

Scott A. Craver

Assistant professor,
Electrical and Computer Engineering
Binghamton University
Binghamton, NY 13902-6000
phone: 607.777.7238
scraver@binghamton.edu

Research Statement

My research is in the general area of Information Security, with a speciality in information hiding, watermarking, and attacks on both. I have a long history of research in the area, and since becoming an assistant professor I have secured nearly 1.5 million dollars in external funding---including a funded award secured in my first year, as well as the prestigious Presidential Early Career Award for Science and Engineering (PECASE,) awarded by the White House to 100 recipients in 2009.

Research History: My research career began with the discovery of **ambiguity attacks** on watermarking systems, while a mathematics student at Northern Illinois University. The premise is simple: if a random watermark is added to an image to establish its ownership, and it is extracted by comparison to the original, unwatermarked image, an attacker can falsify evidence of ownership by removing a watermark from someone else's image, declaring the doctored image to be the original. In other versions of the attack a watermark is chosen by inspection from the energy of a host signal.

When this idea became a submitted paper, I was tasked with reviewing another paper, on a steganography protocol for key exchange in a censored environment. The protocol had one missing piece, and I suggested in the review that a robust hashing algorithm can be used as an inconspicuous robust broadcast channel. The reviewers recommended that this idea be published separately; the result was the development of **supraliminal channels**, robust broadcast channels suitable for steganographic environments.

Upon entry to the Electrical Engineering Ph.D. program at Princeton University, I developed several methods for **zero-knowledge watermark detection**. These are techniques for proving the presence of a signal without providing information on the signal's embedding or its structure that would allow its removal. One notable property of the solution is that ambiguity attacks are part of the protocol. With the power to counterfeit watermarks, we can conceal a real watermark among a thousand false ones.

One of our most notable research achievements was **breaking SDMI**. The Recording Industry Association of America had long planned to embed robust watermarks in digital music, to act as control signals prohibiting play on computers and in portable music devices. This culminated in an overall architecture with several competing watermark technologies. In the fall of 2000, the RIAA and SDMI, the Secure Digital Music Initiative, instigated a public challenge to attack their candidate watermark systems. Of the four candidates, the one that fared the best would presumably be favored for use in their system. Our team from Princeton University, Rice University, and Xerox broke all the candidates, and in the process I was able to substantially

reverse-engineer one of the watermarks' internal structure. Our team wrote a paper outlining both how we broke the watermarks and how they worked; the level of detail caused the RIAA to send threatening letters to us, to our universities and employers, to the program chair of the Information Hiding Workshop to which it was submitted, and to his employer (the US Navy.) We found ourselves fighting in court for the right to publish under the recently-passed Digital Millennium Copyright Act; the experience was very valuable, and taught me a great deal about the legal side of intellectual property disputes. This also produced new techniques for **detecting echo-hiding watermarks**.

Later research focused more on reverse-engineering. As an assistant professor at Binghamton, I and my students won the first round of another watermarking contest, the international **Break Our Watermarking System (BOWS) contest**. Three images had an unknown watermark, with an online detector; the challenge was to render the mark undetectable with noise below 30.00 dB PSNR. Our attacks employed an unusual and backwards approach: rather than attempting to remove the watermark with minimum noise, we attempted to preserve the watermark with maximum noise. This leaks a great deal of information about the underlying algorithm, and the trick led to several general attacks such as the **noise calipers**, or "snakes on a cone" technique for sensitivity analysis.

My current research is in the development of new supraliminal channels, and in the development of detectors with minimum information leakage under oracle attacks. Both are areas ripe for future research: it has long been known that detectors can be beaten by oracle attacks, but over ten years after the first sensitivity attacks were published, detection regions remain as rudimentary and as vulnerable as always. There have yet to be any results characterizing the information leakage of a detection algorithm, or developing detectors with maximal guessing entropy.

Several other results lie outside this general arc of ideas: I have also developed search algorithms for image and video databases based on space-filling curves, and we are currently working on an implementation of self-destructing unprintable unforwardable email. The latter uses a hardware solution that can identify and decrypt selected regions of a screen area by observing and monitoring the DVI signal between computer and monitor.

Research Philosophy: I have always been a proponent of the doctrine of random research, the idea that one should read research papers on random subjects both inside and outside one's general area. This provides researchers with an unusual toolbox of techniques with unexpected applicability to their own field. Many of my milestones in research result from an unexpected synergy of random ideas. Our development of PHOOSballs for high-dimensional search stemmed from an earlier investigation of space-filling curves and point-set topology, an investigation untainted by any desire for application. Our development of ambiguity attacks on watermarking systems was a result of studying logical paradoxes in legal frameworks. I was fortunate to be reading Douglas Hofstadter's column on the game *Nomic* an hour before I was supposed to explain the NEC watermarking algorithm in Nasir Memon's group meeting, and that is basically where the idea came from.

Future Directions: I have just won the Presidential Early Career Award for Science and Engineering (PECASE) to continue my work in reverse-engineering of detectors. My current goal is to extend the concept of guessing entropy to watermark and other detectors. The difficulty with this problem is that unlike passwords or keys, detector input-output pairs can leak information about the underlying algorithm that is difficult to quantify. This is due to the fact that the detection region cannot be completely discontinuous, but must satisfy error bounds and robustness requirements. Even establishing a bound on “reversing entropy” would have broad application beyond watermarking to any adversarial signal detection scenario.