ACM NEWS

Dutch Police Fight Crime by Cracking PGP Phones

By Arnout Jaspers
July 10, 2018
Comments

Dutch police unit Team High Tech Crime decrypted text messages sent via Blackberry mobile devices that featured supposedly unbreakable PGP encryption. As a result, the police were able to read texts ordering a murder.

Credit: regmedia.co.uk

It was a stunning success for the Dutch National High Tech Crime Unit (NHTCU), a dedicated team within the Dutch National Police Agency that investigates advanced forms of cybercrime.

Last year, the Dutch police unit, also known as Team High Tech Crime (THTC), announced that as part of a criminal investigation, it had decrypted text messages sent via Blackberry mobile devices provided by the company Ennetcom, whose servers had been seized and copied by the police in 2016. The Blackberry phones, which featured supposedly unbreakable PGP encryption, were alleged to be a favorite means of communication for organized crime in the Netherlands.

In May 2018, a person nicknamed 'Noffel' was sentenced to 18 years in prison by a Dutch court, largely based on evidence from the decrypted messages that indicated he had ordered the murder of another person.

The THTC will not comment on how it decrypted the messages. According to Marc Stevens, a cryptographer at the Centrum Wiskunde & Informatica (CWI), the Netherlands' national research institute for mathematics and computer science, the most likely explanation for the decryption is that the individual secret keys for the users were generated and stored on the central server in Canada. A court file from September 2016, when a Canadian judge allowed the transfer the data from the Ennetcom servers to the Dutch police, indicates as much, without going into detail.

Said Stevens, "PGP encryption, by itself, is still unbroken, but it doesn't offer any security if your private keys are not secure as well." Users, he added, should always generate their own private keys.

A peculiarity of this kind of digital evidence gathering is that no human ever looks at all the data. The Ennetcom server data contained millions of text messages generated by thousands of users. The evidence in this case resulted from several rounds of automated search of the servers, with some human tinkering in between. This raises the question: how can the judge be sure that the evidence presented was a fair representation of the total amount of data? Maybe the search of the data overlooked important accomplices; maybe the search was biased in favor of incriminating evidence. A simple message like "Only kidding" can completely change the meaning of a previous message, and could be overlooked. Even if a suspect remembers this, it might go against his or her broader interest to acknowledge he sent any of these messages.

Gigabytes of data from the Ennetcom servers were processed and analyzed using Hansken, a forensic search engine developed at the Netherlands Forensic Institute (NFI). In general, such processing involves recovering deleted files from hard disks, unpacking archives, creating an index system, and so on.

The THCT then used Hansken to identify "criminal conversations" from the data. First, only messages in Dutch were selected. Even this is not trivial, because these text messages are full of slang and typos.

The thinking was that conversations classified as "criminal" would contain one or more keywords or nicknames that the THTC had compiled into a long list. The prosecutor claimed users of the purportedly secure Ennetcom phones obviously would not use criminal words like 'drugs' or 'kill' in their texts, but would replace them with code words like 'fruit' and 'put to sleep'.

After the first round of automated search, the list of keywords was expanded, to more than 100, and two more rounds of search were carried out.

During the court case against Noffel, independent data scientist David Graus was asked to examine this process. After studying documentation provided by the NFI and working with Hansken, Graus said he had had serious doubts about two main points: the crude methods to select conversations from the database, and the lack of software documentation because "Hansken is a machine that produces evidence from data, but the machine changes all the time."

Hansken was created by a team of digital forensics experts at the NFI, where team member Harm van Beek said, 'Hansken will never be finished; it's an arms race. Every year, there are new smartphones, new operating systems; we have to keep up with that.' On average, a new version comes out every three weeks, van Beek said.

The NFI describes Hansken as Digital Forensics as a Service "that processes multiple terabytes of digital material in a forensic context and gives easy and secure access to the processed results."

When searching text, Hansken still uses the decades-old weighting scheme term frequency–inverse document frequency (Tf-idf), an indicator of how often an individual term shows up within a text. According to Graus, much better context-based algorithms have become available which can search text for certain subjects while avoiding the arbitrary selection process based on dozens of keywords. Also, Hansken does not do advanced image recognition so it cannot, for instance, look for a weapon or a certain face in a set of pictures.

The NFI is working on improving Hansken, but cautiously. Team member Erwin van Eijk explained, "We must give a forensic validation of every automatic search method, and that is far from easy. We cannot just choose what to keep or ignore."

Forensics needs a high-recall search, which is quite different from the high-precision search that search engines like Google are good at providing. From a Google search, you might expect a short list of results featuring the most relevant hits for your query. Police detectives, however, want to be sure not a shred of evidence is overlooked by taking false leads for granted. Said van Eijk, "On Instagram, you see this one nice selfie, but all the failed, blurry selfies are still on that person's cellphone, and one of them may contain evidence."

For a few years now, Hansken has been used by the Dutch police to analyze all captured digital data. The largest amount of data handled with Hansken for a single case was 140 terabytes. Hansken can be used on a standalone computer, but in the Netherlands, it is cloud-based. Every time the police capture devices with huge amounts of data, a copy is uploaded to Hansken. Said van Beek, "In the Netherlands, uploading terabytes of data is usually not the bottleneck, because reading the data from a hard disk is slower than the upload speed of our network."

The data of all the cases Hansken has ever handled, amounting to more than a petabyte, is encrypted and stored in government datacenters, where it remains accessible on line to many forensic investigators simultaneously, if necessary.

Hansken is a treasure trove of data for the police, which is exactly what worries data scientist Graus. He said he was particularly amazed by the poor accountability of the process.

The private messages of thousands of Ennetcom users are now at the disposal of Dutch law enforcement. At least for criminal cases, there are legal limits to what fishing expeditions the police can perform on them.

Intelligence agencies collect huge amounts of data, too. Recently, a law was passed in the Netherlands giving them a much wider scope to collect Internet data. Nothing is publicly known about how they zoom in on suspects, but imagine terabytes of data from Facebook and Twitter being subjected to a keyword search with words like 'fruit' and phrases like 'put to sleep' to identify criminal or terrorist conversations, with the persons having those conversations winding up on a secret watch list.

Said Graus, "I sure hope they have something better than this."

Arnout Jaspers *is a freelance science writer based in Leiden, the Netherlands.*

No entries found