No place to hide

Danny Bradbury

Received 1 November 2005; revised 1 November 2005; accepted 1 November 2005

Surveillance technology is developing both at street level and at a macroscopic level — but will it do the job, and at what cost?

The Reality Mining project at MIT’s Media lab could enable others to predict your behaviour with almost perfect accuracy, thus destroying the shreds that remain of your personal privacy. Not that the researchers intend this, but the eventual use of a research concept doesn’t always mirror its original purpose.

Students at MIT’s Media lab and the adjacent Sloan Business School are using 100 special Nokia cellphones. Each carries a parasite application that logs everything from cellphone tower ID to call logs and nearby handsets to which the phone has established a Bluetooth connection. The phone uploads the data to a central database that uses an inference engine to work out trends in behaviour. Researchers believe that they can predict with remarkable accuracy the movements of most individuals. In many cases, they get it right over 90% of the time.

The ability to identify individuals and their social networks with such accuracy could be invaluable for epidemiology studies and disease tracking, where data about the movements and relationships of individuals could be used to predict and control the spread of contagious diseases, for example. This carries significant implications for privacy and surveillance.

The result could be a scenario in which everything about your communication activities could be stored, from the times that you logged on through to the data about the WiFi access point that you access the internet from. Such data could in turn lead people to know which coffee shop you’re usually in at a certain time; for instance, tying the MAC address of your laptop to the wireless LAN’s data logs. And the system would also know who else was logged on and in that coffee shop with you, at the same time.

Conspiracy manna

Of course, this scenario is manna from heaven for conspiracy theorists. The use of ‘life logs’ of mobile and other communications media and physical movements can identify the large percentage of people whose movements are predictable and those others whose lives are more erratic and difficult to predict. Given that people are creatures of habit, there is nevertheless something depressing in the fact that nine times out of 10, a database knows exactly what someone will be doing over the next few hours, or even days. Depressing, but not necessarily surprising.

But what self-respecting conspiracy nut wouldn’t suffer paranoidal paralysis knowing that, thanks to your cellphone, the authorities knew where you were going and what you were doing even before you did it?

Of course, there’s the off switch, but who wants to give up the convenience that a cellphone and laptop PC offer. Besides, it’s a huge logistical leap. Before it could happen, mobile phone companies

E-mail address: danny@itjournalist.com

1742-2876/S - see front matter © 2005 Danny Bradbury. Published by Elsevier Ltd. All rights reserved.
would have to keep all your records securely and
give them to the authorities on request. This is
gigabytes of data per person per year. Multiply
that by more than two billion cellphone users, 1.2
billion fixed line telephone users and another
billion-odd PC users.

File and forget?

Most people believe this scenario is ludicrous from
both a privacy and a technology perspective.
There’s simply too much data for it to be stored
reasonably. Besides it’s a gross infringement of
privacy laws.

But if it’s inconceivable, why is the European
Commission working on it?

The EC is currently mulling the unification of
data retention legislation. The draft directive
would make network operators and Internet service
providers keep mobile and fixed telephony data for
up to a year, and Internet communication data for
six months. A separate proposal calls for the re-
tention of data for up to three years. UK service
providers have already been retaining data volun-
tarily, says advocacy group Privacy International.

The proposals are couched in anti-terrorist
language that argues that we must be more
vigilant in our use of technology for surveillance
if we are to stop the next September 11 or July 7
attack.

Foundation shift

Gus Hosein, a senior fellow at Privacy International,
disagrees with the premise and the action. He has
been fighting government attempts to pass data
retention law for half a decade. His worry is that
by permitting the government to gather such
extensive data sets we could lay the foundation
for future legislators to use it for preventative
purposes. Instead of looking for existing suspects
they would be analysing behaviour patterns, social
networks and trends in information flow that could
lead them to potential terrorists. The foundation
of criminal justice then shifts fundamentally from
a presumption of innocence until proven guilty to
one of guilt by association until proven innocent.

"It just takes a single switch in the law that no-
one will ever notice, and then yes, we’re at that
point," he says.

Those who argue that they have nothing to hide
and therefore nothing to fear from surveillance are
missing the point, says Hosein. His fear is that the
use of such data, and in some cases the data itself,
will be erroneous.

"Incompetence has a large part to play in
surveillance. That’s why we see the wrong people
being stopped getting on airplanes," he warns.
"Incompetence might miss terrorists but might
catch me. That’s why people who think they don’t
have anything to fear actually do."

Fatal error

On the other hand, one can hope that better
surveillance technology will avoid incidents such
as the assassination of Jean Charles de Menezes,
a 27-year-old Brazilian electrician whom police
mistook for a suicide bomber, or Harry Stanley,
whom police shot dead because they believed
a chair leg he carried was a sawn-off shotgun
wrapped in plastic.

The use of large-scale databases for surveil-
ance through data mining is well-documented,
especially in the US where now-dead projects like
the CAPPS-II travel database and the Total In-
formation Awareness (TIA) initiative attempted to
accrue huge volumes of data.

Other approaches to surveillance, supported by
firms like predictive analytics software vendor
SPSS and machine translation firm Language Weaver,
aim to mine unstructured data such as Internet
pages and newsgroup postings for potential clues.
The two firms combined to produce a version of
SPSS’s text mining tool that can analyse millions
of documents in native languages. This looks for
cues by grouping together words into concepts
using linguistic analysis, says Olivier Jouve, vice-
president of data mining and text mining at SPSS.

The likely introduction of biometric ID cards in
the UK (the ID Card Bill had just passed its third
reading in the House of Commons at the time of
writing) could provide an easy way to tag aggre-
gated surveillance data to an individual, says Phil
Booth, national co-ordinator for anti-ID card lo-
bying group No2ID.

Playing tag

The National ID Registry Number (NIRN) could be
used to match data from sources like the Criminal
Records Bureau against cross-tabbed longitudinal
studies in benefits and pensions and the health
service, he warns. However, the accuracy of some
of the data in these institutions is known to be
questionable. In addition, the security of the ID
card registration process itself has yet to be
validated. This suggests the potential for wild goose
chases based purely on data mining is huge. Not
only will it be costly for the authorities to pursue
them, but the cost in money, time and emotional distress for innocent suspects will be enormous.

Like Hosein, Booth warns of the potential for inaccuracy when it comes to mining large datasets to look for anomalous behaviour. However, the equivalent alternative, scanning CCTV images to discover suspicious characters, is unlikely to be effective. Currently CCTV footage is used mostly in investigations after an attack or other crime.

But the problem remains: how do we maintain our privacy while protecting ourselves against attacks?

Neil Fisher is director of security solutions for QinetiQ, a commercial spin-off from the MoD’s Defence Evaluation and Research Agency. He proposes combining historical datasets and surveillance equipment near potential targets to help spot anomalies automatically.

“You must have an automated system. It might pick up the features of the face, compare it against a watch list or some other triggers that you may be looking for,” he says.

Companies such as Visionics offer real-time systems that match faces. In high traffic areas such as airports, they can scan and search watchlists automatically. Other new surveillance technologies include millimetre wave cameras (QinetiQ offers one), which detect natural radiation as it reflects from objects. Such technologies could detect weapons, says Fisher, but they could also help to assess unique features in an individual, such as the volume of their lung cavity, which could also be referenced against a central database for identification.

As both close-to-target and more abstract surveillance technologies develop, Hosein worries that they will continue to come together. “Previously there were these invasive devices that watched people, and 10 years ago we started talking about creating data mining systems. Now we’re seeing the fusion of the two,” he says.

Finger the usual suspects

But how effective will this marriage be against a new generation of criminal with no track record? The four July 7 bombers were described by police as ‘cleanskins’, i.e. they had no previous records. Politically motivated groups have no shortage of impressionable youths who can be turned to the cause and used once to carry out an attack. Abstract data mining might be less effective in spotting such threats.

Fisher reckons smarter near-target surveillance systems can also be used to look for anomalous behaviour close to the site and time of an attack. This should at least prompt law enforcers to pull in suspicious characters. In the pipeline is non-intrusive surveillance equipment that can measure everything from the human heartbeat to stress pheromones. This may make it possible to spot people who are nervous, for example, or who were emotionally upset because they had fought with their boss. Perhaps a smart camera might identify a car loitering in an area deemed vulnerable to attack. It could then flag it to officers who could investigate whether a potential attacker was on a reconnaissance trip, or someone was simply waiting for his wife to finish shopping.

Would the July 7 bombers have been caught by improved near-target or data mining-based surveillance systems? It is impossible to say. But perhaps the more important question is whether the expense of coming up with increasingly sophisticated and intrusive surveillance technologies is going to prove counter-productive.

In the 1980s, Ronald Reagan outsold the monolithic USSR on defence and bankrupted it. Thanks to microchips and software we now face an asymmetric threat in that the cost and effectiveness of current weapons systems and procedures are inappropriate and disproportional to the imminent threat. The post-modern Al-Qaeda-style terror gangs have shown how effective organisation on a network basis can be against the incumbent monolithic political and law enforcement institutions.

The Al-Qaeda attack on the World Trade Centre on 11 September 2001 triggered the invasion of Iraq and the ousting of its president, Saddam Hussein. As the number of US fatalities there approaches the 2752 innocents killed in the World Trade Centre, one wonders who is really paying the bill and what the final cost will be, technological aid or not, unless there is a political settlement.