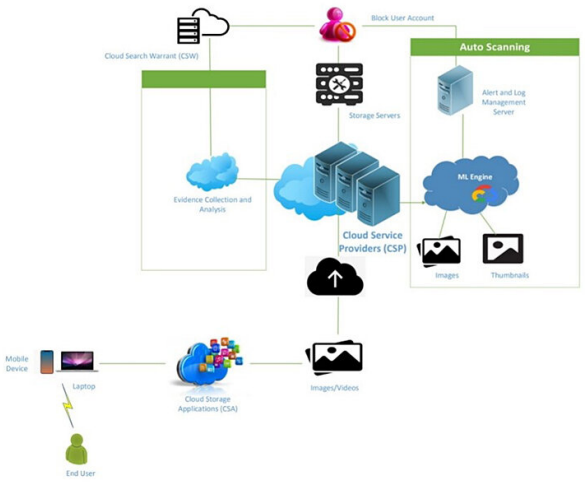
Machine learning advances new tool to fight cybercrime in the cloud

6 November 2019, by Chris Adam



Credit: Purdue University

Increased adoption of cloud applications, such as Dropbox and Google Drive, by private users has increased concern about use of cloud information for cybercrimes such as child exploitation, illegal drug trafficking and illegal firearm transactions.

Researchers at Purdue University have developed a cloud forensic model using machine learning to collect digital evidence related to illegal activities on cloud storage applications.

"It is crucial to detect illegal cloud activities in motion," said Fahad Salamh, a Ph.D. student in the Purdue Polytechnic Institute, who helped create the system. "Our technology identifies and analyzes in real time incidents related to these cybercrimes through transactions uploaded to cloud storage applications."

Salamh worked on the technology with Marcus Rogers and Umit Karabiyik, professors in Polytechnic who specialize in computer and information technology.

The Purdue system deploys deep learning models to classify child exploitation, illegal drug trafficking and illegal firearms transactions uploaded to cloud storage applications and report illegal activities via a forensic evidence collection system.

The process begins when a cloud storage application user uploads a media file, either image or video. The pre-trained machine learning models scan both images and thumbnails to look for signs of cybercrimes.

Through identifying and analyzing these incidents using machine learning, cloud service providers can collect alerted logs, block the associated accounts and report them to law enforcement based on a cloud search warrant request.

"It is important to automate the process of digital forensic and incident response in order to cope with advanced technology and sophisticated hiding techniques and to reduce the mass storage of digital evidence on cases involving cloud storage applications," Salamh said. "Cloud environments challenge investigators in identifying the ownership of uploaded media files because of their network architecture and data processing."

The Purdue team tested more than 1,500 images, and the model accurately classified an image roughly 96% of the time.

Provided by Purdue University