C17 Stego App Database (DB) and the Prevalence of Mobile Steganography

Jennifer Newman, PhD, Iowa State University, Ames, IA 50011; Li Lin, BS, Iowa State University, Ames, IA 50011; Wenhao Chen, BS, Iowa State University, Ames, IA 50011; Yong Guan, PhD, Iowa State University, Ames, IA 50011; Stephanie Reinders, BA, Iowa State University, Ames, IA 50011; Min Wu, PhD, University of Maryland, College Park, MD 20742*

Learning Overview: After attending this presentation, attendees will be aware of the first mobile stego app database for image forensics and the current lack of software to detect the prevalence of stego use on mobile devices.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by providing forensic criteria under which the dataset was designed and collected and will demonstrate that such a carefully curated and provenanced dataset can provide software developers with image data to benchmark tools they create on currently available stego data from mobile apps.

Tools for digital image analysis are advancing toward mainstream use in practical forensics. It is well known that reference datasets are necessary to help researchers develop and benchmark solutions for the variety and scale of data encountered by forensic practitioners.[1] For example, steganography has no known reference dataset to benchmark software that detects stego images occurring "in the wild." While steganography is starting to consider more realistic scenarios, there is currently no known application that can test for steganography content originating from mobile device photographs created from stego apps on smart phones. Thus, the prevalence of steganography on mobile devices cannot be determined. This presentation introduces the first database consisting of mobile device photographs and stego images produced from stego apps on the phones, including a rich set of side information. The database contains images available to software developers to create steg detection programs—or steganalysis algorithms—that are more effective in detecting stego images produced by mobile apps. Once these tools are created and tested, it may be possible to begin investigating the prevalence of steganography use on mobile devices.

StegoAppDB, a steganography apps forensics image database, contains more than 810,000 innocent and stego images from ten different cell phone models (24 distinct devices) with detailed provenanced data including a wide range of International Organization of Standardization (ISO) speed and exposure settings, Exchangeable Image File Format (EXIF) data, stego Android™ Package Kits (APKs), message information, embedding rate, and other information. The data was collected according to a set of forensic criteria (authentication, representation, evaluation) and was free of copyright or privacy issues. The database is currently available to the public.

The acquisition procedure for original images will be discussed, including the camera app Cameraw, in both Android™ and iOS® systems.[2] The application allows researchers to gather multiple images per scene and saves each image simultaneously in both Digital Negative (DNG) and high-quality Joint Photographic Experts Group (JPEG) formats. From original images, stego images are created using five Android™ stego applications and one iPhone® stego application. Included are stego apps that write signatures and others that use random embedding methodologies. The extensive reverse engineering will be discussed, including source code modification and binary code instrumentation, the large amounts of innocent and stego image data that was generated for the benchmarking database. StegoAppDB provides cover-stego image pairs for each stego image, including applications that change input image dimensions, so that machine learning algorithms can be deployed for steganography detection.

Descriptive statistics of the database will be presented in addition to results of several experiments to substantiate the database's investigatory nature. Applying three software programs to stego images from mobile apps—Stego Hunt, DC3 StegDetect, and Provos StegDetect—it will be shown that they are not adequate to detect stego images from modern stego apps.[3-5] The user interface, queries, and the download process of the website will be demonstrated. New data will continue to be added to the database on a monthly basis with the retained devices. While designed for steganography, possible uses of StegoAppDB to other digital image forensic topics will also be discussed.

Reference(s):

1. National Research Council. Strengthening Forensic Science in the United States: *A Path Forward*. National Academies Press, 2009.
2. Wenhao Chen. *Cameraw, an Android™ Camera App for Digital Image Forensics*. Technical Report, Center for Statistics and Applications in Forensic Evidence (CSAFE), Iowa State University, 2017.
3. WetStone Technologies, Inc. StegoHunt. https://www.wetstonetech.com/products/stegohunt/.
4. Personal correspondence with W. Eber, 2017.
5. N. Provos and P. Honeyman. Detecting Steganographic Content on the Internet. In: *Network and Distributed System Security Symposium*. San Diego, California, Feb. 2002. Internet Society.

Steganography, Forensic Image Database, Steganalysis