C24 Using Rapid Differential Forensics Algorithm to Speed Transmission of Large Files Around the World

Mark D. Guido, MS, The MITRE Corporation, Mclean, VA 22102; Rob H. Schmicker, BS, The MITRE Corporation, McLean, MD 22102; Brandon Adler, MS, The MITRE Corporation, Mclean, VA 22102; Tristan Fletcher, BS, The MITRE Corporation, Mclean, VA 22102*

Learning Overview: The goal of this presentation is to describe an extension to the previously presented Rapid Differential Forensics Imaging algorithm to address a common issue affecting forensic laboratories and practitioners around the world: their ability to efficiently obtain forensic acquisitions and co-locate them with their computing power.

Impact on the Forensic Science Community: This presentation will impact the forensic science community by demonstrating a reference architecture that both saves significant amounts of time and money comparatively to how major laboratories are performing the practice today and also reduces the time to analytical results.

In 2016, an automated differential forensic imaging technique and algorithm was introduced that uses baseline datasets and hash comparisons to limit the amount of data sent from a mobile device to an acquisition endpoint. It produced forensically validated bit-for-bit copies of device storage in significantly reduced amounts of time compared to commercial products.

As is common practice today, forensic laboratories centralize their images and forensic computing power to minimize the time it takes to process the image and perform forensic analysis. The processes used, surrounding case details, and employed forensic analysis techniques may be considered sensitive to an investigation, which yields a general reluctance to decentralize or use cloud computing resources. Major laboratories often opt to physically mail large forensic images to their computing destination, causing significant delays in returning actionable results to assist an investigator or try to prioritize transmitted data to send, essentially taking a guess as to where relevant data is located on the hard drive to prioritize those areas to send, causing similar problems that they were trying to avoid during the acquisition phase. The United States military may similarly attempt to relocate forensic images to a centralized environment to return actionable, timely results to soldiers that may need it. In some environments, the soldiers may only have a low bandwidth (potentially unidirectional) transmission medium (e.g., satellite transmission) available to them to relocate their acquired data.

By applying an automated differential forensic imaging algorithm, the amount of time needed to perform a transmission from a sender to an endpoint is reduced. It eliminates sending duplicate or common parts of image files, including seen-before operating system file parts and empty unallocated space, as well as removing random, at-rest encryption from empty portions of the transmitted disk image. For unidirectional communications, the addition of forward error correction provides assurances that when data is lost or mangled, we can be assured we can still reconstruct it using the additionally transmitted parity information.

Further, when implemented on a cloud infrastructure such as Amazon® Web Services (AWS), it significantly reduces the time and cost it takes to send large files from one region to another and closes the gap between acquisition and returning analytical results to investigators. Although very fast, this inter-region computing can be expensive if utilizing AWS-provided communication mechanisms. By utilizing an algorithm to only send the data required by the algorithm to the destination, a savings on resources leads to reductions in time and cost savings.

Differential, Rapid Transmission, Cloud

*Presenting Author