. Of course, if you flag 100% of the neighborhoods in the city, you will correctly "predict" every crime that occurs, but you will not generate any useful information. Mohler found that if he flagged the 10% of the city at highest risk each day, he correctly predicted the location of 660, or 25%, of the 2627 burglaries in the data set.

At this point, Zach Friend, a crime analyst for the Santa Cruz Police Department, read an article in *The Los Angeles Times* about Mohler's work and contacted him to see if he would like to try his model in action. Mohler spent a few months developing the software and working with the police to decide what the output should look like. The system went online in July 2011.

Every day at 4:00 PM, Friend collects the crime reports for that day and feeds them into the software. The computer works out the probability of a burglary in each 500-foot by 500-foot grid cell in Santa Cruz over the next day. Friend identifies the ten grid cells with the highest probability and tells police officers to pay special attention to those areas. This may mean just driving their cruisers through the areas during the times when they are not doing anything else. According to

Policing

continued from page 2

Friend, about 20% of the blocks flagged by the computer are places that the police would not have anticipated as high-risk.

Two months into the project, the data look very encouraging. Six arrests have been made as a result of the computer "tips," including one that was prominently featured in a *New York Times* article. In a downtown garage that had been flagged as high-risk, police found and arrested two women who were peering into cars. According to the *Times*, one of the women had an arrest warrant outstanding, and the other was carrying illegal drugs. Perhaps more importantly, burglaries were down 27% in Santa Cruz in July 2011 compared with July 2010. This reversed the trend of the first half of 2011, when the burglary rate had been higher than the previous year's.

In a department hit by staffing reductions, Friend says, predictive policing "allows us to be more effective with our resources." The reaction within and outside the department has been very positive. In November, the Santa Cruz predictive policing experiment made *Time* magazine's list of the top 50 inventions of the year.

Unfortunately, as Martin Short points out, the Santa Cruz experiment is not "scientifically rigorous," because it does not compare the results of predictive policing to those of business as usual. (Perhaps burglaries would have dropped by 27% anyway.) The limitation was at the request of Santa Cruz, which in any case was arguably not large enough to conduct a full-blown randomized controlled trial. Los Angeles definitely is large enough, and that will be the next step for Mohler (now at Santa Clara University) and the UCLA team. The Los Angeles Police Department started a randomized controlled trial of Mohler's predictive policing software in October. Mohler expects to release results in May.

Gang Violence

So far, the researchers have emphasized burglaries over violent crimes like murder. According to Friend, there is a reason for this: "Criminologists have found that [property crime] is a predictable act and that you can deter it simply by having a police presence in the area." Violent crime is harder to predict and to deter.

Gang-related violence might be an exception to this rule. Undergraduate research groups at UCLA have studied mathematical models of gang-related crimes in the Hollenbeck area of Los Angeles, which has about 30 active street gangs. Gang-related crimes are highly clustered, and the computer can often identify a gang as the likely perpetrator. From a mathematical point of view, "These are well-behaved gangs!" says Kym Louie, a senior at Harvey Mudd College who has worked on the project for three summers.

Bertozzi and George Tita, a criminologist at the University of California at Irvine, have worked on agent-based models of gang rivalries, or what Bertozzi calls "bottom-up models." Such models have proven useful for illustrating how a gang-rivalry network might form. While they still cannot make predictions about individual people, the researchers hope to incorporate additional individual-level data from the LAPD to make their models of gang activities and retaliatory behavior even more realistic.

According to Brantingham, one key to the success of the UCLA research program has been the supportive attitude of the Los Angeles Police Department, especially in the early years, when the research was very theoretical. "We've been very careful not to over-promise," he says. "We've never pushed the idea of mathematics as a silver bullet. Good science has to be done in small, incremental steps. Now, after about a half dozen years, we are moving in a more practical direction, and we feel cautiously optimistic."

Dana Mackenzie writes from Santa Cruz, California.