Database for steganalysis of forensic evidence

February 13 2018

Jennifer Newman (bottom right) is leading the effort to build a photo database for steganalysis research. Credit: Christopher Gannon

There is nothing striking or remarkable about the tens of thousands of pictures an Iowa State University research team has spent the past 18

months collecting for a new database.

Jennifer Newman, an associate professor of math leading the team on behalf of the Center for Statistics and Applications in Forensic Evidence based at Iowa State, says they are not interested in aesthetics or photos worthy of framing. In fact, the dimly lit and overexposed pictures make the database valuable for researchers interested in steganalysis.

"The pictures in the database are of ordinary scenes, and we purposely collected poor quality pictures in different camera-setting environments," Newman said. "We've found the settings can influence how well the steganalysis detection works."

Steganalysis is of growing interest to forensic investigators analyzing digital photos, which might contain hidden content related to a criminal case. Steganography – a covert method of sending messages without detection—dates back to ancient Greece. It has evolved with technology, Newman said. Now, anyone with a simple algorithm – available through a variety of free apps – can manipulate the code of a picture or any digital file, such as a Word document or PDF, to conceal a message.

Unlike encryption, which does not mask the presence of a message, steganography "hides the message in plain sight" so someone looking at the photo would not suspect it contains hidden text, Newman said. Helping forensic investigators identify and decode these messages is the ultimate goal for researchers, but first they must establish a foundation for steganalysis, or steg detection.

That is where Newman and colleagues Yong Guan, an associate professor of electrical and computer engineering at ISU; and Min Wu, at the University of Maryland, along with a team of ISU graduate and undergraduate students are focusing their efforts. Armed with nearly two dozen different cell phone models, the researchers have taken more than

50,000 original images for the StegoDB database, expected to launch in late March. Now researchers around the world have a catalog of cell phone images to use for training and testing steganalysis detection tools.

"If you don't have the database to test images, where do you even start? We hope other researchers can look at the data we gather and develop a quantitative statistical analysis that is useful in a court of law," Newman said.

Technology creates a need

Existing public steganalysis databases only include images from still cameras. With the widespread use of cell phone cameras, ISU researchers recognized the need for data specific to cell phone images to help in the development and testing of steganalysis detection tools. Newman says because camera settings, which are unique to each phone brand and model, can affect a tool's reliability, it is necessary to test for inconsistencies.

Considering the multitude of phone brands and models, that is a daunting task, but Newman says the team is up for the challenge. One of CSAFE's goals is to introduce a statistical framework for evidence in criminal cases that does not currently exist, she said. According to a 2009 National Academy of Sciences report, only DNA had reliable, quantifiable error rates associated with the conclusions. Newman says that means there is no statistical validity or scientific support for evidence such as fingerprints, bite marks, bullets or stego images.

"This is valuable evidence that can help rule out suspects in a case, but you need to identify the strength or lack of strength of the evidence," Newman said. "As scientists, our work is focused on quantifying what is known and what is unknown. We are setting boundaries of what is statistically known and statistically uncertain for these forensic arenas."

Determining these statistical boundaries then allows an expert witness to testify to a degree of certainty that a photo collected as evidence came from a specific phone model and the likelihood that it contains hidden content, she said. It will take years of research to reach this point, but Newman says the work may benefit investigators in child pornography and intellectual property theft cases. As an example, she explained more and more companies are interested in technology to scan outgoing electronic files to detect if an employee is covertly sending sensitive information to someone outside of the company.

New area for exploration

Newman says steganalysis is a relatively young field with several unexplored applications. This creates opportunities for researchers to explore and contribute to the framework. She and her colleagues presented some of their initial findings at conferences on computer forensics and multimedia images. The work would not be possible without the team's combined expertise in math, statistics, computer science and engineering and physics.

The team will continue collecting photos taken with different phone models to expand the database. Newman says for it to be effective, researchers need a wide variety of images for greater applicability.

Provided by Iowa State University