Smarter Phones for All

The amount of information we can transmit though the air is limited by the laws of physics, but the mathematics of signal processing lets us squeeze more data into the same amount of space. As a result, we get better, cheaper and faster mobile phone calls.

There are more active mobile phones in the UK than people: over 70 million at the last count, thanks to multiple handset or SIM card ownership. What most don't realise is that the mobile communications industry is only made possible by the mathematical study of signal processing, which allows us to extract useful information from the noisy, invisible sea of radio signals above our heads. The rise of smartphones and mobile internet will introduce new challenges to the mobile networks, but cutting-edge mathematics is set to provide cheaper, more energy efficient and better quality communications for all.

Mobile phones transmit signals to a nearby base station via radio waves, a small part of the electromagnetic spectrum that also includes visible light, microwaves, and X-rays. Waves in the spectrum are described by their frequency, the number of oscillations per second, and two mobiles attempting to communicate with a base station using the same frequency will interfere with each other. This places a fundamental physical limit on how many mobile users can squeeze into the available frequency bands, making the spectrum a scarce natural resource that must be regulated by the Government to ensure it is used fairly.

The spectrum is also an incredibly valuable resource, as demonstrated by the Government's £22.5 billion auction of the 3G frequency band in 2000. Just as tangible resources such as gold or oil become more expensive when demand rises, so too will demand for mobile communications increase

the spectrum's value, but unlike these other resources the spectrum can be used more efficiently with the help of mathematical research.

> ## "It's like finding a mathematical equation for turning thin air into **gold**."

Mobile networks, like all forms of communication, are underpinned by a branch of mathematics called information theory. It was founded in the late 1940s by the American mathematician Claude Shannon, who realised there is an upper limit on the amount of information that you can send over a communications channel, such as a radio frequency band, before errors start to creep in. Reaching this "Shannon limit" requires a mathematical description of the message called an error-correcting code, but for decades the best codes could only achieve around half-capacity. It wasn't until the 1990s that researchers came close to unlocking the full potential of communications channels, with a new method called turbo codes.

Now, mathematicians in the UK are developing methods for even better

communications. Although turbocodes help to get the most out of a communications channel, there are also factors in the physical world that affect channel performance. Signals in a complex urban environment scatter when they bounce off buildings, causing echoes that take longer to reach their destination. These delayed echoes can interfere and cancel out when they meet again at the receiver, leading to dead zones and dropped phone calls. But rather than viewing this as a problem, engineers and mathematicians such as John McWhirter at Cardiff University have figured out ways to make these multiple-path effects actually improve signal transmission.

These improvements exploit an important new development in broadcasting technology called MIMO, or "multiple-input and multiple-output". MIMO uses arrays of "smart" radio antennas in both the transmitter and receiver, combined with software that tunes in on the direction in which a signal is strongest. The process is similar to tuning an FM radio into your favourite station, but rather than twiddling the radio dial, a mathematical algorithm rapidly tries different array configurations until it finds the best signal. Currently used algorithms are inefficient because they can't solve the problem without an intermediate stage, but in recent years McWhirter and his colleagues have developed methods to tackle it directly.

Institute of **mathematics** & its applications

EPSRC Pioneering research and skills

Transmitting a signal with a MIMO system is a problem in linear algebra, a branch of mathematics involving grids called matrices. The matrix in a MIMO system describes how a signal is transformed during transmission due to noise and other errors, and recovering the original transmitted signal from the altered received signal means undoing these changes. In mathematical terms, this is called "inverting the matrix". It's easy to invert simple matrices because they are just grids of numbers, but the matrices describing a MIMO system are anything but simple.

Due to the time-delays introduced by multiple-path effects, MIMO matrices aren't just grids of numbers, but actually contain a long mathematical expression called a "polynomial" in each cell. These are the best way to describe the complicated interactions of multiple-path MIMO signals, so finding an algorithm to invert them is a top priority. The current solution converts the polynomial matrix into a series of simple grids that can

more easily be inverted, but this comes at a price because it can use up more of the precious spectrum.

McWhirter hopes that removing this extra step will allow a more efficient use of the available spectrum. His new algorithms for inverting polynomial matrices directly could enable more mobile phone users to make better quality calls while also using less energy, an important goal when levels of mobile communication are constantly increasing. In terms of exploiting natural resources, it's like finding a mathematical equation for turning thin air into gold.

With smartphones and mobile internet changing the way we talk and access information, the future of mobile communications promises to be exciting, but fast, low-power and high-quality connections won't be enabled by technology alone. Mathematicians will help network operators make these changes by packing

more information into the limited spectrum with techniques such as turbo codes and polynomial matrix inversion, thus enabling the industry to continue meeting the insatiable demand for communication on the go.

TECHNICAL SUPPLEMENT

Error-correcting and turbo codes

In 1948, Claude Shannon showed that it is always possible to send error-free information over a noisy communications channel, provided you use a code with some redundancy. A basic code might simply repeat each section of the message three times; e.g. the string 101 would be transmitted as 101101101. If the person on the other end receives 101111101 instead, they know something has gone wrong but can still extract the correct string.

This simple code is not very good, however. The repetition makes transmitting a message take three times longer, and multiple errors render it useless – does 111111101 correspond to 111 with one error or 101 with two? Shannon proved that much more efficient codes existed that would enable the maximum rate of data transmission (the "Shannon limit"), but for decades even the best codes could only send little more than 1 kilobyte per second.

The breakthrough came in 1993, when Claude Berrou, Alain Glavieux and Punya Thitimajshima announced a new type of code that came very close to the Shannon limit. Nicknamed turbo codes, they use pairs of encoders and decoders to boost transmission speeds. The two encoders transmit different versions of the same message and the two decoders make statistical guesses about the bits in each versions. Both decoders rate how confident they are that a particular bit is a 0 or a 1 and compare their answers. After a number of iterations, the two decoders agree and the original message is recovered.

Polynomial matrix inversion in MIMO systems

The transmitted and received signals in a MIMO system are represented by two different vectors, with each number in the vector corresponding to one of the many receivers or transmitters in the system, and the channel is represented by a matrix describing ho w the signal changes during transmission.

It's easy to find the received signal vector by multiplying the transmitted signal vector and the channel matrix together, but the problem mobile phones need to solve is the exact opposite: working out the unknown transmitted vector from the received vector and the channel matrix. This involves inverting the channel matrix; a more difficult problem than multiplication, but one that is easily solved for simple matrices.

Polynomial matrices are currently inverted by first reducing them to a set of simple scalar matrices using a technique called orthogonal frequency-division multiplexing (OFDM), then inverting these scalar matrices with traditional algorithms. Each scalar matrix corresponds to a particular set of frequencies within the frequency range of the communication channel, but separating the frequencies in this way results in a loss of degrees of freedom and a less efficient use of the spectrum.

References

Berrou, C., Glavieux, A. & Thitimajshima, P. (1993) Near Shannon limit error-correcting coding and decoding: Turbo-codes. In: *Proceedings of the IEEE International Conference on Communications*, Geneva, Switzerland, May 1993, pp. 1064-1070. DOI: 10.1109/ICC.1993.397441

Shannon, C.E. (1948) A mathematical theory of communication. *The Bell System Technical Journal* 27(1), 379-423. DOI: 10.1145/584091.584093

Foster, J., McWhirter, J.G., Davies, M. & Chambers, J.A. (2010) An algorithm for calculating the QR and singular value decompositions of polynomial matrices. *IEEE Transactions on Signal Processing*, 58 (3), 1263-1274. DOI:10.1109/TSP.2009.2034325

McWhirter, J.G., Baxter, P.D., Cooper, T., Redif, S. & Foster, J. (2007) An EVD algorithm for para-Hermitian polynomial matrices. *IEEE Transactions on Signal Processing*, 55 (5), 2158-2169. DOI: 10.1109/TSP.2007.893222