April 12, 2000

Protecting Intellectual Property in the Digital Era

Encryption

We've studied various methods during this semester. In practice, encryption methods used to protect intellectual property may rely both on "software" or "hardware". We've discussed some possible software methods relying on the difficulty of certain algorithms to be unravelled without knowledge of their keys. Special purpose hardware has been designed to forestall copying or access to intellectual property. This could take the form of special computer chips encased in decoder boxes or special attachments for computers ("dongles"). These can be unwieldy and unpopular. Digital property when converted to plaintext can perhaps just be copied unless additional intricate arrangements are made. Also, decryption can be elaborate and take time, and its intricacies can be another level of complexity subject to malfunction during delivery of the property. Additionally, of course, encryption schemes can be broken. The most notorious recent case is the cracking of the DVD protection scheme.

Here's a paragraph from an official Memorandum of the U.S. District Court published in January 2000 (source: a web page at `www.eff.org`):

> DVDs contain motion pictures in digital form, which presents an enhanced risk of unauthorized reproduction and distribution because digital copies made from DVDs do not degrade from generation to generation. Concerned about this risk, motion picture companies, including plaintiffs, insisted upon the development of an access control and copy prevention system to inhibit the unauthorized reproduction and distribution of motion pictures before they released films in the DVD format. The means now in use, Content Scramble System or CSS, is an encryption-based security and authentication system that requires the use of appropriately configured hardware such as a DVD player or a computer DVD drive to decrypt, unscramble and play back, but not copy, motion pictures on DVDs. CSS has been licensed to hundreds of DVD player manufacturers and DVD content distributors in the United States and around the world.

The licensees are required to keep the algorithms and keys secret, and to install tamper-proof hardware in each player which would contain the keys needed to read those DVD's legal for the player's area in the world. The DVD organization divided the world into several distinct areas: discs and players could each work only in designated areas.

CSS turns out to be a rather simple encryption scheme relying on 40 bit keys (not adequate by current encryption standards). The system was reverse engineered: that is, observation of the inputs and outputs of the system were used to figure out how the system worked. Also, one of the manufacturers mistakenly allowed some of the keys to be seen. Even extremely secure cryptosystems may be vulnerable when users make mistakes.

The following quotes are taken from pages on `http://www.wired.com`. The first is from, of course, a representative of those wishing to protect their intellectual property. The second is from the "crackers".

> "The circulation through the Internet of the illegal and inappropriate software is against the stream of copyright protection. Toshiba, which has led the establishment of the DVD format and is the chair-company of the DVD Forum, feels it is a great pity," wrote Masaki Mikura, manager of the strategic partnership and licensing division at Toshiba Ltd.
>
> Johansen and his partners were able to guess more than 170 working keys by trial and error before finally just giving up to go do something else. "I wonder how much they paid for someone to actually develop that weak algorithm," said Johansen. "It's a very weak encryption algorithm."

OVER

Steganography

From the web page `http://www.cl.cam.ac.uk/~fapp2/steganography/`:

> While cryptography is about protecting the content of messages ... **steganography** is about concealing their very existence. It comes from Greek roots and literally means "covered writing," and is usually interpreted to mean hiding information in other information.
>
> Until recently, information hiding techniques received very much less attention from the research community and from industry than cryptography, but this is changing rapidly ... The main driving force is concern over protecting copyright; as audio, video and other works become available in digital form, it may be that the ease with which perfect copies can be made will lead to large-scale unauthorized copying which will undermine the music, film, book and software publishing industries. There has therefore been significant recent research into "watermarking" (hidden copyright messages) and "fingerprinting" (hidden serial numbers or a set of characteristics that tend to distinguish an object from other similar objects); the idea is that the latter can be used to detect copyright violators and the former to prosecute them.
>
>

From the webpage `http://members.tripod.com/steganography/stego.html`:

> **What is Steganography?**
>
> In an ideal world we would all be able to openly send encrypted mail or files to each other with no fear of reprisals. However there are often cases when this is not possible, either because you are working for a company that does not allow encrypted email or perhaps the local government does not approve of encrypted communication (a reality in some parts of the world). This is where steganography can come into play.
>
> Steganography simply takes one piece of information and hides it within another. Computer files (images, sounds recordings, even disks) contain unused or insignificant areas of data. Steganography takes advantage of these areas, replacing them with information (encrypted mail, for instance). The files can then be sent or transported without anyone knowing what really lies inside of them. An image of the space shuttle landing might contain a private letter to a friend. A recording of a short sentence might contain your company's plans for a secret new product. Steganography can also be used to place a hidden "trademark" in images, music, and software, a technique referred to as watermarking.
>
>

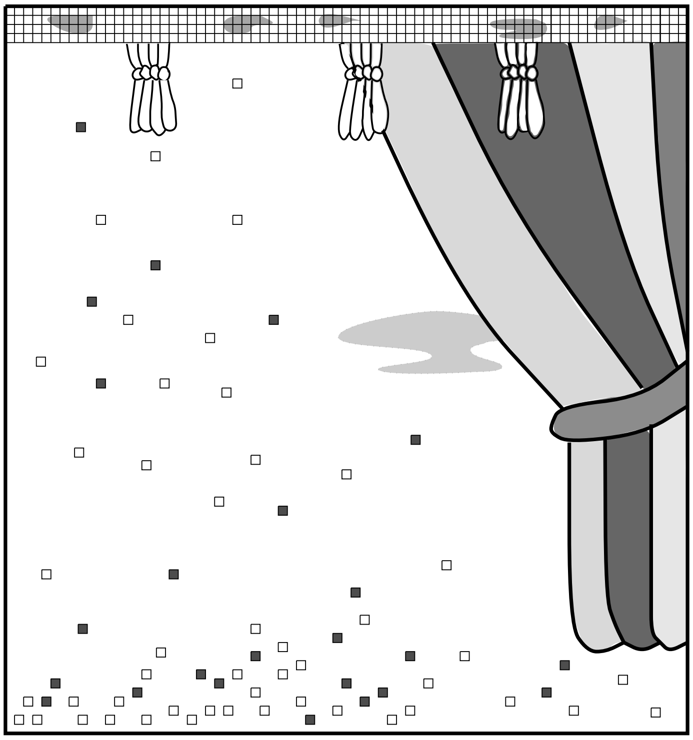
Steganography is not a commonly used word, but steganography is more commonly applied than is realized. These paragraphs are from a message on the cypherpunks mailing list:

> ... when the potential for counterfeiting of valuable documents on color copiers/xerographic printers became apparent in Japan (where such machines first appeared) manufacturers were concerned about negative governmental reaction to such technology. In an effort to stave off legislative efforts to restrict such devices, various ID systems began being implemented at that point. At one stage for at least one U.S. manufacturer, this was as crude as a serial number etched on the underside of the imaging area glass!
>
> Modern systems, which are now reportedly implemented universally, use much more sophisticated methods, encoding the ID effectively as "noise" repeatedly throughout the image, making it impossible to circumvent the system through copying or printing over a small portion of the image area, or by cutting off portions of printed documents. ...
>
> To read these IDs, the document in question is scanned and the "noise" decoded via a secret and proprietary algorithm. In the case of Xerox-manufactured equipment, only Xerox has the means to do this, and they require a court order to do so (except for some specific government agencies, for whom they no longer require court authorizations). I'm told that the number of requests Xerox receives for this service is on the order of a couple a week from within the U.S.
>
> In the copier case, that ID technology being used for color copies *could* be adapted to black and white copies and prints as well. The addition of cheap GPS [global positioning systems: my note] units to copiers could provide not only valid date/time stamps, but also the physical *locations* of the units, all of which could be invisibly encoded within the printed images.
>
> Pressures to extend the surveillance of commercial copyright enforcement take such concepts out of the realm of science-fiction, and into the range of actual future possibilities.

There are lots of bits in most digital property whose alteration would hardly be noticed by a human observer. Huge music files have lots of room: the ear can hardly hear that well. And pictures . . . so much can be hidden in even the simplest pictures.

A problem in steganography

Describe carefully at least 3 ways to hide 20 bits in the picture below, without harming its artistic *integrity or quality**. 20 bits is enough to encode more than a million distinct numbers since 2^{20} is 1,048,576.



BITSTORM

* Assuming there is any.