DeCrypto Pro: Deep Learning Based Cryptomining Malware Detection Using Performance Counters

Ganapathy Mani, Vikram Pasumarti, Bharat Bhargava, Justin King

2020 IEEE ACSOS

Problem Addressed in this Paper:

APTs are a class of cyberattacks where the attackers will stay in the system for extended periods of time to observe and execute their malicious activities. They can execute those actions step-by-step by waiting for a long time between each step, making it difficult for detection by antivirus or learning and prediction models.

Cryptojacking is a term defined where collaborative attackers run cryptocurrency mineros on victims system without their authorization and utilize their CPUs computational power to mine cryptocurrencies. Cryptomining is the process of generating wealth by creating and verifying new blockchain-based cryptocurrencies.

Each block in a cryptocurrency blockchain has three elements:

- Data (transactions and their relevant information such as sender, receiver and reward)
- Unique Hash Signature
- Unique Hash Signature of previous Block.

When there is a new transaction, it is bundled into a block, everyone in the network must verify the block's information before it is linked as another block to the major blockchain. In order for the new block to be linked to the major blockchain, the right hash value must be calculated. Finding the right hash given a large set of random numbers is arbitrary and a brute force mechanism. Solving this blockchain puzzle is a computationally intensive process that requires extensive computing architecture and powerful machines such as GPUs, Application-Specific Integrated Circuits (ASICs), and advance FPGAs for increasing the probability of finding the right PoW. As the

cryptocurrency blockchain grows longer, the difficulty of finding a new block goes higher. Hence, there is a high demand for more and more computing power.

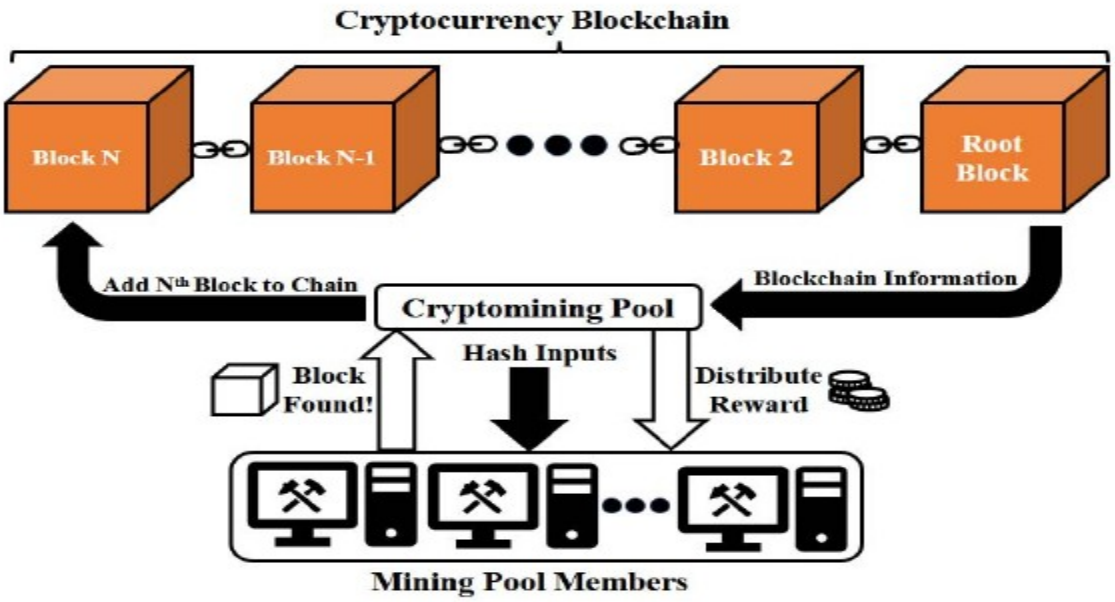

Figure 1: Cryptomining Workflow

Fig. 1 shows the workflow of mining cryptocurrencies using cryptomining pools. The cryptomining pools will have the whole distributed blockchain ledger and act as intermediaries with pool members. A mining pool service issues the hash inputs that are needed to be computed for finding the best PoW to all the systems that have signed up with the mining pool service. If an appropriate block is found, it is added to the major blockchain of the cryptocurrency. In return for providing the computing resources, mining users get the equally distributed reward and pooling service gets its service fee. Pooling services are easy to access and this creates a significant potential for deploying these miners in a large number of systems. Attackers can deploy these mining services through native applications or browser-based web applications. where they can mine any particular cryptocurrency in the background without authorization from the victim.

They propose a robust solution for cryptojacking detection and prediction by operating context profiling—a type of behavior profiling where applications are classified based on their holistic effect of their operating context. Through experiments, we show that learning models of DeCrypto Pro provide high accuracy with significantly less false positive and false negative rates.

Motivation for the Problem:

CryptoCurrencies are digital assets designed for secure and transparent transactions facilitated by Blockchain public ledgers. Cryptojacking uses computing power of the victim for mining of cryptocurrencies. Symantec detected 8 million cryptojacking attacks in just three months from Dec 2017 to Feb 2018. The demed increase in use of cryptocurrencies such types of attack needed to be handled as victim have no idea where it's CPU exhausted.

Motivation for the Solution:

Many detection technique use cryptomining's high CPU usage and overheating of the system as triggers to alert the antivirus of potential mining activity. But cryptominers can evade detection through the following techniques and their properties.

- They can set the threshold for CPU usage therefore have minute control over execution and operating environment parameters.
- Employ drive-by mining technique where they stay on the system for a very short time and use the CPU, before moving on to the next victim. They will come back again and repeat the same process.
- Most cryptomining algorithms such as CryptoNight, only perform cryptographic calculations, bitwise operations, and encryption operations. They neither change nor disrupt the system's behavior in any significant manner. This makes detection difficult since benign applications such as compression or encoding applications perform similar cryptographic, bitwise, and encryption operations.

Theory:

Performance counters are small programs that count, monitor, and measure events in the system. They provide information on how a particular operating system or a service, an application or a driver is performing. Windows provides a unique set of performance counters that contain 68 properties. These 68 counters can produce several hundred values depends on the number of processor cores in the CPU. In this research, each sample in the data set contains 245 values.

DeCrypto Pro provides a streamlined research approach for detecting as well as predicting evasive cryptomining actions. DeCrypto Pro uses resilient performance counters data for effective prediction with efficient model selection, which reduces the

computational resource usage. This will help mission-critical systems to focus on important tasks rather than spending resources on security operations. In order to make an efficient detection of cryptomining malware, we employ a model selection utility function that can select an optimal model given computational resources such as processor frequency and memory as well as accuracy and F1 scores of the model training.

Implementation:

Experiments were set up on 3 Windows machines with various configurations on processing frequency (2.40, 2.90, 2.30 GHz), memory size (16, 8, 8 GB), and number of processor cores (2, 4, 5). Since we aim to capture the system status signature of PoW algorithm such as CryptoNight's signature, we mainly focus on bitwise, cryptographic, and processor-specific encryption operations. Thus we consider compression software (7-Zip, SecureZip, PeaZip, WinRAR, WinZip, and Freemake) as our benign examples and cryptomining applications (XMRig, XMR-Stak, Coinhive, Computta, and GUIminer) as malicious example.

We used Windows PowerShell to obtain the performance counter data. There are three settings for collecting operating context performance counter samples for training. Compression software is for benign examples, cryptomining software running for malicious examples and both compression and cryptomining software running for malicious examples. Once the data is collected, we performed Min-Max feature scaling to normalize the data. We label each data row of performance counters data as 0 or 1 for multi-class classification. All of the null values are removed from the data set.

They implemented k-NN and RF using the Scikit-learn version 0.21.2 Python library. For model selection framework, we simulated random selection of models as well as selection based on utility function's output. Each model selection is performed with retraining of the models in random intervals with newly obtained data.