11/02/2021

Lecture 10

Topic: Steganography

Steganography conceals the existence of the message, whereas Cryptography renders the message unintelligible to outsiders by various transformations of the text

Consider an environment where the very use of encrypted messages causes suspicion. If a nefarious government or Internet service provider (ISP) is looking for encrypted messages, they can easily find them. Consider the following text file; what else is it likely to be if not encrypted?

qANQR1DBwU4D/TlT68XXuiUQCADfj2o4b4aFYBcWumA7hR1Wvz9rbv2BR6WbEUsy
ZBIEFtjyqCd96qF38sp9IQiJIKlNaZfx2GLRWikPZwchUXxB+AA5+lqsG/ELBvRa
c9XefaYpbbAZ6z6LkOQ+eE0XASe7aEEPfdxvZZT37dVyiyxuBBRYNLN8Bphdr2zv
z/9Ak4/OLnLiJRk05/2UNE5Z0a+3lcvITMmfGajvRhkXqocavPOKiin3hv7+Vx88
uLLem2/fQHZhGcQvkqZVqXx8SmNw5gzuvwjV1WHj9muDGBY0MkjiZIRI7azWnoU9
3KCnmpR60VO4rDRAS5uGl9fioSvze+q8XqxubaNsgdKkoD+tB/4u4c4tznLfw1L2
YBS+dzFDw5desMFSo7JkecAS4NB9jAu9K+f7PTAsesCBNETDd49BTOFFTWWavAfE
gLYcPrcn4s3EriUgvL3OzPR4P1chNu6sa3ZJkTBbriDoA3VpnqG3hxqfNyOlqAka
mJJuQ53Ob9ThaFH8YcE/VqUFdw+bQtrAJ6NpjIxi/x0FfOInhC/bBw7pDLXBFNaX
HdlLQRPQdrmnWskKznOSarxq4GjpRTQo4hpCRJJ5aU7tZO9HPTZXFG6iRITOwa47
AR5nvkEKoIAjW5HaDKiJriuWLdtN4OXecWvxFsjR32ebz76U8aLpAK87GZEyTzBx
dV+lH0hwyT/y1cZQ/E5USePP4oKWF4uqquPee1OPeFMBo4CvuGyhZXD/18Ft/53Y
WIebvdiCqsOoabK3jEfdGExce63zDIO=
=MpRf

The message above is a sentence in English that is encrypted using Pretty Good Privacy (PGP), probably the most commonly used e-mail encryption software today. Besides being nonsensical to a casual reader, the other indication that this is encrypted is that the characters comprising the message appear more-or-less at random and do not adhere to the relative frequency counts that one would expect in a non-encrypted message.

In this context, the cover_medium is the file in which we will hide the hidden_data, which may also be encrypted using the stego_key. The resultant file is the stego_medium (which will, of course. be the same type of file as the cover_medium). The cover_medium (and, thus, the stego_medium) are typically image or audio files.

In this lecture I will focus on image files and will, therefore, refer to the cover_image and stego_image.

Images typically use either 8-bit or 24-bit color. When using 8-bit color, there is a definition of up to 256 colors forming a palette for this image, each color denoted by an 8-bit value. A 24-bit color scheme, as the term suggests, uses 24 bits per pixel and provides a much better set of colors. In this case, each pix is represented by three bytes, each byte representing the intensity of the three primary colors red, green, and blue (RGB), respectively.

Images typically use either 8-bit or 24-bit color. When using 8-bit color, there is a definition of up to 256 colors forming a palette for this image, each color denoted by an 8-bit value. A 24-bit color scheme, as the term suggests, uses 24 bits per pixel and provides a much better set of colors. In this case, each pix is represented by three bytes, each byte representing the intensity of the three primary colors red, green, and blue (RGB), respectively.

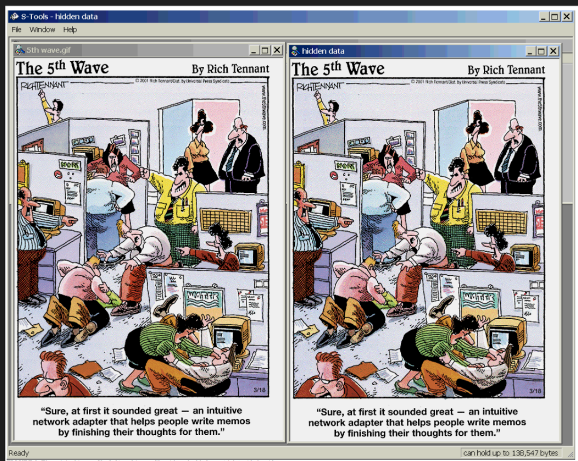

FIGURE 2. The original image file (left) and image file with embedded text (right), side by side.

Now suppose we want to "hide" the following 9 bits of data (the hidden data is usually compressed prior to being hidden): 101101101. If we overlay these 9 bits over the LSB of the 9 bytes above, we get the following

4 changes

10010101	00001101	11001001		10010101	00001100	11001001
10010110	00001111	11001010		10010111	00001110	11001011
10011111	00010000	11001011		10011111	00010000	11001011

While much of the steganography employed today is quite high-tech, steganography itself can make use of many low-tech methods.

One common, almost obvious, form of steganography is called a null cipher. In this type of stego, the hidden message is formed by taking the first (or other fixed) letter of each word in the cover message. Consider this cablegram that might have been sent by a journalist/spy from the U.S. to Europe during World War I:

PRESIDENT'S EMBARGO RULING SHOULD HAVE IMMEDIATE NOTICE. GRAVE SITUATION AFFECTING INTERNATIONAL LAW. STATEMENT FORESHADOWS RUIN OF MANY NEUTRALS. YELLOW JOURNALS UNIFYING NATIONAL EXCITEMENT IMMENSELY.

The first letters of each word form the character string: PERSHINGSAILSFROMNYJUNEI.

A little imagination and some spaces yields the real message:

PERSHING SAILS FROM NY JUNE I.

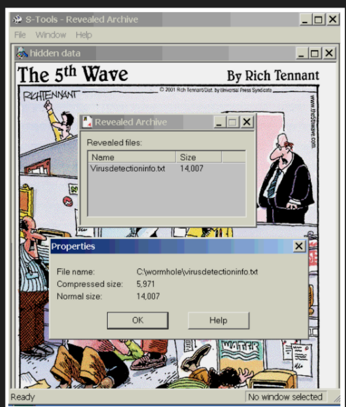

FIGURE 3. Extracting hidden information from the image file.

Another form of steganography uses a template (e.g., a piece of paper with holes cut in it) or a set of preselected locations on the page to hide a message. In this case, obviously, the sender and receiver must use the same template or rules. Consider this note:

THE MOST COMMON WORK ANIMAL IS THE HORSE. THEY CAN BE USED TO FERRY EQUIPMENT TO AND FROM WORKERS OR TO PULL A PLOW. BE CAREFUL, THOUGH, BECAUSE SOME HAVE SANK UP TO THEIR KNEES IN MUD OR SAND, SUCH AS AN INCIDENT AT THE BURLINGTON FACTORY LAST YEAR. BUT HORSES REMAIN A SIGNIFICANT FIND. ON A FARM, AN ALTERNATE WORK ANIMAL MIGHT BE A BURRO BUT THEY ARE NOT AS COMFORTABLE AS A TRANSPORT ANIMAL.

Applying a template or rule as to which words to read to this message might yield the following:

HORSE
FERRY
SANK
IN BURLINGTON
FIND
ALTERNATE
TRANSPORT