Abstract. We present a number of insights into information hiding. It was widely believed that public key steganography was impossible; we show how to do it. We then look at a number of possible approaches to the theoretical security of hidden communications. This turns out to hinge on the inefficiency of practical compression algorithms, and one of the most important parameters is whether the opponent is active or passive (i.e., whether the censor can add noise, or will merely allow or disallow a whole messages). However, there are covertexts whose compression characteristics are such that even an active opponent cannot always eliminate hidden channels completely.

1 Introduction

Steganography is about concealing the existence of messages, and it goes back to ancient times. Kahn tells of a classical Chinese practice of embedding a code ideogram at a prearranged place in a dispatch; of the warning the Greeks received of Xerxes' intentions from a message underneath the wax of a writing tablet; and a trick of dotting successive letters in a covertext with secret ink, due to Aeneas the Tactician [8].

The opponent may be passive, and merely observe the covertext, but he may also be active. In the US post office during the second world war, postal censors deleted lovers’ X’s, shifted watch hands, and replaced items such as loose stamps and blank paper. They also rephrased telegrams; in one case, a censor changed ‘father is dead’ to ‘father is deceased’, which elicited the reply ‘is father dead or deceased?’

The study of this subject in the open scientific literature may be traced to Simmons, who in 1983 formulated it as the prisoners’ problem [16]: Alice and Bob are in jail, and wish to hatch an escape plan. All their communications pass through the warden, Willy. If Willy sees any encrypted messages, he will frustrate their plan by putting them into solitary confinement. So they must find some way of hiding their ciphertext in an innocuous looking covertext. As in the related field of cryptography, we assume that the mechanism in use is known to the warden, and so the security must rely solely on a secret key.

There are many real life applications of steganography. Apparently, during the 1980’s, British Prime Minister Margaret Thatcher became so irritated at
press leaks of cabinet documents that she had the word processors programmed
to encode their identity in the word spacing of documents, so that disloyal minis-
ters could be traced. Similar techniques are now undergoing trials in an electronic
publishing project, with a view to hiding copyright messages and serial numbers
in documents [10].

Simmons' real application was more exotic — the verification of nuclear arms
control treaties. The US and the USSR wanted to place sensors in each others' nuclear
facilities that would transmit certain information (such as the number of
missiles) but not reveal other kinds of information (such as their location). This
forced a careful study of the ways in which one country's equipment might smuggle
out the forbidden information past the other country's monitoring facilities
[17, 19].

Steganography must not be confused with cryptography, where we transform
the message so as to make its meaning obscure to a person who intercepts it. Such
protection is often not enough: the detection of enciphered message traffic
between a soldier and a hostile government, or between a known drug-smuggler
and someone not yet under suspicion, has obvious implications.

However, we still have no comprehensive theory of steganography, in the way
that Shannon gave us a theory of encryption [15] and Simmons of authentication
[18]. In this article, we will try to move a few small steps towards such a theory.

2 The State of the Art

A number of computer programs are available that will embed a ciphertext file
in an image. The better systems assume that both sender and receiver share a
key and use a conventional cryptographic keystream generator [13] to expand
this into a long pseudo-random keystream. The keystream is then used to select
pixels in which the bits of the ciphertext are embedded.

Of course, not every pixel may be suitable for encoding ciphertext: changes
to pixels in large fields of monochrome colour, or that lie on sharply defined
boundaries, might be visible. So some systems have an algorithm that determines
whether a candidate pixel can be used by checking that the variance in luminosity
of the eight surrounding pixels is neither very high (as on a boundary) nor very
low (as in a monochrome field). A bit can be embedded in a pixel that passes this
test by some rule such as setting its low order bit to the parity of the surrounding
pixels (though in practice one might use something slightly more complicated to
avoid leaving telltale statistics).

Of course, the more bits per pixel, the less correlated the low order bits will
be with neighbouring bits and with higher order bits in the same pixel. Some
quantitative measurements of the correlations between pixels on different bit
planes in digital video may be found in [20]. In effect, the bits that Alice can use
to embed covert data are redundant in that Willy will be unaware that they have
been altered. It follows that they might be removed by an efficient compression
scheme, if one exists for the image or other covertext in use.