GCHQ taps fibre-optic cables for secret access to world's communications

21. Juni 2013

British spy agency collects and stores vast quantities of global email messages, Facebook posts, internet histories and calls, and shares them with NSA, latest documents from Edward Snowden reveal

Ewen MacAskill, Julian Borger, Nick Hopkins, Nick Davies and James Ball, The Guardian

Britain's spy agency GCHQ has secretly gained access to the network of cables which carry the world's phone calls and internet traffic and has started to process vast streams of sensitive personal information which it is sharing with its American partner, the National Security Agency (NSA).

The sheer scale of the agency's ambition is reflected in the titles of its two principal components: Mastering the Internet and Global Telecoms Exploitation, aimed at scooping up as much online and telephone traffic as possible. This is all being carried out without any form of public acknowledgement or debate.

One key innovation has been GCHQ's ability to tap into and store huge volumes of data drawn from fibre-optic cables for up to 30 days so that it can be sifted and analysed. That operation, codenamed Tempora, has been running for some 18 months.

GCHQ and the NSA are consequently able to access and process vast quantities of communications between entirely innocent people, as well as targeted suspects.

This includes recordings of phone calls, the content of email messages, entries on Facebook and the history of any internet user's access to websites – all of which is deemed legal, even though the warrant system was supposed to limit interception to a specified range of targets.

The existence of the programme has been disclosed in documents shown to the Guardian by the NSA whistleblower Edward Snowden as part of his attempt to expose what he has called "the largest programme of suspicionless surveillance in human history".

"It's not just a US problem. The UK has a huge dog in this fight," Snowden told the Guardian. "They [GCHQ] are worse than the US."

However, on Friday a source with knowledge of intelligence argued that the data was collected legally under a system of safeguards, and had provided material that had led to significant breakthroughs in detecting and preventing serious crime.

Britain's technical capacity to tap into the cables that carry the world's communications – referred to in the documents as special source exploitation – has made GCHQ an intelligence superpower.

By 2010, two years after the project was first trialled, it was able to boast it had the "biggest internet access" of any member of the Five Eyes electronic eavesdropping alliance, comprising the US, UK, Canada, Australia and New Zealand.

UK officials could also claim GCHQ "produces larger amounts of metadata than NSA". (Metadata describes basic information on who has been contacting whom, without detailing the content.)

By May last year 300 analysts from GCHQ, and 250 from the NSA, had been assigned to sift through the flood of data.

The Americans were given guidelines for its use, but were told in legal briefings by GCHQ lawyers: "We have a light oversight regime compared with the US".

When it came to judging the necessity and proportionality of what they were allowed to look for, would-be American users were told it was "your call".

The Guardian understands that a total of 850,000 NSA employees and US private contractors with top secret clearance had access to GCHQ databases.

The documents reveal that by last year GCHQ was handling 600m "telephone events" each day, had tapped more than 200 fibre-optic cables and was able to process data from at least 46 of them at a time.

Each of the cables carries data at a rate of 10 gigabits per second, so the tapped cables had the capacity, in theory, to deliver more than 21 petabytes a day – equivalent to sending all the information in all the books in the British Library 192 times every 24 hours.

And the scale of the programme is constantly increasing as more cables are tapped and GCHQ data storage facilities in the UK and abroad are expanded with the aim of processing terabits (thousands of gigabits) of data at a time.

For the 2 billion users of the world wide web, Tempora represents a window on to their everyday lives, sucking up every form of communication from the fibre-optic cables that ring the world.

The NSA has meanwhile opened a second window, in the form of the Prism operation, revealed earlier this month by the Guardian, from which it secured access to the internal systems of global companies that service the internet.

The GCHQ mass tapping operation has been built up over five years by attaching intercept probes to transatlantic fibre-optic cables where they land on British shores carrying data to western Europe from telephone exchanges and internet servers in north America.

This was done under secret agreements with commercial companies, described in one document as "intercept partners".

The papers seen by the Guardian suggest some companies have been paid for the cost of their co-operation and GCHQ went to great lengths to keep their names secret. They were assigned "sensitive relationship teams" and staff were urged in one internal guidance paper to disguise the origin of "special source" material in their reports for fear that the role of the companies as intercept partners would cause "high-level political fallout".

The source with knowledge of intelligence said on Friday the companies were obliged to co-operate in this operation. They are forbidden from revealing the existence of warrants compelling them to allow GCHQ access to the cables.

"There's an overarching condition of the licensing of the companies that they have to co-operate in this. Should they decline, we can compel them to do so. They have no choice."

The source said that although GCHQ was collecting a "vast haystack of data" what they were looking for was "needles".

"Essentially, we have a process that allows us to select a small number of needles in a haystack. We are not looking at every piece of straw. There are certain triggers that allow you to discard or not examine a lot of data so you are just looking at needles. If you had the impression we are reading millions of emails, we are not. There is no intention in this whole programme to use it for looking at UK domestic traffic – British people talking to each other," the source said.

He explained that when such "needles" were found a log was made and the interception commissioner could see that log.

"The criteria are security, terror, organised crime. And economic well-being. There's an auditing process to go back through the logs and see if it was justified or not. The vast majority of the data is discarded without being looked at … we simply don't have the resources."

However, the legitimacy of the operation is in doubt. According to GCHQ's legal advice, it was given the go-ahead by applying old law to new technology. The 2000 Regulation of Investigatory Powers Act (Ripa) requires the tapping of defined targets to be authorised by a warrant signed by the home secretary or foreign secretary.

However, an obscure clause allows the foreign secretary to sign a certificate for the interception of broad categories of material, as long as one end of the monitored communications is abroad. But the nature of modern fibre-optic communications means that a proportion of internal UK traffic is relayed abroad and then returns through the cables.

Parliament passed the Ripa law to allow GCHQ to trawl for information, but it did so 13 years ago with no inkling of the scale on which GCHQ would attempt to exploit the certificates, enabling it to gather and process data regardless of whether it belongs to identified targets.

The categories of material have included fraud, drug trafficking and terrorism, but the criteria at any one time are secret and are not subject to any public debate. GCHQ's compliance with the certificates is audited by the agency itself, but the results of those audits are also secret.

An indication of how broad the dragnet can be was laid bare in advice from GCHQ's lawyers, who said it would be impossible to list the total number of people targeted because "this would be an infinite list which we couldn't manage".

There is an investigatory powers tribunal to look into complaints that the data gathered by GCHQ has been improperly used, but the agency reassured NSA analysts in the early days of the programme, in 2009: "So far they have always found in our favour".

Historically, the spy agencies have intercepted international communications by focusing on microwave towers and satellites. The NSA's intercept station at Menwith Hill in North Yorkshire played a leading role in this. One internal document quotes the head of the NSA, Lieutenant

General Keith Alexander, on a visit to Menwith Hill in June 2008, asking: "Why can't we collect all the signals all the time? Sounds like a good summer project for Menwith."

By then, however, satellite interception accounted for only a small part of the network traffic. Most of it now travels on fibre-optic cables, and the UK's position on the western edge of Europe gave it natural access to cables emerging from the Atlantic.

The data collected provides a powerful tool in the hands of the security agencies, enabling them to sift for evidence of serious crime. According to the source, it has allowed them to discover new techniques used by terrorists to avoid security checks and to identify terrorists planning atrocities. It has also been used against child exploitation networks and in the field of cyberdefence.

It was claimed on Friday that it directly led to the arrest and imprisonment of a cell in the Midlands who were planning co-ordinated attacks; to the arrest of five Luton-based individuals preparing acts of terror, and to the arrest of three London-based people planning attacks prior to the Olympics.

As the probes began to generate data, GCHQ set up a three-year trial at the GCHQ station in Bude, Cornwall. By the summer of 2011, GCHQ had probes attached to more than 200 internet links, each carrying data at 10 gigabits a second. "This is a massive amount of data!" as one internal slideshow put it. That summer, it brought NSA analysts into the Bude trials. In the autumn of 2011, it launched Tempora as a mainstream programme, shared with the Americans.

The intercept probes on the transatlantic cables gave GCHQ access to its special source exploitation. Tempora allowed the agency to set up internet buffers so it could not simply watch the data live but also store it – for three days in the case of content and 30 days for metadata.

"Internet buffers represent an exciting opportunity to get direct access to enormous amounts of GCHQ's special source data," one document explained.

The processing centres apply a series of sophisticated computer programmes in order to filter the material through what is known as MVR – massive volume reduction. The first filter immediately rejects high-volume, low-value traffic, such as peer-to-peer downloads, which reduces the volume by about 30%. Others pull out packets of information relating to "selectors" – search terms including subjects, phone numbers and email addresses of interest. Some 40,000 of these were chosen by GCHQ and 31,000 by the NSA. Most of the information extracted is "content", such as recordings of phone calls or the substance of email messages. The rest is metadata.

The GCHQ documents that the Guardian has seen illustrate a constant effort to build up storage capacity at the stations at Cheltenham, Bude and at one overseas location, as well a search for ways to maintain the agency's comparative advantage as the world's leading communications companies increasingly route their cables through Asia to cut costs. Meanwhile, technical work is ongoing to expand GCHQ's capacity to ingest data from new super cables carrying data at 100 gigabits a second. As one training slide told new users: "You are in an enviable position – have fun and make the most of it."