Eighth Annual Industrial Control System Security Workshop, December 6th, 2022

Device-Centric Ransomware Detection using Machine Learning-Based Memory Forensics for Smart Inverters

Authors: Alycia Jenkins, BoHyun Ahn, Alve Akash

Advisor: Prof. Taesic Kim

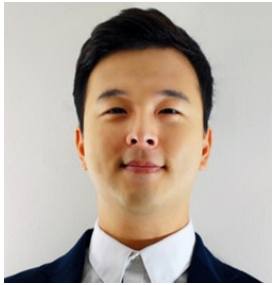
Cyber-Power and Energy System Laboratory

Texas A&M University-Kingsville

December 6, 2022

Alycia M Jenkins (First Author)

Master's Student in Computer Science

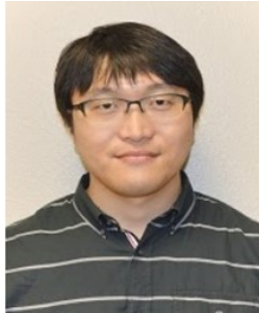


BoHyun Ahn

Ph.D. Student in Electrical Engineering

Alve Rahman Akash

Master's Student in Electrical Engineering



Dr. Taesic Kim, Ph.D. (Director)

Associate Professor and Lab Director

TEXAS A&M UNIVERSITY KINGSVILLE

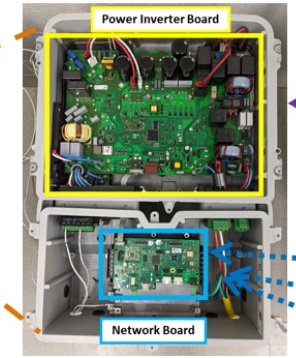
Outline

1. Introduction
2. Security Threat Modeling for Smart Inverters
    1. System Identification
    2. Attack Modeling
    3. Penetration Testing
3. Proposed Device-Level Memory Forensics for Smart Inverters
    1. Proposed Memory Forensics Method
    2. Binary Files Collection and Artifacts
    3. Deep Transfer Learning
4. Experimental Validation
5. Conclusion

Introduction

Smart inverter: Networked power electronics device that transfers generated power from distrbuted energy resources (DERs) to the power grid

A Smart Inverter-based PV Farm



A Smart Inverter Device

Smart inverters in a solar farm

PV systems

Wind energy systems

Energy Storage Systems

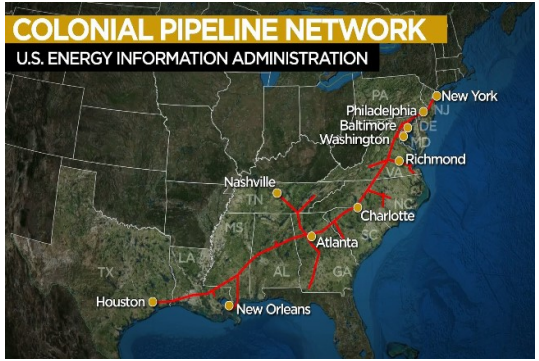
EV Charging Systems

DERs

" Numerous smart inverters will be installed, and smart inverters are becoming major devices in the future power grids (inverter-dominant power grid)!!"

Introduction

Ransomware: A malware to encrypt a target system's files making them unusable. Then the attackers demand a significant amount of ransom in exchange for decryptions



Colonial Pipeline Company (Paid $5 million in cryptocurrency)

"Recently, ransomware attacks have targeted industrial control systems (ICS) and increased about 500% from 2018 to 2020 (Dragos)."

Introduction

Ransomware targeting Power Grid

Renewable Energy Cyber Incidents

- 2014: **SolarWorld AG:** Chinese cyber espionage for economic advantage
- March 5, 2019: **sPower:** Denial-of-service attack
- Feb. – April 2020: **Azerbaijani wind turbines:** PoetRAT malware
- April 18, 2020: **EDP Renewables:** Ransomware
- June 2021: **Invenergy:** REvil ransomware
- August 2021: **ERG:** LockBit 2.0 ransomware
- Sep. 2021: **Swedish renewable manager**: LockBit 2.0 ransomware
- Nov 19, 2021: **Vestas:** LockBit 2.0 ransomware
- Feb. 24, 2022: **Enercon:** Russian state-sponsored SATCOM attack
- March 31, 2022: **Nordex Group:** Conti ransomware
- April 11, 2022: **Deutsche Windtechnik:** Ransomware
- April – June 2022: **South China Sea wind turbines:** Chinese ScanBox malware
- August 28, 2022: **GSA:** BlackCat ransomware

Credit: Idaho National Lab.

Ransomware Attack on Wind Farm IT System

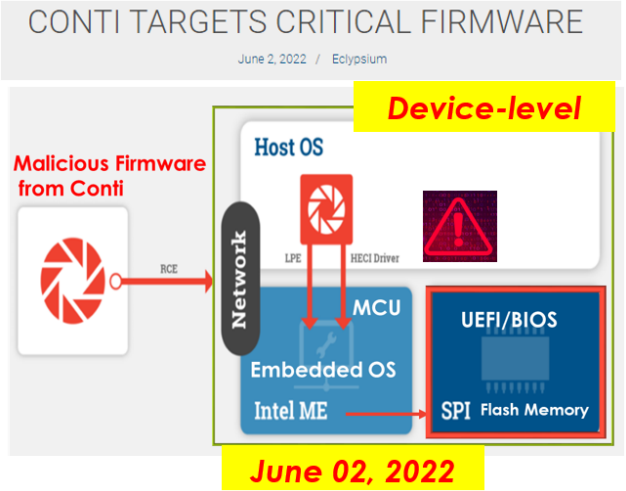
Ransomware Impacts in Power Grid:

- Leaking and selling of confidential data leading to more cyber-attacks in the future
- Damage to field device/system control processes and consequently the creation of public safety hazards
- Substantial disturbances to the power grid operation
- Harm in the form of economic loss

"Smart grid cybersecurity and engineering teams must prepare now to combat more frequent ransomware attacks!!"

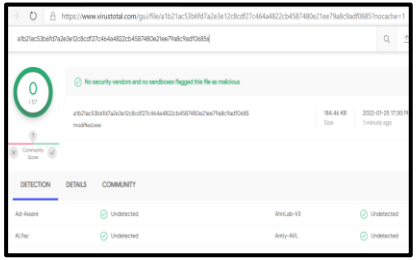
Introduction

Device-level (firmware) Ransomware Attack



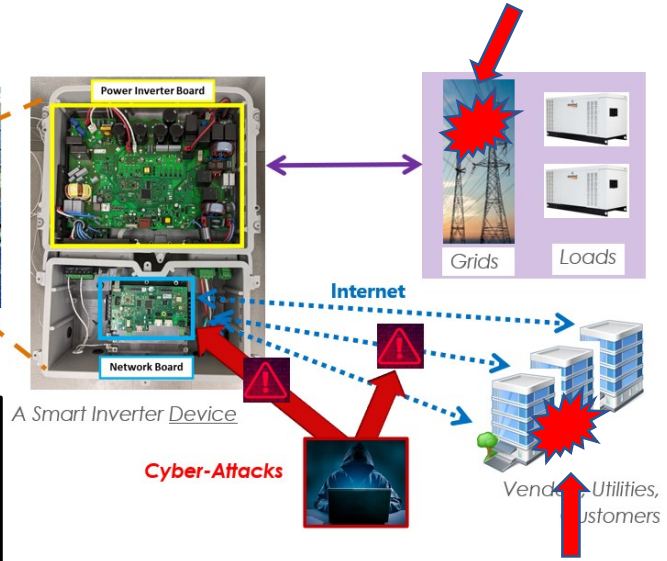
Conti ransomware targeting firmware

A Smart Inverter-based PV Farm



Detection failure of an obfuscated firmware malware

Coordinated smart inverters attack



A Smart Inverter Device

Smart inverter worm attack to the utility control server

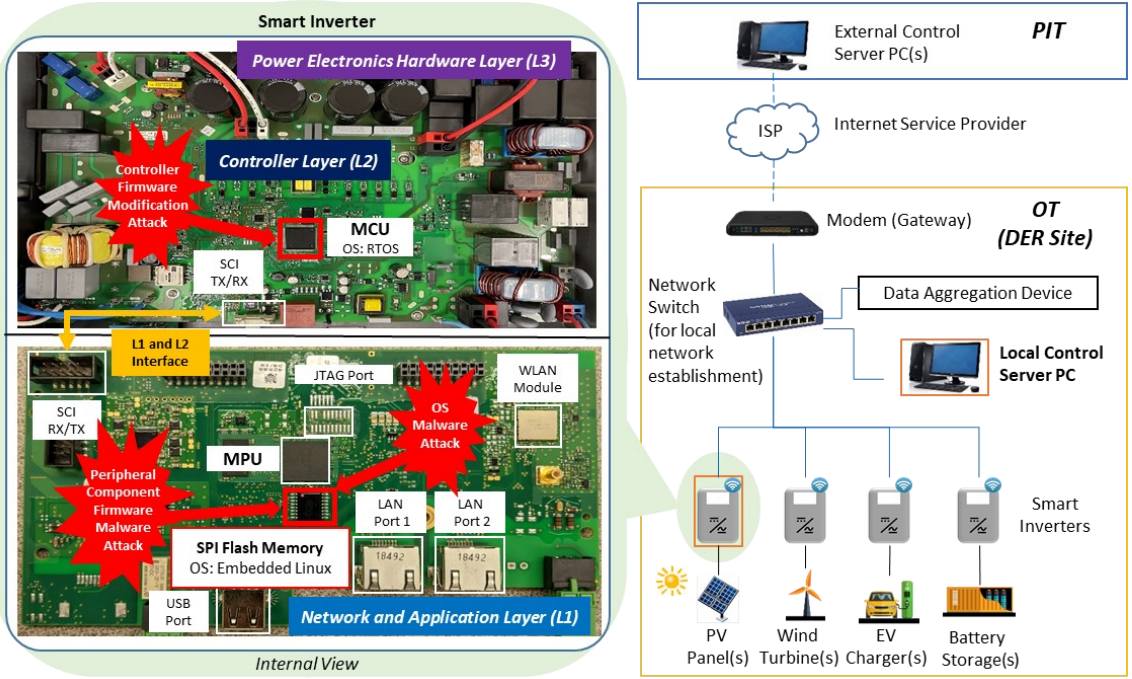
"Device-level ransomware attack (new attack vector) has been less studied."

Outline

1. Introduction
2. Security Threat Modeling for Smart Inverters
    1. System Identification
    2. Attack Modeling
    3. Penetration Testing
3. Proposed Device-Level Memory Forensics for Smart Inverters
4. Experimental Validation
5. Conclusion

Security Threat Modeling for Smart Inverters

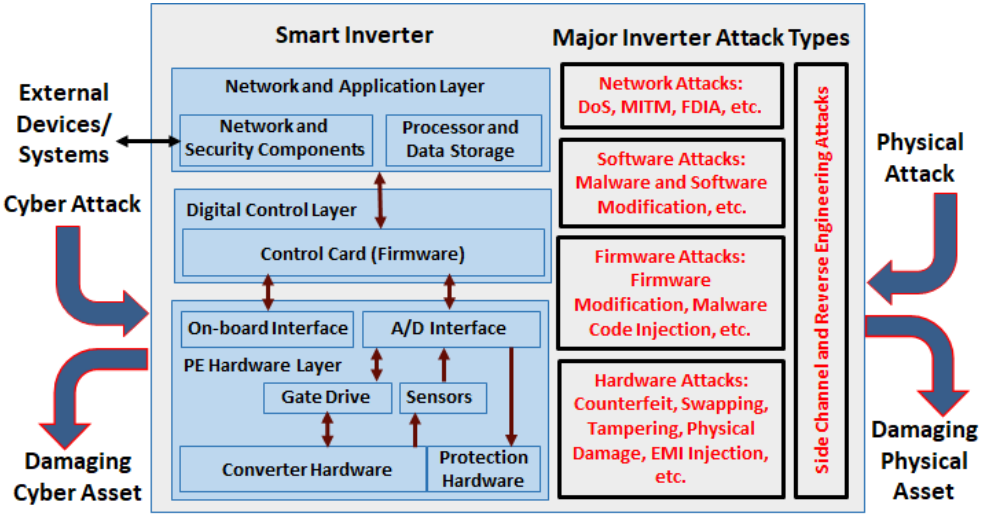
- **Commercial Smart Inverter**
  - Network and Application Layer (L1)
    - Microprocessor unit (MPU) and SPI flash memory
    - Embedded Linux OS and a bootloader for embedded systems (e.g., U-Boot) are installed in the flash memory
    - Bi-directional communication process through LAN ports and WLAN module
  - Controller Layer (L2)
    - Relatively lower computational resource with a microcontroller unit (MCU) for signal processing and power stage regulations
  - Power Electronics Hardware Layer (L3)
    - Power Electronics hardware, protection circuits, sensors, etc.



A commercial smart inverter system architecture

Security Threat Modeling for Smart Inverters

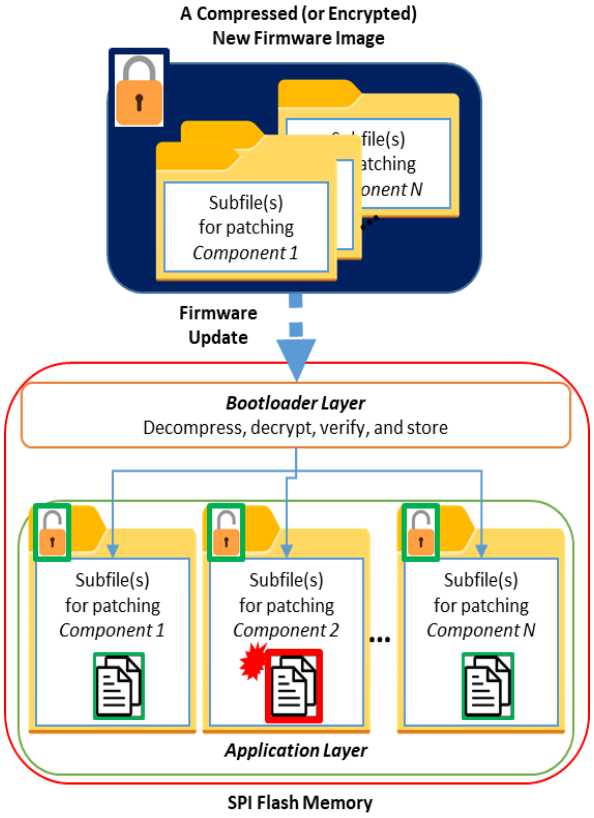
- **Security Sensitive Layer of a Smart Inverter**
  - Network and Application Layer (L1)
    - Network attacks
    - OS Malware Attack
    - Firmware Malware Attack
  - Controller Layer (L2)
    - Controller Firmware Modification Attack
  - Power Electronics Hardware Layer (L3)
    - Counterfeit and tampering
    - EMI injection
    - Supply chain attacks
- **Physical Memory Access**
  - Access SPI flash memory and tamer with firmware image during a system development stage, a supply chain process or maintenance process
  - JTAG debugging pins, UART debugging pins, flash memory chip pins
  - Memory access tools can connect to the proper pins to dump and flash a whole firmware image



Security sensitive layer of a smart inverter

Security Threat Modeling for Smart Inverters

- **Firmware Update**
  - Vendor sends a compressed (or encrypted) new firmware image
  - Device bootloader decompresses, decrypts, verifies, and stores the received image
  - Subfile(s) from the image patches corresponding target components
- **Firmware Update Attack**
  - Insider's can reactivate the Netcat function or the Telnet protocol allowing unintended external remote access
  - Firmware image include correct checksum data or code signing signature created by the vendor
  - Target device-tailored malware can be downloaded

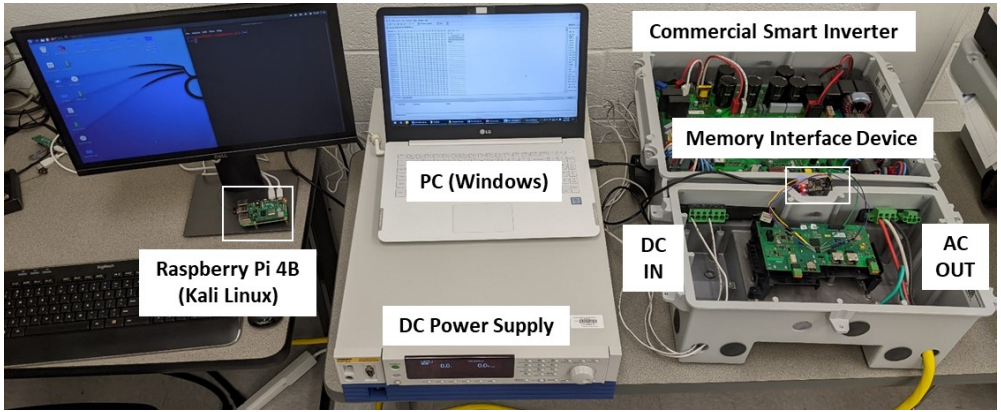


Malicious Firmware update process to SPI flash memory

Security Threat Modeling for Smart Inverters

Penetration Testing:
Once the target binary image stored in SPI flash memory is dumped by a memory access tool and decompressed into system files (e.g., squashfs), the smart inverter could be penetrated by the following techniques:

1. Firmware Image Tempering
2. CRC-32 Checksum Collision
3. Byte Order Obfuscation Bypass
4. Boot Header Information Tampering



Experimental Setup

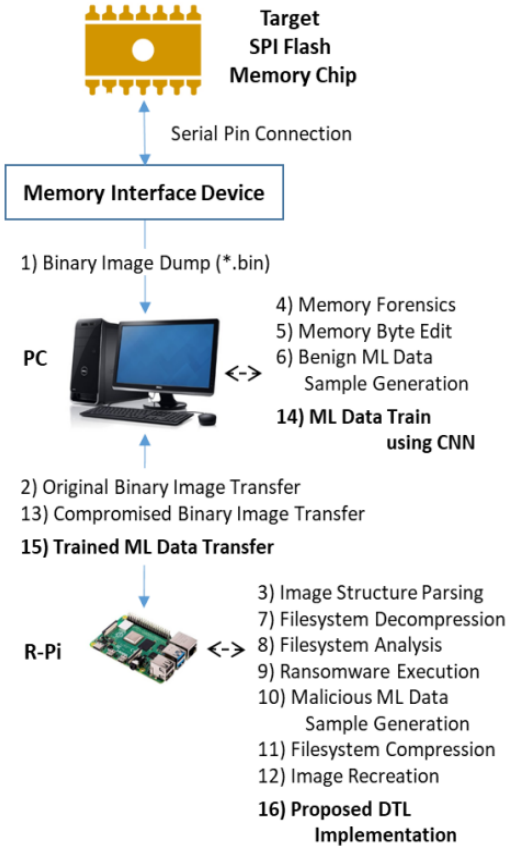
Outline

1. Introduction
2. Security Threat Modeling for Smart Inverters
3. Proposed Device-Level Memory Forensics for Smart Inverters
4. Experimental Validation
5. Conclusion

Proposed Memory Forensics for Smart Inverters

ML-based Memory Forensics:

1. Overall process of the proposed ML-based memory forensics
2. Binary File Collection and Artifact
3. Deep transfer learning



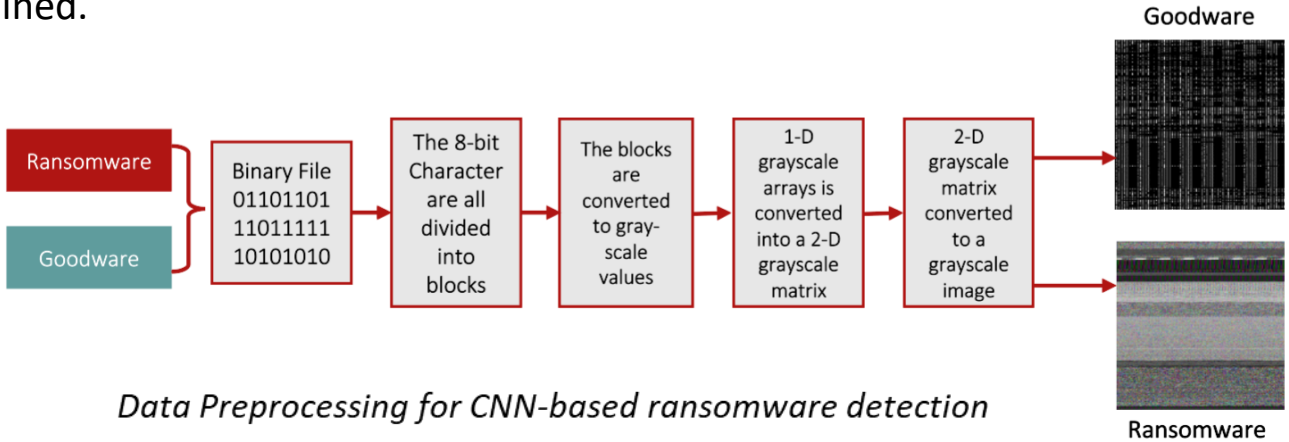
Overall processor of the proposed ML-based memory forensics method

Proposed Memory Forensics for Smart Inverters

Binary File Collection and Artifacts

- Benign binary samples: Benign binary samples are collected from a publicly available embedded Linux file system (e.g., SquashFS) that had ben extracted from network devices (e.g., routers). The smart inverter's file system extracted from the memory is also included in these samples.

- Ransomware binary samples: After having decryption processes of the file system collections, each file system is encrypted by a real-world Linux version of Conti ransomware. Then, they are compressed again to complete ransomware binary sample creations. As a result, one hundred non-repetitive samples are obtained.

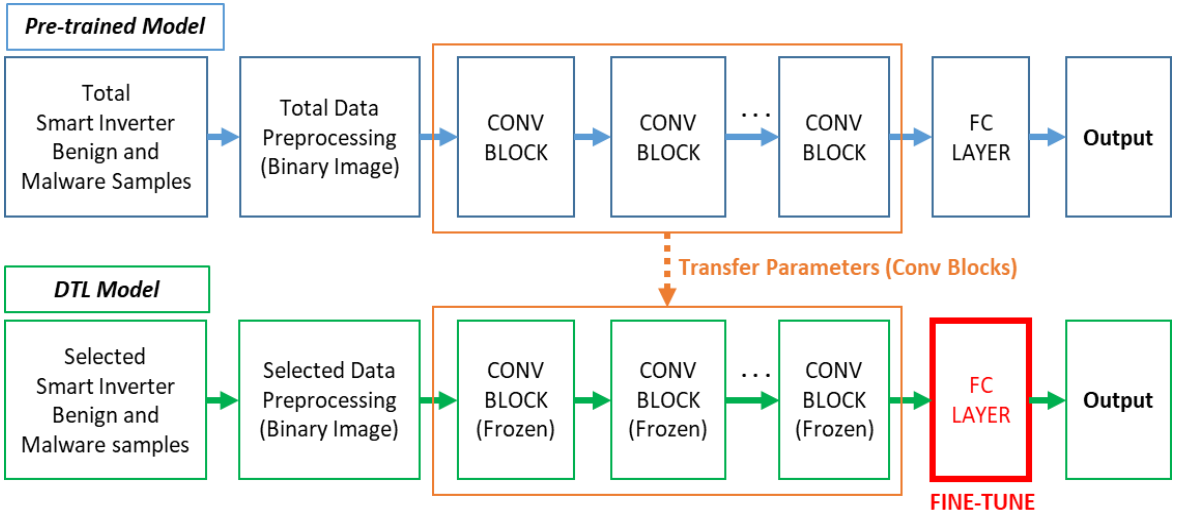


Data Preprocessing for CNN-based ransomware detection

Proposed Memory Forensics for Smart Inverters

Deep Transfer Learning

- A DTL algorithm is applied to create the smart inverter-centric ransomware detection algorithm by retraining a basis CNN-based ransomware detection model with the generated.
- The DTL method uses three convolutional 2D layers and three maxpooling 2D layers fitted over 640 parameters.



Deep transfer learning [14]

[14] S. Alvee, B. Ahn, S. Ahmad, K. Kim, T. Kim, and J. Zeng, "Device-centric firmware malware detection for smart inverters using deep transfer learning," in Proc. 2022 Figure 8: Experimental results of ransomware detection . 7 IEEE Design Methodologies for Power Electronics, Bath U.K., Sep. 1-2, 2022, pp. 1-5.

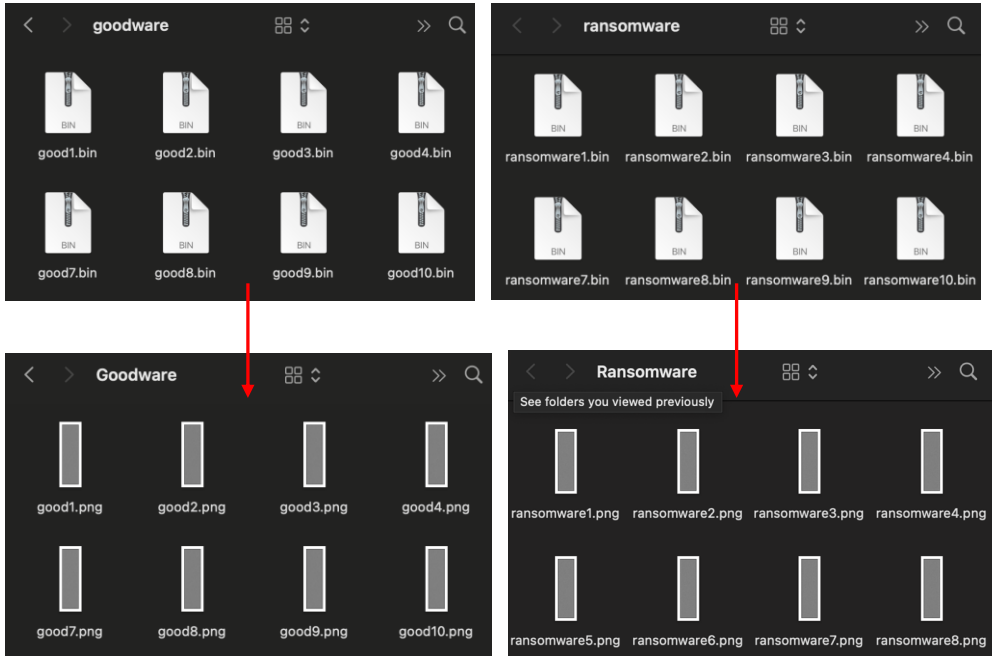
Proposed Memory Forensics for Smart Inverters

DTL Experimental Setup

- DTL model experiment was executed using a Macintosh Operating System with the Apple M1 SOC, 7-Core GPU, and 8 GB of RAM. A total of 100 benign firmware files (Goodware) and 100 ransomware are used for training (60%) and validation (40%).
- The proposed DTL model is implemented in an IoT device (R-Pi with 64-bit ARM OS, internal GPU (896 MB), and 8GB RAM) using the test image files

Data Preprocessing for DTL

- Extract binary data and convert them into image arrays. The RGB image array is then converted to training data. Since the transfer learning model was trained over grayscale training images, converted using the above-mentioned method, the RGB images were imported as grayscale images to better initiate the model.

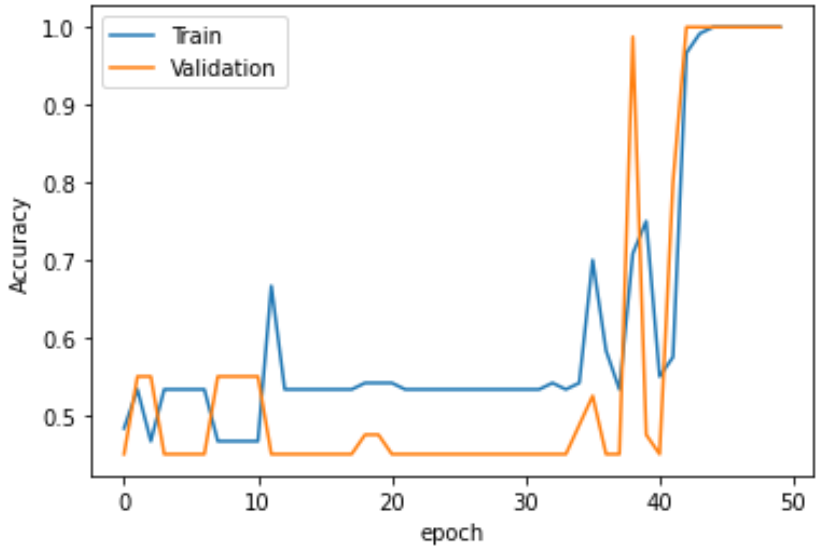


Data preprocessing for DTL

Proposed Memory Forensics for Smart Inverters

DTL Modeling Training and Validation

- The model comprises a total of 132,097 trainable parameters. The model executes and predicts the ransomware and goodware in 1 minute and 44.5 seconds while being run on the Apple M1 architecture.
- The model was trained for 50 epochs, where it converged to over 98% accuracy in the last 10 epoch
- However, the VGG16 basis model shows 91.31% of ransomware detection accuracy for the device-specified ransomware samples, although the basis ransomware detection algorithm provides high detection accuracy for the general ransomware files (> 96%).



Model training and validation results

Outline

1. Introduction
2. Security Threat Modeling for Smart Inverters
3. Proposed Device-Level Memory Forensics for Smart Inverters
4. Experimental Validation
5. Conclusion

Experimental Validation

Ransomware Attack Results

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F   Decoded text
00000000  68 73 71 73 3 timestamp 5C D3 DE 5C 00 00 02 00   hsqs>...\ÓÞ\....
00000010  84 00 00 00 04 00 11 00 C0 00 03 00 04 00 00 00   „.......À.......
00000020  BC 0E F2 24 00 00 00 00 B0 B3 B0 00 00 00 00 00   ¼.ò$....°³°.....
00000030  A8 B3 B0 00 00 00 00 00 FF FF FF F1              F ¨³°.....ÿÿÿÿÿÿÿ
00000040  48 4D B0 00 00 00 00 00 E0 75 B0 0( Image size    0 HM°.....àu°.....
00000050  96 A8 B0 00 00 00 00 00 8A B3 B0 00 00 00 00 00   –¨°.....Š³°.....
```

(a)

```
Offset(h) 00 01 02 03 04 05 06 07 08 09 0A 0B 0C 0D 0E 0F   Decoded text
00000000  68 73 71 73 3 timestamp 84 B6 28 63 00 00 02 00   hsqs?...„¶(c....
00000010  84 00 00 00 04 00 11 00 C0 00 03 00 04 00 00 00   „.......À.......
00000020  DC 0E 02 24 00 00 00 00 DC C0 B0 00 00 00 00 00   Ü..$....ÜÀ°.....
00000030  D4 C0 B0 00 00 00 00 00 FF FF FF F1              F ÔÀ°.....ÿÿÿÿÿÿÿ
00000040  8C 61 B0 00 00 00 00 00 44 89 B0 0( Image size    0 Œa°.....D‰°.....
00000050  82 B9 B0 00 00 00 00 00 B6 C0 B0 00 00 00 00 00   ‚¹°.....¶À°.....
```

(b)

```
003435C0  2A BB ED E9 74 F7 B4 79 50 EB CD 9A 9F F1 D3 39   *»íét÷´yPëÍšŸñÓ9
003435D0  D8 00 24 A2 50 66 07 47 DA ED 5E 36 F9 68 C4 F5   Ø.$¢Pf.GÚí^6ùhÄõ
003435E0  06 34 C0 41 6C 6C 20 6F 66 20 79 6F 75 72 20 66   .4ÀAll of your f
003435F0  69 6C 65 73 20 61 72 65 20 63 75 72 72 65 6E 74   iles are current
00343600  6C 79 20 65 6E 63 72 79 70 74 65 64 20 62 79 20   ly encrypted by
00343610  43 4F 4E 54 49 20 73 74 72 61 69 6E 2E 20 49 66   CONTI strain. If
00343620  20 79 6F 75 20 64 6F 6E 27 74 20 6B 6E 6F 77 20    you don't know
00343630  77 68 6F 20 77 65 20 61 72 65 20 2D 20 6A 75 73   who we are - jus
00343640  74 20 22 47 6F 6F 67 6C 65 20 69 74 2E 22 0D 0A   t "Google it."..
00343650  0D 0A 41 73 20 79 6F 75 20 61 6C 72 65 61 64 79   ..As you already
00343660  20 6B 6E 6F 77 2C 20 61 6C 6C 20 6F 66 20 79 6F    know, all of yo
00343670  75 72 20 64 61 74 61 20 68 61 73 20 62 65 65 6E   ur data has been
00343680  20 65 6E 63 72 79 70 74 65 64 20 62 79 20 6F 75    encrypted by ou
00343690  72 20 73 6F 66 74 77 61 72 65 2E 20 0D 0A 49 74   r software. ..It
003436A0  20 63 61 6E 6E 6F 74 20 62 65 20 72 65 63 6F 76    cannot be recov
003436B0  65 72 65 64 20 62 79 20 61 6E 79 20 6D 65 61 6E   ered by any mean
003436C0  73 20 77 69 74 68 6F 75 74 20 63 6F 6E 74 61 63   s without contac
003436D0  74 69 6E 67 20 6F 75 72 20 74 65 61 6D 20 64 69   ting our team di
003436E0  72 65 63 74 6C 79 2E 0D 0A 0D 0A 44 4F 4E 27 54   rectly.....DON'T
```

(c)

Ransomware Attack Results: (a) original firmware data from the SPI flash memory, (b) compromised firmware data by ransomware, and (c) ransomware attack indication on the compromised firmware

Experimental Validation

Ransomware Detection Results

The trained model is installed in an R-Pi (i.e., security module in a smart inverter). Variants of 4 ransomware and 1 goodware files are additionally created to validate a real-time ransomware attack detection using the proposed method.

```
In [12]: prediction = model.predict(X, batch_size=1)
         print(prediction,'\n')  # will be a list in a list.

         for x in prediction:
             print(class_category[int(x)])

         [[1.]
          [1.]
          [1.]
          [0.]
          [1.]]

         Ransomware
         Ransomware
         Ransomware
         Goodware
         Ransomware
```

Experimental results of ransomware detection

Conclusion

- Provides ransomware threat modeling and proposed ML-based memory forensics method.
- The proposed ML-based memory forensics method can successfully detec malicous ransomware attacks targeting the smart inverters with high accuracy.
- The method can be implemented in the smart inverter with additional security processors enabling device-centric ransomware/malware defense.

Future works include:

1. Proactive detection using prediction of ransomware variants and deobfuscation techniques
2. Ransomware key detection and recovery
3. Developing recovery techniques for smart inverters
4. Cyber-resilience control for mitigating impacts of powerbots implementing automatic detection for the files

Acknowledgments

This work was supported in part by the Department of Energy (DoE) under award No. DE-EE0009026, National Science Foundation (NSF) under award No. EEC-1359414 and CNS2219733.

Thank you!

Question?