End-to-End Deep Neural Networks and Transfer Learning for Automatic Analysis of Nation-State APT
# DEEP INSTINCT IN A NUTSHELL

## CYBER TECHNOLOGY PIONEERS
- Founded in 2015
- The first company to apply deep learning to cyber security
- The only omni-cybersecurity company to offer endpoint & mobile security cross-OS over one unified omni-cyber platform

## STRATEGIC INVESTORS
- NVIDIA
- Coatue

## INTELLECTUAL PROPERTY
- Unique proprietary deep learning framework for cybersecurity
- Protected with 3 registered patents & 3 trademarks

## HIGHLY EXPERIENCE MANAGEMENT
- **Key People**
  - Chairman: Lane Bess
  - CEO: Guy Caspi
  - CTO: Dr Eli David
  - CPO: Nadav Maman
- **90 full time employees**
- **Deep learning scientists**, comprised of PhDs, MScs, computation intelligent AI.
- Cybersecurity veterans - Ex IDF and Israel intelligence Service

## RECOGNITION
- Continuously recognized as one of the *leading deep learning companies* in the world!
- 5 technology innovation awards in short time

## Global Operation
- New York
- Tel Aviv
- Singapore
- Tokyo
- Sydney
• Advanced persistent threat (APT): stealthy and continuous attack targeting a specific entity

• Static vs dynamic analysis: object code analysis (no execution) vs behavior analysis during execution (API calls...)

• Malware family: collection of malwares sharing common code and functionalities (often from the same authors)
• Advanced persistent threats (APT): developed by nation-states

• Nation-state APT attribution classifier: Russia vs China
  • Transfer Learning from APT family classifier

• Input: sandbox reports as raw features

• Classifier: deep neural network (DNN)
Agenda

- Problem definition
- Contributions
- Challenges with classifying APTs
- Proposed solution
- Experimental evaluation and results
Problem Definition

• Nation-State APT authorship attribution

• Why?
  ➢ Act of war

• China vs. Russia
  ➢ Datasets from other nation-states are too small
Nation State APT Families

• Cosmic Duke (Russia)
  ➢ An information stealing malware
  ➢ Used in targeted attacks against government and military targets.
• Net-Traveler (China)
  ➢ A keylogging and information gathering malware
  ➢ Used in targeted attacks against high-profile targets.
• Havex (Russia)
  ➢ a Remote Administration malware (RAT)
  ➢ Used in targeted attacks against industrial sector targets, and has also infected SCADA.
Main Contributions

• First nation-state APT attribution classifier & family classifier

• Largest nation-state developed APT dataset ever collected

• Learning high level feature abstractions using DNN
  • Transfer Learning: APT family classifier -> APT attribution
Challenges Classifying Nation-State APTs

• Multiple nation-state campaigns

• Multiple authors / cyber units

• Evasive APTs

• Small dataset

• False attribution prevention (targeted adversarial attack)
Proposed Solution

• Sandbox reports (behavior of the APT when run dynamically)
• Most common 50K words in reports as raw features
• Boolean input vectors
• 10-layer fully connected DNN with softmax output
  • 50,000- 2,000-1,000-1,000-1,000-1,000-1,000-1,000-500-2
    neurons
• 20% input noise and 50% dropout (hidden layers)

• Transfer learning method:
  1. Train an APT Family Classifier
  2. Replace top later (multi-class families -> binary attribution)
  3. Re-train top layer with new labels
Experimental Evaluation

• Training set: 1,600 files from each class

• Validation set: 200 files taken from the training set

• Test set: 500 files from each class

• Complete separation between APT families in test and train set
  ➢ Generalization is more challenging
  ➢ Results are more applicable to the real world
t-SNE: APT Families vs Attribution

APT Families (t-SNE of top layer)

APT Attribution (t-SNE of top layer)
Results

- 98.6% accuracy on the test set (APT families are different from training set)
  - Each nation-state has different methodologies for developing APTs
- 97.8% accuracy on the test set using transfer learning
  - Feature abstractions for dynamic behavior can be used for attacker profiling
  - More emphasis on a certain injection technique in cyber unit training would make it more prominent in the families
• Nation-state APT attribution is a critical question

• Challenging dataset
  ➢ Small size
  ➢ Multiple authors
  ➢ Attribution prevention techniques

• However, DNN combined with high input noise and high dropout rate leads to good generalization (98.6% accuracy)
Want to Dig Deeper?

• The original paper, presented in ICANN 2017, is available at: https://arxiv.org/abs/1711.09666

• The extended revision discussed here would be published in a special issue of Entropy Journal (ISSN 1099-4300, OpenAccess): "Selected Papers from the 26th International Conference on Artificial Neural Networks - ICANN 2017"
Questions?
Want to join the first company to apply deep learning to cybersecurity?

Contact us at: careers@deepinstinct.com