Steganography: How to Send a Secret Message

By Bryan Clair

8 October 2001

This may seem to be an ordinary beginning to an ordinary article. It is not. There's a secret message hidden here, in this very paragraph. It's not in view, and its source is modern. But the art of hiding messages is an ancient one, known as steganography.

Steganography is the dark cousin of cryptography, the use of codes. While cryptography provides privacy, steganography is intended to provide secrecy. Privacy is what you need when you use your credit card on the Internet -- you don't want your number revealed to the public. For this, you use cryptography, and send a coded pile of gibberish that only the web site can decipher. Though your code may be unbreakable, any hacker can look and see you've sent a message. For true secrecy, you don't want anyone to know you're sending a message at all.

Early steganography was messy. Before phones, before mail, before horses, messages were sent on foot. If you wanted to hide a message, you had two choices: have the messenger memorize it, or hide it on the messenger. In fact, the Chinese wrote messages on silk and encased them in balls of wax. The wax ball, "la wan," could then be hidden *in* the messenger.

Herodotus, an entertaining but less than reliable Greek historian, reports a more ingenious method. Histaeus, ruler of Miletus, wanted to send a message to his friend Aristagorus, urging revolt against the Persians. Histaeus shaved the head of his most trusted slave, then tattooed a message on the slave's scalp. After the hair grew back, the slave was sent to Aristagorus with the message safely hidden.

Later in Herodotus' histories, the Spartans received word that Xerxes was preparing to invade Greece. Their informant, Demeratus, was a Greek in exile in Persia. Fearing discovery, Demeratus wrote his message on the wood backing of a wax tablet. He then hid the message underneath a fresh layer of wax. The apparently blank tablet sailed easily past sentries on the road.

A more subtle method, nearly as old, is to use invisible ink. Described as early as the first century AD, invisible inks were commonly used for serious communications until WWII. The simplest are organic compounds, such as lemon juice, milk, or urine, all of which turn dark when held over a flame. In 1641, Bishop John Wilkins suggested onion juice, alum, ammonia salts, and for glow-in-the dark writing the "distilled Juice of Glowworms." Modern invisible inks fluoresce under ultraviolet light and are used as anti-counterfeit devices. For example, "VOID" is printed on checks and other official documents in an ink that appears under the strong ultraviolet light used for photocopies.

During the American revolution, both sides made extensive use of chemical inks that required special developers to detect, though the British had discovered the American formula by 1777. Throughout World War II, the two sides raced to create new secret inks and to find developers for the ink of the enemy. In the end, though, the volume of communications rendered invisible ink impractical.

With the advent of photography, microfilm was created as a way to store a large amount of information in a very small space. In both world wars, the Germans used "microdots" to hide information, a technique which J. Edgar Hoover called "the enemy's masterpiece of espionage." A secret message was photographed, reduced to the size of a printed period, then pasted into an innocuous cover message, magazine, or newspaper. The Americans caught on only when tipped by a double agent: "Watch out for the dots -- lots and lots of little dots."

Modern updates to these ideas use computers to make the hidden message even less noticeable. For example, laser printers can adjust spacing of lines and characters by less than 1/300th of an inch. To hide a zero, leave a standard space, and to hide a one leave 1/300th of an inch more than usual. Varying the spacing over an entire document can hide a short binary message that is undetectable by the human eye. Even better, this sort of trick stands up well to repeated photocopying.

All of these approaches to steganography have one thing in common -- they hide the secret message in the physical object which is sent. The cover message is merely a distraction, and could be anything. Of the innumerable variations on this theme, none will work for electronic communications because only the pure information of the cover message is transmitted. Nevertheless, there is plenty of room to hide secret information in a not-so-secret message. It just takes ingenuity.

The monk Johannes Trithemius, considered one of the founders of modern cryptography, had ingenuity in spades. His three volume work *Steganographia,* written around 1500, describes an extensive system for concealing secret messages within innocuous texts. On its surface, the book seems to be a magical text, and the initial reaction in the 16th century was so strong that *Steganographia* was only circulated privately until publication in 1606. But less than five years ago, Jim Reeds of AT&T Labs deciphered mysterious codes in the third volume, showing that Trithemius' work is more a treatise on cryptology than demonology. Reeds' fascinating account of the code breaking process is quite readable.

One of Trithemius' schemes was to conceal messages in long invocations of the names of angels, with the secret message appearing as a pattern of letters within the words. For example, as every other letter in every other word:

padiel a**p**o**r**s**y** mesarpon o**me**u**a**s peludyn m**a**l**p**re**a**x**o**

which reveals "prymus apex."

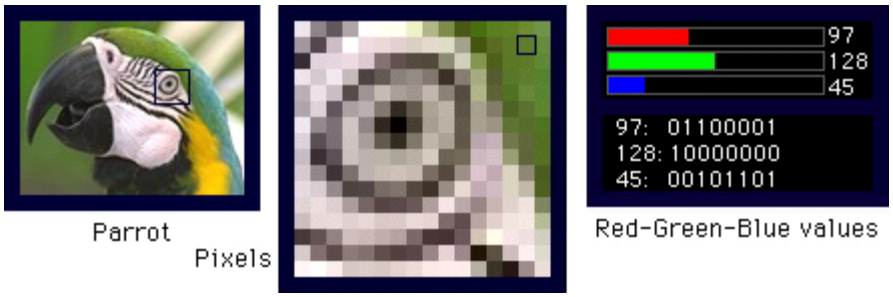
Another clever invention in *Steganographia* was the "Ave Maria" cipher. The book contains a series of tables, each of which has a list of words, one per letter. To code a message, the message letters are replaced by the corresponding words. If the tables are used in order, one table per letter, then the coded message will appear to be an

innocent prayer.

The modern version of Trithemius' scheme is undoubtedly SpamMimic. This simple system hides a short text message in a letter that looks exactly like spam, which is as ubiquitous on the Internet today as innocent prayers were in the 16th century. SpamMimic uses a "grammar" to make the messages. For example, a simple sentence in English is constructed with a subject, verb, and object, in that order. Given lists of 26 subjects, 26 verbs, and 26 objects, we could construct a three word sentence that encodes a three letter message. If you carefully prescribe a set of rules, you can make a grammar that describes spam.

Unfortunately, for serious users, every scheme we've seen is unacceptable. All are well known, and once a technique is suspected the hidden messages are easy to discover. Worse, a ten page document whose line spacing spells out a secret message is completely incriminating, even if the message is in an unbreakable code. A good steganographic technique should provide secrecy even if everyone knows it's being used.

The key innovation in recent years was to choose an innocent looking cover that contains plenty of random information, called white noise. You can hear white noise as a the nearly silent hiss of a blank tape playing. The secret message replaces the white noise, and if done properly it will appear to be as random as the noise was. The most popular methods use digitized photographs, so let's explore these techniques in some depth. Digitized photographs and video also harbor plenty of white noise. A digitized photograph is stored as an array of colored dots, called pixels. Each pixel typically has three numbers associated with it, one each for red, green, and blue intensities, and these values often range from 0-255. Each number is stored as eight bits (zeros and ones), with a one worth 128 in the most significant bit (on the left), then 64, 32, 16, 8, 4, 2, and a one in the least significant bit (on the right) worth just 1.



A difference of one or two in the intensities is imperceptible, and, in fact, a digitized picture can still look good if the least significant four bits of intensity are altered -- a change of up to 16 in the color's value. This gives plenty of space to hide a secret message. Text is usually stored with 8 bits per letter, so we could hide 1.5 letters in each pixel of the cover photo. A 640x480 pixel image, the size of a small computer monitor, can hold over 400,000 characters. That's a whole novel hidden in one modest photo!

Hiding a secret photo in a cover picture is even easier. Line them up, pixel by pixel. Take the important four bits of each color value for each pixel in the secret photo (the left ones). Replace the unimportant four bits in the cover photo (the right ones). The

cover photo won't change much, you won't lose much of the secret photo, but to an untrained eye you're sending a completely innocuous picture.

Unfortunately, anyone who cares to find your hidden image probably has a trained eye. The intensity values in the original cover image were white noise, i.e. random. The new values are strongly patterned, because they represent significant information of the secret image. This is the sort of change which is easily detectable by statistics. So the final trick to good steganography is make the message look random before hiding it.

One solution is simply to encode the message before hiding it. Using a good code, the coded message will appear just as random as the picture data it is replacing. Another approach is to spread the hidden information randomly over the photo. "Pseudo-random number" generators take a starting value, called a seed, and produce a string of numbers which appear random. For example, pick a number between 0 and 16 for a seed. Multiply your seed by 3, add 1, and take the remainder after division by 17. Repeat, repeat, repeat. Unless you picked 8, you'll find yourself somewhere in the sequence 1, 4, 13, 6, 2, 7, 5, 16, 15, 12, 3, 10, 14, 9, 11, 0, 1, 4, . . . which appears somewhat random. To spread a hidden message randomly over a cover picture, use the pseudo-random sequence of numbers as the pixel order. Descrambling the photo requires knowing the seed that started the pseudo-random number generator.

Here's a sample. The bear above is an adorable glow-in-the-dark skeleton costumed bear. The bear below is the same photo, now containing a hidden secret picture. To see the secret photo, get yourself a copy of S-Tools by Andy Brown and decrypt using the secret password "strange." Or, click here.

With these new techniques, a hidden message is indistinguishable from white noise. Even if the message is suspected, there is no proof of its existence. To actually prove there was a message, and not just randomness, the code needs to be cracked or the random number seed guessed. This feature of modern steganography is called "plausible deniability."

All of this sounds fairly nefarious, and in fact the obvious uses of steganography are for things like espionage. But there are a number of peaceful applications. The simplest and oldest are used in map making, where cartographers sometimes add a tiny fictional street to their maps, allowing them to prosecute copycats. A similar trick is to add fictional names to mailing lists as a check against unauthorized resellers.

Most of the newer applications use steganography like a watermark, to protect a copyright on information. Photo collections, sold on CD, often have hidden messages in the photos which allow detection of unauthorized use. The same technique applied to DVDs is even more effective, since the industry builds DVD recorders to detect and disallow copying of protected DVDs.

Even biological data, stored on DNA, may be a candidate for hidden messages, as biotech companies seek to prevent unauthorized use of their genetically engineered material. The technology is already in place for this: three New York researchers successfully hid a secret message in a DNA sequence and sent it across the country. Sound like science fiction? A secret message in DNA provided *Star Trek*'s explanation for the dubious fact that all aliens seem to be humans in prosthetic makeup!

Maybe, as in *Star Trek*, there really is a message hidden somewhere for humans to find. In the real world, the place to look for such a message is space, and humans have been looking for quite some time. Marconi, the inventor of radio, speculated that

strange signals heard by his company might be signals from another planet. To his credit, he was hearing these signals years before his competitors, but today they are known to be caused by lightning strikes.

In 1924, Mars passed relatively close to Earth, and the U.S. Army and Navy actually ordered their stations to quiet transmissions and listen for signals. They found nothing. In 1960, Dr. Frank Drake and a cadre of radio technicians used their 85 foot radio telescope for one of the first extensive studies of signals from space. They listened to Tau Ceti and Epsilon Erdani for 150 hours, and found nothing.

Today, the search for messages from space is underway on an unbelievable scale. The SETI@home project, based in Berkeley, has convinced millions of people to use their home computers in the search for signals. Their simple marketing trick was to package the calculations in a nifty screensaver, and now SETI@home is the largest computation in history. They've been looking for more than two years, with a telescope a thousand feet wide, but still they have found nothing.

Why have they found nothing? Maybe they haven't searched enough. But there is a dilemma here, the dilemma that empowers steganography. You never know if a message is hidden. You can search and search, but when you've found nothing you can only conclude: Maybe I didn't look hard enough, but maybe there is nothing to find.

Reader Comments

Bryan Clair is a professor of mathematics at Saint Louis University. His previous publications in *Strange Horizons* can be found in our Archive.

Further Reading

A comprehensive steganography website is F. Petitcolas' page, which has history and current research. To get a copy of S-Tools, or other shareware steganography programs, the best place to look is StegoArchive.Com.

The comprehensive history of cryptography reference is D. Kahn's *The Codebreakers*. An article from the journal *Cryptologia* has some interesting 16th and 17th century history.

For more on document marking by altering line spacing, look at the "Copyright Protection for the Electronic Distribution of Text Documents" paper on N. Maxemchuk's page. For more on DNA based steganography, try this article by Ivars Peterson.

Top