COMPUTING / SOFTWARE

FEATURE

Beyond C.S.I.: The Rise of Computational Forensics

Pattern recognition and other computational methods can reduce the bias inherent in traditional criminal forensics

By SARGUR N. SRIHARI / DECEMBER 2010

Image: EFE/El Pais/AP Photo

TERROR ERROR: Closed-circuit television captured this image just seconds after three bombs exploded at Madrid's Atocha train station on 11 March 2004. The FBI claimed that a fingerprint at one of the bomb sites was that of U.S. citizen Brandon Mayfield but later retracted its accusation.

On 6 May 2004, a Portland, Oregon, lawyer named Brandon Mayfield was arrested for his alleged involvement in the terrorist bombings of four commuter trains in Madrid. The attacks killed 191 people and injured 2000 others. But Mayfield had never been to Spain, and his passport at the time was expired. The sole evidence against him was a partial fingerprint found on a plastic bag in a van used by the bombers. The FBI's Integrated Automated Fingerprint Identification System had identified Mayfield as a possible match, and three FBI fingerprint experts as well as an outside analyst confirmed the identification.

The analysts knew that Mayfield had converted to Islam, was married to an Egyptian woman, and had once represented a man in a child custody case who later turned out to be part of a jihadist group. That information swayed the FBI inquiry in Mayfield's direction.

Spanish authorities, however, argued that the fingerprint belonged not to Mayfield but to an Algerian with a criminal record, Spanish residency, and terrorist links. They were right. It took almost three weeks from his arrest, but Mayfield was cleared of the charges and released from federal custody. The U.S. government eventually agreed to pay him US $2 million for the mistake and issued a formal apology.

Such high-profile cases grab the headlines and our attention, but they also point to an underlying problem with fingerprints—and with shoe prints, handwriting, and nearly every other form of classical forensics data. "The fact is that many forensic tests…have never been exposed to stringent scientific scrutiny," a committee convened by the U.S. National Academy of Sciences concluded last year. One of the main problems with forensics evidence is that it must be analyzed and interpreted by a person, whose own theory of the crime can introduce a bias in the results. There can also be significant uncertainty in the analyst's conclusions, but oftentimes that uncertainty is never quantified or conveyed to judges and juries.

And yet, these traditional forms of forensics evidence can be very helpful, provided they can be looked at objectively and the uncertainty of the results can be measured and properly explained. The relatively new field of computational forensics has sprung up to address those needs.

Computational forensics is not yet mainstream. But lots of research goes on in academic settings, such as at my own lab at the State University of New York at Buffalo, and eventually the courts may allow these techniques to be applied in criminal trials. Clearly, any method that can improve the analysis of evidence would be a good thing.

On the popular television series "C.S.I." and its spin-offs, attractive forensics experts speedily uncover long trails of evidence that point unfailingly to the true villains. In the real world, crime scene investigations are rarely so cut and dried (and the people carrying them out are rarely so glamorous). More often, forensics experts must rely on a few murky clues, none of which is definitive in itself. Even then, as the Brandon Mayfield case revealed, the biases of the analysts can lead to erroneous conclusions.

To appreciate how computational methods could avoid such blunders, it's helpful to understand how forensics data are used. In general, a forensics investigation attempts to match crime scene evidence with a known source. The first step is an analysis of the evidence, followed by a comparison of the evidence with data from known sources, and finally an independent verification of the results.

The type of evidence determines how exactly this process unfolds. An incredibly wide variety of classical forensics data exists: impression patterns (such as latent fingerprints, shoe prints, and tire treads), nonimpression patterns (for example, bloodstains and speech), trace evidence (including paint, dust, and pollen), biological evidence (such as DNA and hair), and toxicological evidence (including drugs or alcohol in the victim's bloodstream).

Then there's the whole burgeoning field of digital forensics, which involves the use of data extracted from personal computers, cellphones, digital cameras, and the like. [For more information on these topics, see *IEEE Spectrum*, "Cellphone Crime Solvers," July 2010, and "Seeing Is Not Believing," August 2009.]

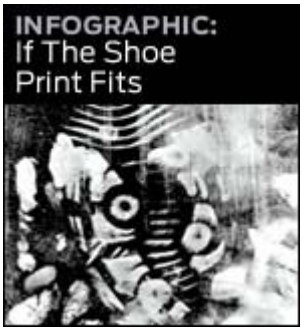
During two years of deliberation by the National Academy's forensic science committee (of which I was a member), a troubling picture emerged. A large part of current forensics practice is skill and art rather than science, and the influences present in a typical law-enforcement setting are not conducive to doing the best science. Also, many of the methods have never been scientifically validated. And the wide variation in forensic data often makes interpretation exceedingly difficult.

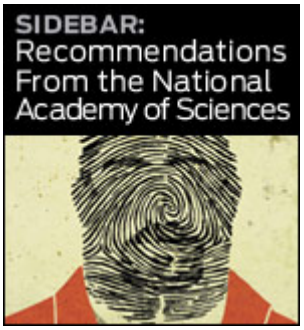
Among all the classical forensics methods, the committee concluded, only DNA analysis has been shown to be scientifically rigorous [see sidebar, "The DNA Difference"]. But committee members identified several areas where the greater use of computers and automation could eliminate bias and minimize error.

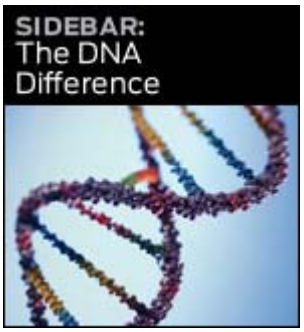
Images: US Department of Justice **MISPRINT:** This latent fingerprint [above] was found on a bag of detonators connected with the 2004 Madrid bombings. Markings indicate where an FBI fingerprint examiner determined the print to be similar to Brandon Mayfield's [below]. Spanish officials disputed the finding and said the print was that of an Algerian national.

Photo: Nike

SEEN AT THE SCENE: Of all the footwear on the market, the Nike Air Force 1 sneaker is the most often encountered at U. S. crime scenes, turning up in about 17 percent of cases.







Some degree of automation already exists. In the Madrid bombing case, FBI analysts first examined the crime scene fingerprint by hand. Then they ran the print through the bureau's fingerprint database, which with some 400 million records from 66 million people is said to be the largest biometric database in the world. The search produced 20 matches, on which the analysts did side-by-side comparisons with the crime scene print. Eventually they deemed Mayfield's the closest match, with 15 points of similarity. (Mayfield's prints dated from 1984, when as a teenager in Kansas he had been arrested for burglary.) It's more or less like using a search engine such as Google: You rely on the search engine's algorithms to return results matching your query, but invariably you have to go through the list to find what you're looking for. The main difference, of course, is that in a criminal investigation the stakes are much higher.

So how might greater automation of classical forensics techniques help? New algorithms and software could improve things in a number of ways. One important area is to quantify the chance that the evidence is unique by applying various probability models.

Let's look at the uniqueness problem as it applies to fingerprints. As children, we learn that no two people have the same fingerprints. That may indeed be true, but in a sufficiently large pool of people, the likelihood of two of them having very similar fingerprints is quite high. Beginning in the 1800s, scientists began looking for distinct features within the ridges of fingerprints—such as whorls, arches, single loops, and double loops—that would help distinguish one person's prints from another's. However, such broad features occur with surprising frequency. Left loops, for example, are found in 30 percent of all fingerprints, and right loops are nearly as common. So fingerprint analysts also look at more detailed features, known as minutiae, such as the locations and directions of ridge endings and bifurcations.

Fingerprint-matching algorithms mostly search for and compare minutiae rather than the larger features, because minutiae are less common and yet are still stable and can be reliably extracted and compared by computer. So the key question in quantifying the individuality of fingerprint evidence is, what are the odds that two randomly chosen samples will have a high degree of similarity?

Using statistical modeling based on the frequencies of specific minutiae in the population, the theoretical probability that a random pair of fingerprints will exhibit a match of, say, 12 of 36 minutiae works out to 1 in 100 billion. While that sounds extraordinary, keep in mind that there is almost always uncertainty in the quality of the minutiae data. For instance, a smudged or partial fingerprint can lead a human analyst (or a computer, for that matter) to flag a similarity

where none may in fact exist. Also, more research is needed to determine the exact frequencies with which certain fingerprint features occur. At present, nobody really knows how likely it is for two people's prints to match closely or the extent to which trained fingerprint analysts can tell them apart. Indeed, experts sometimes disagree on what constitutes a match.

To address these issues, my research group and others are trying to figure out how to calculate the probability of one person's fingerprints randomly matching those of any other in a given population. That's a key thing to know. If a fingerprint found at the crime scene corresponds reasonably well to a print from the accused, the prosecution hypothesis is that they are from the same person, while the defense hypothesis is that they are from different people. With the right computational tools you can calculate two probability values, one for the prosecution hypothesis and one for the defense hypothesis. And the ratio of the former to the latter shows just how incriminating (or not) this evidence is.

Computational forensics can also be used to narrow down the range of possible matches against a database of cataloged patterns. To do that, you need a way to quantify the similarity between the query and each entry in the database. These similarity values are then used to rank the database entries and retrieve the closest ones for further comparison. Of course, the process becomes more complicated when the database contains millions or even hundreds of millions of entries. But then, computers are much better suited than people to such tedious and repetitive search tasks.

Here's how an automated database search could work with shoe prints, which are among the most common types of evidence found at crime scenes [see illustration, "If the Shoe Print Fits"]. The shoe print database may consist of photographs of the outsoles provided by shoe manufacturers or vendors. Each shoe print in the database is indexed using a feature vector or description, similar to those used in computer vision applications for representing elements in an image. For the shoe print, the feature description identifies the straight edges in the print and patterns such as ellipses, circles, or herringbones.

The query print taken from the crime scene is usually of much poorer quality. For example, the print may be nearly indistinguishable from the surface on which it was made, and so it first needs to be enhanced. One way to do this is with an image-processing technique known as intelligent thresholding, which automatically breaks up an image into its component segments and separates the foreground (the print) from the background (the surface). Once the shoe print has been enhanced, a corresponding feature description is extracted. The search algorithm then computes the similarity between the query and each database shoe print. Lastly, the database entries are ranked to reveal the closest matches.

Currently, forensics examiners who analyze shoe prints have no means of conducting such an automated search. The state of the art is a database package called SoleMate, developed by Foster & Freeman, in Evesham, England, and sold to law-enforcement agencies worldwide for about $7000. It works like this: The user renders onscreen a close approximation of the crime scene shoe print by selecting common features and locations—such as a star pattern near the big toe—from drop-down menus. SoleMate then compares the result to its database of more than 24 000 sports, work, and casual shoes, which it regularly updates using data from shoe companies. Clearly, though, having to redraw the shoe print is needlessly time-consuming. My team is currently working with Foster & Freeman on ways to automate its software.

Of course, fingerprints and shoe prints aren't the only kinds of evidence that could be evaluated automatically. For some cases, you might need to compare handwriting samples—to identify the source of a threat note, for instance.

Analyzing handwriting is less straightforward than fingerprint analysis, because a person's handwriting varies somewhat from day to day and can change dramatically over the course of years. Also, handwriting can be forged in a way that fingerprints and DNA cannot. Still, there are ways that computers can aid in handwriting comparisons.

A handwriting sample has global or macro features, such as line spacing and slant, as well as micro features like the shapes of letters and whole words. For a computer to be able to compare a handwritten paragraph, the sample must first be processed to extract individual lines of text, individual words within those lines, and individual letters and combinations of letters within those words.

From these basic components, the computer can quantify the features present in the sample, and a measure of similarity between it and that of a known sample can then be calculated. At present, there's no way to entirely automate this process, because tasks like the recognizing and isolating of individual letters are quite complex and better handled by humans than by machine.

But handwriting-recognition software continues to improve, as shown at the U.S. Postal Service, where 95 percent of all handwritten addresses can now be read by machine. In fact, forensic handwriting analysis builds directly on methods developed for recognizing postal addresses, and a commercially available forensics program called CEDAR-FOX incorporates work that my group did in this area. There are some key differences between the two applications, however. While both need to separate lines and words of text and recognize individual letter shapes, address recognition isn't really concerned with the handwriting differences between individuals, whereas in forensics it is these very differences that need to be captured. Even so, it seems only a matter of time before computers make greater inroads into handwriting forensics.

What computational forensics—or any forensics method, really—cannot do is determine whether a suspect did or did not commit the offense. That's a matter for a judge and jury to decide. At trial, the role of a forensics expert is to testify whether the profile drawn from the evidence matches that of the suspect or of an unrelated person. It's still common, though, for a lawyer to ask a forensics expert to speculate on matters further afield, such as the type of activity that might have led to the presence of the evidence at the crime scene. And some of those experts will freely render such opinions—but they shouldn't.

Will traditionally trained forensics examiners welcome the further incursion of computers into their work? I believe they will. Those I've met have been eager to learn about new technologies to improve what they do (even while I'm pointing out the shortcomings of their current practices). Likewise, those of us active in computational forensics pursue these techniques not solely for the intellectual challenge but because we firmly believe that our work will one day help to prevent the wrongful conviction and punishment of innocent people. Such injustices are tragedies, not only because those convicted suffer but also because the guilty remain at large.

This article originally appeared in print as "Computing the Scene of a Crime".

About the Author

Sargur N. Srihari, author of "Computing the Scene of a Crime", is an IEEE Fellow and a professor at the State University of New York at Buffalo. Compiling large sets of forensics data can be costly and time-consuming, but he's learned to be resourceful. One unexpected source: the footwear purveyor Zappos.com, which offers detailed photos online of each shoe it sells, including the sole. "You'd think most people wouldn't be interested in that level of detail," Srihari observes. "But it's a very nice database for researchers like me."