The Many Colors of Multimedia Security

Protection of artistic content from illegal distribution involves significant gray areas in terms of methods and laws.

Digital multimedia (whether it be audio, video, or still photography and art) is exposed to a broad spectrum of security problems. From the standpoint of the media provider, protection of materials from unauthorized distribution or modification is a primary concern. At the delivery end, recipients want to ensure that downloads are virus-free and legitimately obtained. Ironically, encryption and digital branding tools can be employed both for securing multimedia as well as for circumventing laws pertaining to content and use.

The most popular of these techniques are steganography (the art and science of embedding secret messages within text, sound, or imagery) and watermarking (the addition of an unremovable identifier to tag the content, indicating ownership). Although such methods have been around for millenia, they have progressed a long way since the backward playing of the "Revolution 9" track of the Beatles' *White* album evoked an eery message about Paul McCartney's supposed demise, or since Sherlock Holmes examined paper imprints to determine the origin of documents.

Applications of steganography and watermarking can include feature location (identification of subcomponents within a data set), captioning, time-stamping, and tamper-proofing (demonstration that original contents have not been altered). As it turns out, photographs taken of natural scenes are not truly random to begin with, according to Dartmouth's Hany Farid, who claims that inherent regularities make it possible to perform statistical analysis to determine whether or not an image has been altered, thus making forgeries mathematically detectable (see www.cs.dartmouth.edu/~farid/). The detection and extraction of embedded data may also be assisted with steganalysis tools. An extensive archive of steganography and steganalysis software is available at www.stegoarchive.com.

Data overlays are often used to conceal illegal content, such as child pornography or terrorist target maps. To assist in stemming this trend, files that are commonly traded or shared have been logged by law enforcement agencies so that the process of identifying particular files on storage devices can now be performed algorithmically. Having experienced sorting through thousands of images

MICHAEL SCHRÖTER

Poorly designed file-sharing services could be thought of as a type of honeypot—luring users in to obtain desirable multimedia content, while providing access for others to extract their private information.

manually some years ago as a computer expert in a murder case, I can attest that it is well appreciated by prosecution and defense teams alike that many of these types of files no longer have to be viewed individually.

Characteristics involved with data embedding [3] include:

- Visibility: embedded data may be intentionally detectable or imperceptible, but either way it should not detract from or degrade the primary media content.
- Robustness (or fragility): the ability of the data to withstand signal-processing attacks (such as compression, rescaling, and format conversions like digital-to-analog conversion).
- Error correction and detection: recovery is possible from small losses or an indication is provided that coded information damage has occurred.
- Header independence: data is encoded directly into the content of the file to allow survival between file format transfers.
- Self-clocking (or blind) coding: extraction does not require reference to the masking information or signal. (Adaptive coding algorithms use content from the masking data to perform hiding, usually through a transform-based method.)
- Asymmetrical coding: the process used to extract the information is not as time or resource consuming as the process used to insert it, to allow for quick access to the data.

Particular embedding techniques offer different trade-offs in the criteria listed here. In addition, encryption algorithms (such as XOR, DES, 3DES, IDEA, and AES) can be used along with steganography to further conceal the obscured data. In such cases, the password is usually hidden in the file, so the identification of it, plus the type of encryption used, is necessary to extract the hidden message.

One popular method used for data encoding in digital audio files is Least Significant Bit (LSB), where the low bit of each sound sample is successively replaced by digits from a binary string. Another is phase coding, where the phase of an audio signal is replaced by a reference phase used to modulate a data signal) [1]. Echoes can even be altered in order to conceal information. All these techniques rely on the fact that the perception of sound is rather inaccurate, and the data modifications used to hide message information are engineered such that they fall within what would normally be considered noise. Similarly, the human visual system is relatively insensitive to such things as small or continuous changes in brightness. So image steganography can employ bit and phase coding or more advanced techniques (like patchwork, texture block, and affine coding) along with encryption and compression, to modulate luminance while achieving a high degree of immunity from detection.

But when all such techniques fail as security mechanisms, media providers have another way of protecting their property—lawsuits made possible under the 1998 Digital Millennium Copyright Act (DMCA) [7]. This controversial bill states that "no person shall circumvent a technological measure that effectively controls access to a work" other than the copyright holder and those granted permission to make noninfringing uses. It is the limited scope of these noninfringing uses, particularly as pertains to their chilling effect on security research and education [4], along with restrictions on the manufacture of components that are "pri-

marily designed or produced for the purpose of circumventing a technological measure that effectively controls access to a work" that has been hotly debated in classrooms and courtrooms (as well as in *Communications* [8]).

In that vein, the Recording Industry Association of America (RIAA) continues to receive a mixture of praise and criticism for its practice of suing thousands of individuals under allegations of illegal file sharing. So far, nearly 700 cases have been settled (often out of court), with an average "monetary compensation" agreement of approximately $3,000. Targets of RIAA's Internet piracy accusations have included 12-year-old Briana LaHara of New York (whose mother paid a $2,000 settlement) and 66-year-old Massachusetts grandmother Sarah Seabury Ward, who was slammed with a $300 million lawsuit by seven major record labels (including Sony, BMG, Warner Brothers, and Arista) for allegedly downloading more than 2,000 rock and hip-hop tunes. Ward (represented by the Electronic Frontier Foundation) had her case eventually dropped because RIAA claimed she used KaZaA, a file-sharing service that was not accessible from the Macintosh computer she owned.

Legitimate file sharing was also deemed risky when researchers N.S. Good and A. Krekelberg discovered that peer-to-peer networks, such as KaZaA, can permit the sharing of all files on a user's hard drive without the owner's knowledge [5]. While the vast majority of participants believed that only multimedia files could be shared via P2P, actually over 60% of users were discovered to have inadvertently set up their accounts such that their Microsoft Outlook inboxes could be viewed. Other files that were accessible by KaZaA clients included Web browser cookies, data from financial software, and credit card information stored in Excel spreadsheets. Although a warning regarding the sharing of copyrighted files appears on KaZaA's home page, one has to wade through the security and privacy section to find the statement: "It is highly recommended that you do not share your entire hard drive or 'My Documents' folder." The study concluded the interface had privacy defaults that assumed a more technically savvy customer pool than was actually present, and that even experienced users could make mistakes that enabled file exposure.

In this way, poorly designed file-sharing services could be thought of as a type of honeypot—luring users in to obtain desirable multimedia content, while providing access for others to extract their private information. Access that potentially includes law enforcement officers sniffing around for illegal downloads. Although KaZaA is not believed to have intentionally provided this backdoor service, they also have not instituted remedies as quickly as some might have preferred.

These security issues aside, perhaps file sharing is not so bad as it seems for the recording industry from a marketing standpoint, as was demonstrated in a study by Oberholzer and Strumpf [6]. Since legal arguments have tended to focus on damage to the industry via lost sales due to increased downloading of media files, as supported by industry-funded research, this independent study attempted to see if there was any empirical basis to these claims. Somewhat surprisingly, even though there were over a billion downloads worldwide each week of music files alone, and despite the dip of recorded music CDs shipped in the U.S. by 15% between 2000 and 2002, causality was not able to be established. Using their most pessimistic metric, it appeared that more than 5,000 downloads of a particular item were necessary in order to displace a single sale. The data further revealed that "high-selling albums actually benefit from file sharing." This points to other factors (such as macroeconomics, demographics, changes in recording format, and listening equipment) probably contributing at a higher rate to the decline in sales.

Similarly, the custom of delaying home media (such as DVD) releases to encourage movie theater revenues may be a shot in the

film industry's own foot, as this practice seems to also encourage bootlegging. According to the Motion Picture Association of America (MPAA), between May 2002 and March 2003, at least 28 films became available in the U.S. prior to their actual release dates, largely due to illicit camcorder recordings at early screenings [2]. The MPAA estimates losses at three to four billion dollars annually, with approximately 400,000 illegal movie downloads per day.

But industry analyst Douglas Dixon (see www.manifest-tech.com) has an alternate view—that "theatrical releases really can be considered an extended advertisement for the DVDs, where the real money is made. And the profusion of reissues and special releases of the same title (like the director's cut and platinum editions) suggests there still is plenty of profit from retail sales, even with the existence of bootlegs and file sharing." So it may be prime time for the MPAA to smell the popcorn and work on a way to provide legitimate film download services the way Apple (and others) are doing with music.

Indeed, the courts may be paving the way to a kinder, gentler view of file sharing, with the August 2004 9th U.S. Circuit ruling in a case filed by the MPAA. As Judge Sidney R. Thomas explained, "in the context of this case, the software design is of great import." Defendants Grokster Ltd. and StreamCast Networks Inc. prevailed in part because these peer-to-peer services (unlike the earlier Napster configuration) did not have central servers directing users to copyrighted files. Addressing the loss of revenue issue, the decision also stated "it is prudent for courts to exercise caution before restructuring liability theories for the purpose of addressing specific market abuses, despite their apparent present magnitude."

But such debates are hardly new, as evidenced by John Philip Sousa's eloquent 1906 essay "The Menace of Mechanical Music" (see www27.brinkster.com/phonozoic/menace.htm): "for the life of me I am puzzled to know why the powerful corporations controlling these playing and talking machines are so totally blind to the moral and ethical questions involved. Could anything be more blamable, as a matter of principle, than to take an artist's composition, reproduce it a thousandfold on their machines, and deny him all participation in the large financial returns, by hiding back of the diaphanous pretense that in the guise of a disk or [player-piano paper] roll, his composition is not his property?" Sousa had been an active participant in the U.S. Congressional debate that year over the issue of copyright extensions to include recordings, arguing in favor of the rights of performers, even though he personally felt such embodiments were as "incongruous as canned salmon by a trout brook." Certainly, nearly a century later, he likely would have disagreed with Thomas' *Grokster* decision.

Some industry supporters have suggested that perhaps the future of multimedia security may ultimately be found in Internet Protocol (IP) set-top boxes. According to Microsoft product managers Bill Wittress and Olivier Fontana: "An IP set-top box is a dedicated computing device that serves as an interface between a television set and a broadband network. In addition to decoding and rendering broadcast TV signals, an IP set-top box can provide functionality that includes video-on-demand (VOD), electronic program guide (EPG), digital rights management (DRM), and a variety of interactive and multimedia services" [10]. If a cellular telephone can handle email, take photos, record audio and video clips, and serve as a personal desk accessory, a television should certainly be able to do that plus much, much more. Basically, when one can view a *Wizard of Oz* download, rip an MP3 file of Judy Garland singing "Somewhere Over the Rainbow," and then use it as a ring tone, all without infringing on anyone's copyrights (especially

if it's before the performance goes into the public domain), we'll know that KaZaA (or its successor) is not in Kansas anymore, at least metaphorically, where media content, delivery, and owner compensation are concerned.

The purview of this technology will likely be tied to the development of advanced multimedia codecs (coders/decoders). But, to date, these devices and even the standards under which they are designed and operate (such as the MPEG formats) have largely been deemed proprietary, and patent-pool issues have slowed new format introductions.

Systems that embody digital rights management protocols have also come under fire by the academic community, which has long enjoyed "fair use" exemptions from some copyright restrictions. Princeton's Edward Felten (see www.freedom-to-tinker.com), who claims copyright on his Web materials through the year 2113, and Cambridge's Ross Anderson (see www.cl.cam.ac.uk/users/rja14/), who has released some of his Web documents under the GNU Free Licensing program, are among the most strident opponents of technologies, platforms, and regulatory controls that inherently limit the access and use of media files. As DRM is debated, these battles are being reflected in the marketplace with a restricted set of consumer choices, such as the current dearth of high-definition digital audio and video recorders, especially those that can easily swap formats.

Perhaps greater hope may be found through the cooperation of industry and academia, as seen in the recently released testimony to the U.S. House Judiciary Committee that was co-authored by RIAA's President Cary Sherman and Penn State's President Graham Spanier [9]. In their testimony, they recognized that the demand for downloads has reached the point where bandwidth performance and reliability problems are now overwhelming the abilities of many schools to readily conduct their online instructional activities. Their joint effort in formulating, implementing, and proliferating novel approaches to educate students about copyright infringement, file sharing, and digital media issues, in conjunction with the development of services that permit and encourage legal file distribution, is a commendable large step in the right direction.

Certainly the challenge will be for security technology to keep pace with the kaleidoscopic evolution of multimedia content and services. In art, as in science, creativity is paramount. Hopefully, new paradigms for protecting ownership rights and revenue will continue to emerge through cooperative approaches that view the spectrum of protective mechanisms from different sides of the prism of possibilities. C

References

1. Bender, W., Gruhl, D., Morimoto, N., and Lu, A. Techniques for data hiding. *IBM Systems Journal 35*, 3&4 (1996); www.research.ibm.com/journal/sj/mit/sectiona/bender.html.
2. Byers, S., Cranor, L., Korman, D., McDaniel, P., and Cronin, E. Analysis of security vulnerabilities in the movie production and distribution process. In *Proceedings of the 2003 ACM Workshop on Digital Rights Management* (Washington, D.C., Oct. 27, 2003).
3. Delp, E.J. Multimedia security research at Purdue University; dynamo.ecn.purdue.edu/~ace/water2/digwmk.html.
4. Electronic Frontier Foundation. Unintended consequences: Five years under the DMCA, version 3, (Sept. 24, 2003); www.eff.org/IP/DMCA/unintended_consequences.pdf.
5. Good, N.S. and Krekelberg, A. Usability and privacy: A study of Kazaa P2P file-sharing. In *Proceedings of the Conference on Human Factors in Computing Systems*, ACM SIGCHI, (Apr. 2003).
6. Oberholzer, F. and Strumpf, K. The effect of file sharing on record sales: An empirical analysis. In *Proceedings of the American Economic Association Annual Meeting* (San Diego, CA, Jan. 2004; www.unc.edu/~cigar/papers/FileSharing_March2004.pdf.
7. One Hundred Fifth Congress of the United States of America. *Digital Millennium Copyright Act*. Public Law 105-304, October 28, 1998.
8. Samuelson, P. Why the anticircumvention regulations need revision. *Commun. ACM 42*, 9 (Sept. 1999).
9. Sherman, C.H. and Spanier, G.B. A report to the subcommittees on courts, the Internet, and intellectual property. House Judiciary Committee testimony (Aug. 23, 2004); live.psu.edu/index.php?sec=vs&stor7776&pf=1.
10. Wittress, B. and Fontana, O. *Internet Protocol (IP) set-top boxes.* Microsoft whitepaper (Sept. 9, 2003); www.windowsfordevices.com/articles/AT7000949259.html.

Rebecca T. Mercuri (mercuri@acm.org) is a Radcliffe Institute Fellow at Harvard University, where she is conducting research on transparency and trust in computational systems.