PHONING IN DATA

Far from being just an accessory, mobile phones are starting to be used to collect data in an increasing number of disciplines. **Roberta Kwok** looks into their potential.

When Martin Lukac felt a small earthquake rattle his Los Angeles apartment, he immediately thought of the mobile phone lying on his desk. Two weeks earlier, he had programmed the phone to capture readings from its built-in accelerometer, a sensor originally intended to support features such as games. Now, Lukac — a doctoral student in computer science at the University of California, Los Angeles — transferred the phone's data to his computer and saw the readings plotted as a series of tell-tale spikes. Success! His phone had become a mobile seismometer.

Such moments are happening more and more often these days, as researchers seek out innovative ways to exploit mobile phones. The opportunities are tantalizing. Phones are increasingly being equipped with not only accelerometers, but also cameras, Global Positioning System (GPS) receivers and Internet connectivity. Many of them can support programs devised by anyone, not just the phone's manufacturer, which means that digitally savvy scientists can write and distribute mobile-phone software for everything from monitoring traffic to reporting invasive species.

And perhaps best of all for the budget-conscious researcher, the phones are almost ubiquitous. There are now about six mobile phone subscriptions for every ten people in the world, according to a March report[1] from the International Telecommunication Union, based in Geneva, Switzerland. And the GSM Association, a mobile-communications industry trade group, announced in February that the number of mobile-phone connections worldwide had hit 4 billion and was expected to reach 6 billion by 2013.

"We've really never had a technology other than human observation itself that is as pervasively deployed out in the world," says Deborah Estrin, Lukac's adviser and director of the Center for Embedded Networked Sensing (CENS) at the University of California, Los Angeles.

Despite the challenges in harnessing mobile phones, including privacy protection and unpredictable data flow, projects such as Lukac's are starting to emerge in a number of disciplines, from medical imaging to human behaviour.

Location, location, location

One of the most enticing features of mobile phones for researchers is GPS, which uses satellite data to pinpoint a phone's location. Once mobile phones got GPS, says Quinn Jacobson, a computer engineer at the Nokia Research Center in Palo Alto, California, they suddenly had an "'awareness' of where they were in the world".

This makes mobile phones a natural tool to study road traffic, says Alexandre Bayen, a systems engineer at the University of California, Berkeley, who is collaborating with Jacobson. Today, Bayen says, traffic is often monitored with equipment such as cameras, radar and sensors embedded in the pavement. But mobile phones could provide a cheaper way to collect the information, because scientists can piggyback on the phone companies' existing communications infrastructure. There's no need to "send a crew with a truck to dig a hole in the highway", says Bayen.

In November last year, Bayen's team launched Mobile Millennium: a project to generate real-time traffic estimates with GPS-enabled mobile
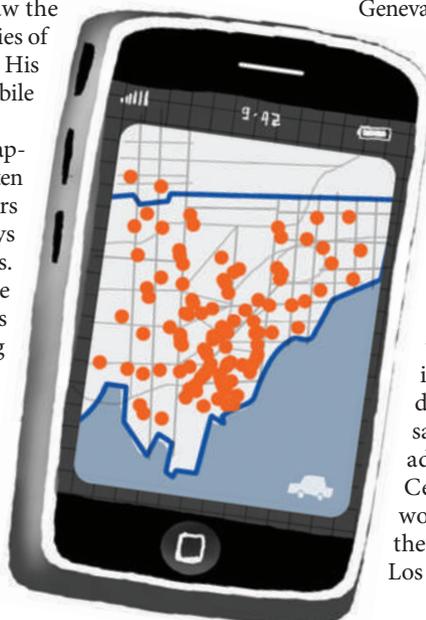
phones in a region near San Francisco. Volunteers download the software and transmit their position and velocity to a central computer system. The system combines the phone data with historical records and other sensor readings, reconstructs traffic flow in the area, and sends the results, suitably anonymized, back to the user's phone for display. So far, the software has been downloaded about 4,500 times, says Bayen, and the team hopes to recruit 10,000 users.

Because mobile phones can travel into areas that lack other sensors, they are revealing traffic flow on smaller, previously unmonitored roads outside the highway system. "For the first time, we're seeing very rich data on these types of roads," says Jacobson. The team conducted a pilot study with 100 cars driving 10–16-kilometre loops in February last year and found that the phones captured velocity patterns similar to those obtained by underground sensors, including the congestion resulting from a five-car accident[2]. When mobile-phone data are fused with other sensor data, "you can get enormous gains in accuracy", says Bayen.

The idea of using mobile-phone data to monitor traffic is not new, but Mobile Millennium has managed larger-scale deployment than most academic research efforts, says Jean-Luc Ygnace, a research engineer at the French National Institute for Transport and Safety Research in Bron. The next challenge will be to recruit enough drivers to get sufficient data over a large road network, he says.

All corners of the world

Mobile phones have managed to penetrate parts of the globe where other infrastructure is absent. In the developing world, mobile-phone subscriptions have skyrocketed from nearly zero a decade ago to 50% of today's population. Fixed telephone lines reach fewer than one-sixth of the people there[1].

Boris Rubinsky first really appreciated this reality when he visited an Indian village at the foothills of the Himalayas in 2005. Children were washing themselves in the river, and animals roamed across the roads. "Suddenly, in the middle of the street, you see a person walking around with a cellular phone," says Rubinsky, a bioengineer at the University of California, Berkeley, and director of the Center
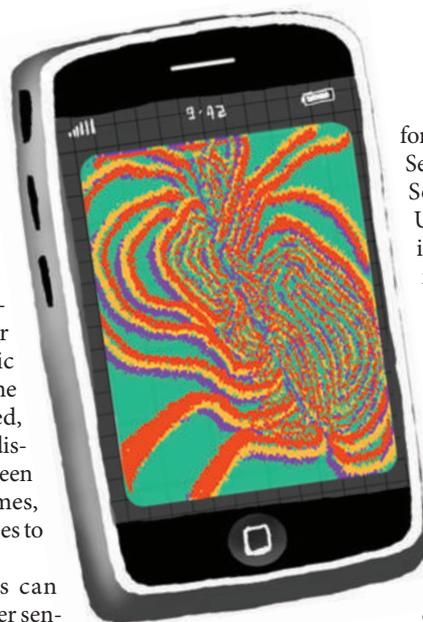
for Bioengineering in the Service of Humanity and Society at the Hebrew University of Jerusalem in Israel. "It dawned on me that cellular phones are everywhere."

Rubinsky decided to use the phones to address a specific problem in the developing world: the lack of access to medical imaging. Imaging equipment typically includes components for data capture, processing and display, which together tend to be expensive and difficult to maintain. But with a mobile phone, Rubinsky reasoned, a doctor in a remote village could transmit raw measurements from a relatively inexpensive data-collection device, such as an ultrasound transducer, to a processing centre in a major city. The centre would then reconstruct the image and transmit it back to the phone.

In a study published last year[3], Rubinsky's team did a proof-of-principle test of their system with a procedure called electrical impedance tomography, in which a device applies current to a patient's tissues and measures the resulting voltage. In this case the voltage readings were from a simulated breast tumour. But the researchers were able to move the data through every step of the process until the image appeared on the phone. "The data are extremely simple, and that's the beauty of it," says Rubinsky. He published another study[4] this month in which the system was used to classify tissue. Rubinsky is planning field trials of his system in Mexico within the next few months to see whether it can detect internal bleeding.

Other groups are creating mobile-phone applications to conduct health surveys, analyse blood samples and report natural disasters. "For those of us working in the developing world, that's the platform of choice," says Gaetano

> "It dawned on me that cellular phones are everywhere."
> — Boris Rubinsky

Borriello, a computer engineer at Google and the University of Washington in Seattle who explores how technology can improve health care in underserved populations.

Eyes on the ground

Mobile phones also have the potential to enhance the role of the citizen scientist. Information collected by non-scientists has traditionally been seen as suspect because it is difficult to verify, says Estrin. But with the help of mobile-phone cameras and GPS, which can label data with a precise location, observations taken by everyday citizens could become more reliable. "The fact that people can do real-time uploading of geo-tagged information changes that story," she says.

Estrin's team is working on a set of projects that will enlist citizens to submit field observations via mobile phone to a central database. One campaign is tracking the appearance of harmful algal blooms; another, scheduled to be opened to the public in July, will monitor invasive species in California's Santa Monica Mountains. Participants will be able to send in geo-tagged pictures, with optional text messages to describe each photo, says Estrin. By next year, the team also plans to supply a phone-based reporting system for Project BudBurst, a citizen-science effort to measure the effect of climate change on plant blooming. Meanwhile, Lukac and his colleagues are hoping to deploy a test set of accelerometer-equipped phones in an earthquake-prone area and eventually establish a mobile seismic network.

Mobile phones may "open up the demographics of the people who could participate" in citizen science, says Jeff Goldman, director of programme development at CENS. Instead of having to remember to enter information or pictures through a website after the fact, people will be able to relay their observations directly from the field. Estrin notes that mobile phones also offer two-way communication, allowing participants to receive reminders and instructions depending on the time or location.

Other researchers are hoping to use mobile phones as communication hubs for external sensors. Sarah Williams, director of the Spatial Information Design Lab at Columbia University in New York, and her colleagues are working to attach air-quality sensors to

mobile phones so that people can send pollution measurements via text message to a central database. And a group led by computer scientist Eiman Kanjo at the University of Cambridge, UK, has developed a system that allows phones to receive data from wireless sensors that measure carbon monoxide concentrations, temperature and other environmental conditions[5]. Kanjo's team gave the mobile sensing system to cycling couriers in Cambridge to collect readings throughout the city and is now analysing the data.

Our phones, our selves

One obvious, but importantm feature of mobile phones is that they are carried by people. So researchers such as Nathan Eagle, a computer scientist at the Santa Fe Institute in New Mexico, can use them to get glimpses into human movement and behaviour.

In an experiment at the Massachusetts Institute of Technology in Cambridge during 2004–05, Eagle and his colleagues recorded call logs and location data from mobile phones carried by 100 students and staff — all volunteers — over nine months. They also detected when people were in close proximity to the users by scanning for radiofrequency signals emitted by nearby mobile phones. Using the phone data, Eagle's team was able to classify students studying business from those studying other subjects with 96% accuracy. If the researchers examined only the first 12 hours of a user's day, they could correctly predict the person's movements for the rest of the day 79% of the time[6].

"We think about behaviour as a very high-dimensional thing," says Eagle. "In reality, and depressingly, you can compress my behaviour down to a few vectors," he says.

Larger-scale experiments are also starting to emerge. Last June, Albert-László Barabási and his colleagues at Northeastern University in Boston, Massachusetts, published a study in *Nature* that analysed the movements of 100,000 mobile-phone users[7]. Eagle is now working with Barabási's group and others to examine phone-operator data from a range of geographic areas, including records for millions of mobile-phone users in Europe and two East African countries. Eventually, Eagle hopes to detect common behavioural patterns, such as changes in movement or calling frequency, that occur during disease outbreaks, which could help alert public-health officials to the early stages of an epidemic.

Eagle's research illustrates how mobile phones can be used to collect accurate, large-scale data about real social interactions, unlike other methods such as interviews or virtual-world observations, says Jon Kleinberg, a computer scientist who studies social networks at Cornell University in Ithaca, New York. Neil Ferguson, a mathematical epidemiologist at Imperial College London, UK, who plans to collaborate with Eagle, says that although mobile-phone use in places such as Africa may not reflect a representative sample of the population, these insights into microscale social networks could help support more fine-grained models of the spread of disease.

At the other end of the spectrum, some groups are exploring public-health applications at the individual level. Researchers at CENS and Intel Research Seattle in Washington, for instance, are developing mobile-phone programs to help users monitor their diet and physical activity.

Hang-ups ahead

The scientists doing these experiments readily admit that mobile phones are not the perfect tool. For one thing, the sensors on a phone are usually not high-quality because they must be small and inexpensive. "You don't want your phone to cost US$20,000," says Bayen. Because researchers cannot dictate what manufacturers include on a phone, the observations that can be performed are limited. And although features such as air-pollution sensors could conceivably be packaged into a mobile phone, people do not always carry their phones in ways that would make them useful as scientific instruments. "Sure, I'd love some air chemistry," says Estrin. "But what does the air chemistry in my purse mean?"

And as mobile phones now provide more information about their users than ever before, researchers must tread carefully to ensure they do not invade users' privacy (see page 968). In some cases, the users are volunteers who agree to be studied. The data from phone operators in Europe and Africa are anonymized, says Eagle. Mobile Millennium relies on a concept called 'virtual trip lines', in which velocity readings are triggered as users pass into predetermined road segments, rather than tracking drivers throughout their journey. The data are encrypted, says Bayen, and Nokia strips out personal information about the user before passing them on to the modelling team.

Once they get the data, researchers must contend with an erratic flow of information. Bayen's group is working to improve mathematical models so that they can handle traffic readings from unpredictable locations, but the reliability of the system fluctuates depending on how many measurements are submitted at the time. "Suddenly you get tons of data, and the next hour you don't get any," says Bayen. Scientists may find it difficult to get enough data to validate their approaches without help from industry, says Jacobson. "A lot of this research needs to be taken out on a commercial scale to even test the fundamental premise," he says.

Despite the challenges, researchers are excited about the possibilities of a planet-wide network of these miniature travelling computers. About 85% of the world's population has access to a mobile signal, says Susan Teltscher, head of the Market Information and Statistics Division at the International Telecommunication Union, and there is still a "huge potential" for more growth. Jacobson envisions that as sensors continue to drop in size, phones could boast even more sophisticated features.

Although mobile phones will not replace traditional scientific instruments, says Estrin, they make up in availability for what they lack in finesse. "If you can't go to the field with the sensor you want," she says, "go with the sensor you have." ∎

Roberta Kwok is a freelance science writer based in California.

> "In reality, and depressingly, you can compress my behaviour down to a few vectors."
> — Nathan Eagle

1. International Telecommunication Union *Measuring the Information Society — The ICT Development Index* (ITU, 2009).
2. Work, D. B., Tossavainen, O.-P., Jacobson, Q. & Bayen, A. M. *Proc. 2009 Am. Control Conf.* (in the press).
3. Granot, Y., Ivorra, A., & Rubinsky, B. *PLoS ONE* **3,** e2075 (2008).
4. Laufer, S. & Rubinsky, B. *PLoS ONE* **4,** e5178 (2009).
5. Kanjo, E. *et al. Pers. Ubiq. Comput.* **12,** 599–607 (2008).
6. Eagle, N. & Pentland, A. S. *Behav. Ecol. Sociobiol.* **63,** 1057–1066 (2009).
7. González, M. C., Hidalgo, C. A. & Barabási, A.-L. *Nature* **453,** 779–782 (2008).