VICE OVER IP

A growing cadre of criminals is hiding secret messages in voice data

By JÓZEF LUBACZ, WOJCIECH MAZURCZYK & KRZYSZTOF SZCZYPIORSKI

MICK WIGGINS

7:00 P.M., SHANGHAI

An employee of an electronic equipment factory uploads a music file to an online file-sharing site. Hidden in the MP3 file (Michael Jackson's album *Thriller*) are schematics of a new mobile phone that will carry the brand of a large American company. Once the employee's Taiwanese collaborators download the file, they start manufacturing counterfeit mobile phones essentially identical to the original—even before the American company can get its version into stores.

3:30 P.M., SOMEWHERE IN AFGHANISTAN

A terrorist hunted by the U.S. Federal Bureau of Investigation posts an excerpt from the motion picture *High School Musical Three: Senior Year* on Facebook. Inside are hidden instructions for a bomb attack on a commuter rail line in southern Europe. Later that day, terrorists based in Athens follow the instructions to plan a rush hour attack that kills hundreds of people.

4:00 A.M., MALIBU, CALIF.

A very famous actor (VFA) has a brief conversation with a well-known director (WKD) over Skype, an application that lets them make free voice calls over the Internet. They discuss the medical problems of VFA's cat in great detail. When the conversation is over, WKD's computer has a sleazy new addition—in a folder on his desktop, there is a picture of a nude teenager, along with her mobile number and the date and time at which WKD will meet her at VFA's pool party for a photo session.

WHAT ALL these scenarios have in common is an information-smuggling technique called steganography—the communication of secret messages inside a perfectly innocent carrier. Think of steganography as meta-encryption: While encryption protects messages from being read by unauthorized parties, steganography lets the sender conceal the fact that he has even sent a message. After the 11 September attacks in 2001, rumors flew that they had been carried out with some help from steganography. A 2001 *New York Times* article described fake eBay listings in which routinely altered pictures of a sewing machine contained malevolent cargo. The link to 9/11 was never proved or disproved, but after those reports, the interest in steganographic techniques and their detection greatly increased.

Steganography use is on the rise, and not just among criminals, hackers, child pornographers, and terrorists. Persecuted citizens and dissidents under authoritarian regimes use it to evade government censorship, and journalists can use it to conceal sources. Investigators even use it on occasion to bait and trap people involved in industrial espionage: In the 1980s, to trace press leaks of cabinet documents, British Prime Minister Margaret Thatcher had government word processors altered to encode a specific user identity in the spaces between words. When leaked material was recovered, the identity of the leaker could be established by analyzing the pattern of those spaces.

Steganography is evolving alongside technology. A few years ago the cutting edge in steganographic tools involved hiding messages inside digital images or sound files, known as carriers, like that *Thriller* MP3. The technique quickly evolved to include video files, which are relatively large and can therefore conceal longer messages.

Now steganography has entered a new era, with stupendously greater potential for mischief. With the latest techniques, the limitations on the length of the message have basically been removed. Consider our example involving the use of Skype. Whereas the first two examples each required a carrier—an MP3 song and a video—there was no such requirement for the transmission of that nude photo. The data were secreted among the bits of a digital Voice over Internet Protocol conversation. In this new era of steganography, the mule that coconspirators are using is not the carrier itself but the communication protocols that govern the carrier's path through the Internet. Here's the advantage: The longer the communicators talk, the longer the secret message (or more detailed the secret image) they can send.

CARRIER EVOLUTION

Steganography has been used for at least 2500 years to disguise secret messages. In its earliest forms, the carriers were physical, but as technology evolved, so did carriers.

494 B.C.

HEAD TATTOO

Histiaeus tattoos a secret message onto a slave's shaved head, waits for the hair to regrow, and sends the slave to the intended recipient, who shaves off the hair to read the message.

480 B.C.

BEESWAX

Demaratus writes a secret message on a wooden tablet to warn the Greeks of Persian attack, and then covers it with many coats of wax.

1558

EGGS

Italian scientist Giambattista della Porta discovers how to hide a message inside a hard-boiled egg: Write on the shell using an ink made from a mixture of alum and vinegar. The solution leaves no trace on the surface,

Most strikingly, the concealment occurs within data whose inherent ephemerality makes the hidden payload nearly impossible to detect, let alone thwart.

We call this new technique network steganography. In our research at the Network Security Group at Warsaw University of Technology, we are studying the ever-evolving spectrum of carrier technologies, the increasing difficulty of detection as more sophisticated carriers leave fewer traces, and the implications of both for law enforcement and homeland security. Our work at Warsaw is literally self-defeating: We figure out the most advanced ways of doing network steganography and then design methods to detect them.

NETWORK STEGANOGRAPHY is a modern version of an old idea. You could argue that steganography helped spark the first major conflict between Greece and the Persian Empire. A classic use of steganography took place in 494 B.C., when Histiaeus, the ruler of Miletus, tried to instigate an Ionian revolt against the Persians. He shaved his favorite slave's head, tattooed it with a message, and waited for the slave's hair to grow back and obscure the tattoo. Then he sent the slave to his destination, where the intended recipient shaved the slave's head and read the message. The ensuing Ionian revolution lasted for half a century. In the 19th and 20th centuries, rapidly evolving warfare and espionage brought many innovations in steganography: Invisible ink, microdots, and Thatcher's word-processor trick are only a few among many.

With today's technology, information can be smuggled in essentially any type of digital file, including JPEGs or bitmaps, MP3s or WAV files, and MPEG movies. More than a hundred such steganographic applications are freely available on the Internet. Many of these programs are slick packages whose use requires no significant technical skills whatsoever. Typically, one mouse click selects the carrier, a second selects the secret information to be sent, and a third sends the message and its secret cargo. All the recipient needs is the same program the sender used; it typically extracts the hidden information within seconds.

ANOUCHKA/ISTOCKPHOTO

Any binary file can be concealed—for instance, pictures in unusual formats, software (a nasty virus, say), or blueprints. The favored carrier files are the most common ones, like JPEGs or MP3s. This emphasis on popular file formats increases the anonymity of the entire transaction, because these file types are so commonplace that they don't stick out.

The one limitation that steganographers have traditionally faced is file size. The rule of thumb is that you can use 10 percent of a carrier file's size to smuggle data. For an ambitious steganographer, that could be a problem: Imagine an electronic equipment factory employee trying to explain to the IT department why he has to send his mother a 100-megabyte picture of the family dog. For that reason, steganographers soon turned to audio and video files. A single 6-minute song, in the MP3 compression format, occupies 30 MB; it's enough to conceal every play Shakespeare ever wrote.

And yet, even with these precautions, conventional steganography still has an Achilles' heel: It leaves a trail. Pictures and other e-mail attachments stored on a company's outgoing e-mail servers retain the offending document. Anything sent has to bounce through some kind of relay and can therefore be captured, in theory.

Steganography poses serious threats to network security mainly by enabling confidential information leakage. The new crop of programs leaves almost no trail. Because they do not hide information inside digital files, instead using the protocol itself, detecting their existence is nearly impossible.

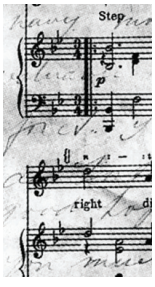
ALL THE new methods manipulate the Internet Protocol (IP), which is a fundamental part of any communication, voice or text based, that takes place on the Internet. The IP specifies how information travels through a network. Like postal service address standards, IP is mainly in charge of making sure that sender and destination addresses are valid, that parcels reach their destinations, and that those parcels conform to certain guidelines. (You can't send e-mail to an Internet address that does not use a 32-bit or 128-bit number, for example.)

but the message is retrieved by removing the shell and reading the egg.

1800s
NEWSPAPER CODE

During the Victorian era, lovers send secret letters by punching holes above certain letters. When the marked letters are combined, the message can be read.

1915
INVISIBLE INK



During World War I, entertainer and German spy Courtney de Rysbach performs in shows all over Britain as a cover for gathering information. Using invisible ink, Rysbach encodes secret messages by writing them in invisible ink on sheets of music.

1941
MICRODOTS

During World War II, German agents photographically shrink a page of text down to a 1-millimeter dot. The microdot is then hidden on top of a period in an otherwise unremarkable letter.

All traffic, be it e-mail or streaming video, travels via a method called packet switching, which parcels out digital data into small chunks, or packets, and sends them over a network shared by countless users. IP also contains the standards for packaging those packets.

Let's say you're sending an e-mail. After you hit the Send button, the packets travel easily through the network, from router to router, to the recipient's in-box. Once these packets reach the recipient, they are reconstituted into the full e-mail.

The important thing is that the packets don't need to reach their destination in any particular order. IP is a "connectionless protocol," which means that one node is free to send packets to another without setting up a prior connection, or circuit. This is a departure from previous methods, such as making a phone call in a public switched telephone network, which first requires synchronization between the two communicating nodes to set up a dedicated and exclusive circuit. Within reason, it doesn't matter when packets arrive or whether they arrive in order.

As you can imagine, this method works better for order-insensitive data like e-mail and static Web pages than it does for voice and video data. Whereas the quality of an e-mail message is immune to traffic obstructions, a network delay of even 20 milliseconds can very much degrade a second or two of video.

To cope with this challenge, network specialists came up with the Voice over Internet Protocol (VoIP). It governs the way voice data is broken up for transmission the same way IP manages messages that are less time sensitive. VoIP enables data packets representing a voice call to be split up and routed over the Internet.

The connection of a VoIP call consists of two phases: the signaling phase, followed by the voice-transport phase. The first phase establishes how the call will be encoded between the sending and receiving computers. During the second phase, data are sent in both directions in streams of packets. Each packet, which covers about 20 milliseconds of conversation, usually contains 20 to 160 bytes of voice data. The connection typically conveys between 20 and 50 such packets per second.

Telephone calls must occur in real time, and significant data delays would make for an awkward conversation. So to ferry a telephone call over the Internet, which was not originally intended for voice communications, VoIP makes use of two more communications protocols, which had to be layered on top of IP: The Real-Time Transport Protocol (RTP) and the User Datagram Protocol (UDP). The RTP gets time-sensitive video and audio data to its destination fast and so has been heavily adopted in much of streaming media, such as telephony, video teleconference applications, and Web-based push-to-talk features. To do that, it relies in turn on the UDP.

Because voice traffic is so time critical, UDP does not bother to check whether the data are reliable, intact, or even in order. So in a VoIP call, packets are sometimes stuck in out

ALL THREE STEGANOGRAPHIC IDEAS WE'VE OUTLINED HERE ARE SO SIMPLE, WE'RE CERTAIN THAT REAL-LIFE APPLICATIONS ARE ALREADY OUT THERE

of sequence. But that's not a big deal because the occasional misplaced packet won't significantly affect the quality of the phone call. The upshot of UDP is that the protocol opens a direct connection between computers with no mediation, harking back to the era of circuit switching: Applications can send data packets to other computers on a connection without previously setting up any special transmission channels or data paths. That means it's completely private.

Compared to old-fashioned telephony, IP is unreliable. That unreliability may result in several classes of error, including data corruption and lost data packets. Steganography exploits those errors.

Because these secret data packets, or "steganograms," are interspersed among many IP packets and don't linger anywhere except in the recipient's computer, there is no easy way for an investigator—who could download a suspect image or analyze an audio file at his convenience—to detect them.

TO BETTER UNDERSTAND what security officials will soon have to deal with, we designed and developed three flavors of network steganography, all of which manipulate IP. The three methods we developed are Lost Audio Packet Steganography, or LACK; Hidden Communication System for Corrupted Networks (HICCUPS); and Protocol Steganography for VoIP application. As their names imply, these techniques exploit lost packets, corrupted packets, and hidden or unused data fields in the VoIP transmission protocol. LACK hides information in packet delays, HICCUPS disguises information as natural "distortion" or noise, and Protocol Steganography hides information in unused data fields.

In regular VoIP telephony, excessively delayed packets containing voice samples are considered useless by the receiver and thus discarded. LACK exploits this mechanism to transmit hidden data. Some of the sender's packets are intentionally delayed, and the steganograms are stowed away inside those delayed packets. To any node that is not "in the know"—that is, a nearby computer that does not have the steganography program installed—they appear useless and are ignored. But if the receiver has the proper software to understand the steganography, it will not discard the excessively delayed packets. It will know that these contain the hidden data [see diagram, "Hidden in the Network"].

The transmission capacity for this scheme depends on the system used to encode the voice and on the quality of the network—specifically, how well it handles packet loss and delays. Using a standard 32-bit-per-second codec, and accounting for a 3 percent packet loss introduced by the network and a 0.5 percent packet loss introduced by LACK itself, a smuggler could transmit about 160 bits per second. At that rate you might be able to transmit a medium-size, 13-kilobyte image or a 2000-word text file during a typical 9- to 13-minute VoIP conversation.

LACK's main selling points are that it is simple to use and hard to detect. The only way it could be detected is if the user tried to hide too many secret packets. In that case, the number of intentionally delayed packets—and therefore the introduced delay—would create a suspiciously abnormal voice connection that might attract the attention of any security officials monitoring the line. If the call was completed before those officials could intercept the packets, however, there would be nothing they could do to try to uncover and assemble the steganograms.

Where LACK relies on lost packets to smuggle steganograms, HICCUPS takes advantage of corrupted packets. HICCUPS is fast. Let's say you have an IEEE 802.11g network with a transmission capacity of 54 megabits per second, with 10 terminals and a 5 percent rate of corrupted frames. Over such a network, you could send hidden data at a rate higher than 200 kilobits per second. That's almost as fast as the ISDN lines that were all the rage in the 1990s.

HICCUPS works on wireless local area networks, such as plain old coffee shop Wi-Fi. In such a wireless environment, data are transmitted by a method called broadcasting, which shuttles data in groups called frames. Like many courier services, broadcasting doesn't concern itself with the contents of the data or whether the data contain errors. When a wireless network detects an error in a frame, the computer simply drops that corrupted frame. The responsibility for detecting dropped frames (and retransmitting them if necessary) is left to the origin and destination terminals.

So in a wireless local-area network, all the user terminals (laptops, for the most part) must have a way of differentiating good packets from corrupted ones. This error-checking mechanism is called the checksum, a kind of signature against which the integrity of the packets can be confirmed. The checksum is a numerical value assigned to a data packet based on the number of bits in that packet. A checksum program uses that value to authenticate that the data hasn't been corrupted.

When the receiver's computer gets a packet, it checks for errors using that packet's checksum. Normally, if the checksum is wrong, the computer discards that packet. But if a terminal has

1980s
WATERMARKING

AP PHOTO

In the 1980s, to trace press leaks of cabinet documents, British Prime Minister Margaret Thatcher has government word processors altered to encode a specific user identity in the spaces between words.

1990s
DIGITAL STEGANOGRAPHY

Researchers develop methods to secretly embed a signature in digital pictures and audio, exploiting the human visual system's varying sensitivity to contrast.

2003
STREAMING VIDEO

Video steganography is similar to image steganography, but more information may be transported in a stream of images.

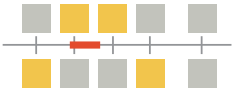
2007
NETWORK STEGANOGRAPHY

New methods focus on using free or unused fields in a protocol's headers.

HIDDEN IN THE NETWORK

HICCUPS
(CORRUPTED PACKETS)

Hides in normal corruption



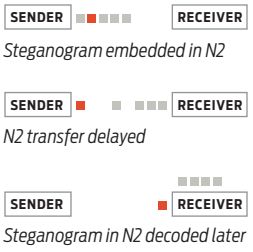
UNCORRUPTED DATA
CORRUPTED DATA
STEGANOGRAM

Highest information density HICCUPS [red] hides in the "noise" of natural distortion [orange] in an otherwise normal VoIP telephone call [gray].

Difficult to use Because this method requires hardware that can generate wrong checksums, it is difficult to use.

200 kilobits per second are transmitted during a typical 9–13 minute VoIP call.

LACK
(LOST AUDIO PACKETS)

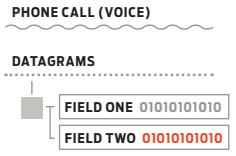


Lowest information density Excessively delayed packets are dropped by the receiver. LACK delays packets on purpose, encodes the hidden data, and decodes the steganograms when they arrive.

Hardest to detect Used carefully, LACK delays only a small percentage of packets.

160 bits per second are transmitted during a typical call.

PROTOCOL STEGANOGRAPHY
(HIDDEN FIELDS)



VoIP breaks voice data into datagram packets. Each packet contains frequently changing fields, such as authentication.

Easiest to use Each bit (phone-call data) contains data fields. Some fields contain frequently changing data, which can be wholly or partially replaced with a steganogram.

Hard to detect By replacing the authentication field, the user sacrifices security.

1–300 bits per second are transmitted during a typical call.

the right steganography program installed, it won't discard these intentionally wrong checksums—instead, it will know that these are precisely the data packets to scan for steganograms.

HICCUPS is more difficult to pull off than LACK. That's because this method requires a wireless card that can control frame checksums (good luck finding one of those at RadioShack). Network cards create checksums at the hardware level. We have applied for a patent in Poland for a HICCUPS-enabled card that can control checksums, but so far we haven't built our own card. Detecting HICCUPS wouldn't be easy. You'd need some way of observing the number of frames with incorrect checksums. If the number of those frames is statistically anomalous, then you might suspect the transmission of hidden information. Another way of detecting HICCUPS would analyze the content of those dropped—and therefore retransmitted—frames in order to detect the differences between the dropped and retransmitted frames. Major differences in these frames would constitute an obvious clue to nefarious goings-on.

Any of these detection methods, of course, would require not only that an investigator be aware that a transmission was about to take place but also that he be equipped with the right equipment, ready to monitor the conversation and intercept bits. Such a situation would be unlikely, to put it mildly.

The third method, Protocol Steganography, is a common name for a group of methods that use another aspect of IP: packet header fields. These fields are like sophisticated address labels that identify the contents of data packets to the recipient. Steganograms can be hidden inside unused, optional, or partial fields, because any data in these fields can be replaced without affecting the connection. Some of the more ham-fisted steganography techniques simply replace the content of the unused or optional fields with steganograms. But that would be relatively easy to detect and even jam.

So, to evade detection by simple analysis, the more sophisticated variant of Protocol Steganography uses fields in which the content changes frequently. For example, some of the more esoteric VoIP fields carry security data for authentication purposes. That little authentication subfield changes frequently during the course of a normal call. A steganogram smuggled inside one of its many randomly changing packets would be extremely hard to detect. Of course, there is a trade-off: The user would also sacrifice security, meaning that his or her conversation could be intercepted more easily.

Minimizing the threat of evolving steganography methods requires an in-depth understanding of how network protocols function and how they can be exploited to hide data. The problem is, however, the complexity of today's network protocols. All three steganographic ideas we've outlined here are so simple, we're certain that real-life applications are sure to come, if they aren't already out there. In fact, much more sophisticated methods will appear as Internet communication evolves from VoIP to other real-time media communications, such as video chat and conferencing.

THE ANONYMITY OF STEGANOGRAPHY might be good for privacy, but it also multiplies the threats to individuals, societies, and states. The trade-off between the benefits and threats involves many complex ethical, legal, and technological issues. We'll leave them for other thinkers and other articles.

What we're trying to do is understand what kind of potential contemporary communication networks have for enabling steganography, and in effect, create new techniques so that we can figure out how to thwart them. Some readers may object to our detailed descriptions of how these methods can be harnessed. But we would counter that unless someone shows how easy all this is, researchers won't understand the urgency and be inspired to develop protective measures. Not only can VoIP steganography be implemented in telephony tools that require a laptop or PC (like Skype), it can also be used in hard phones, such as the Android VoIP-enabled mobile phones that are starting to proliferate. Steganography on a phone is more difficult, because it requires access to the device's operating system, but no one should doubt that committed individuals will have no trouble rising to the challenge. As George Orwell once wrote, "On the whole human beings want to be good, but not too good, and not quite all the time." ❑