Tutorial

Title:
AI based Malware Detection

Abstract:
Today computing devices like laptops, mobile phones, smart devices, etc., have penetrated very deep into our modern society and have become an integral part of our daily lives. Currently, more than half of the world's population uses computers/mobile devices for their professional/personal needs. However, these computing devices are targeted by malware designers encouraged by profits/gains associated with the attack. According to a recent report, monetary losses due to cybercrime are expected to reach 10 trillion dollars annually by 2025. The primary role in providing defense against malware attacks is designed and developed by the anti-malware community (researchers and the anti-virus industry). Traditionally anti-viruses are based on the signature, heuristic, and behavior based detection engines. However, these engines are unable to detect next-generation polymorphic and metamorphic malware. Thus researchers have started developing malware detection engines based on machine learning to complement the existing anti-virus engines. However, there are many open research challenges in these models like adversarial robustness, explainability, fairness, etc., which we are going to discuss in detail during the tutorial.

Duration:
2 hours

Motivation -
Often computer/mobile users call everything that disturbs/corrupts their system a VIRUS without being aware of what it means or accomplishes. This tutorial systematically introduces the different malware varieties, their distinctive properties, different methods of analyzing the malware, and their detection techniques.

Expected audience
Senior undergraduate students, postgraduate students, doctoral students, faculty members, and researchers working or interested in the area of malware analysis and detection.

Outline of contents:

Topics to be covered

1. Introduction to Malware
2. A short history of Malware (virus to malware)
3. 1st Generation Malware
4. 2nd Generation Malware
5. Traditional Malware Detection Systems
6. Static Malware Analysis
7. Challenges in Static Analysis
8. Dynamic Malware Analysis
9. Challenges in Dynamic Analysis
10. Malware Detection as a Classification Problem
11. Challenges in AI based Malware Detection Systems
    a. Adversarial Robustness of Malware Detection Systems
    b. Explainability in Malware Detection Systems
    c. Fairness in Malware Detection Systems
    d. Data Challenges and Routing based Malware Classification
12. Alternate Mechanisms for Malware Detection
    a. Advanced Metamorphic Malware Generation
    b. Malware Normalization as a Defense Strategy
13. Open Research Problems and Future Directions

Key references:

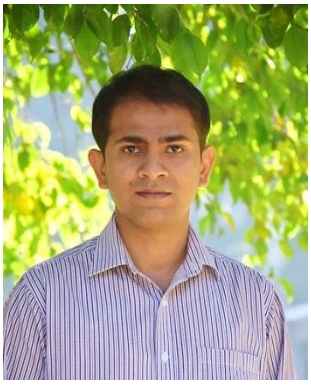
1. Ligh, Michael, Steven Adair, Blake Hartstein, and Matthew Richard. Malware analyst's cookbook and DVD: tools and techniques for fighting malicious code. Wiley Publishing, 2010.
2. Oktavianto, Digit, and Iqbal Muhardianto. Cuckoo malware analysis. Packt Publishing Ltd, 2013.
3. Or-Meir, Ori, Nir Nissim, Yuval Elovici, and Lior Rokach. "Dynamic malware analysis in the modern era—A state of the art survey." ACM Computing Surveys (CSUR) 52, no. 5 (2019): 1-48.
4. Qiu, Junyang, Jun Zhang, Wei Luo, Lei Pan, Surya Nepal, and Yang Xiang. "A survey of android malware detection with deep neural models." ACM Computing Surveys (CSUR) 53, no. 6 (2020): 1-36.
5. Sikorski, Michael, and Andrew Honig. Practical malware analysis: the hands-on guide to dissecting malicious software. no starch press, 2012.
6. Willems, Carsten, Thorsten Holz, and Felix Freiling. "Toward automated dynamic malware analysis using cwsandbox." IEEE Security & Privacy 5, no. 2 (2007): 32-39.
7. Ye, Yanfang, Tao Li, Donald Adjeroh, and S. Sitharama Iyengar. "A survey on malware detection using data mining techniques." ACM Computing Surveys (CSUR) 50, no. 3 (2017): 1-40.

List of speakers:

- **Mohit Sewak, Principal Applied Researcher, Microsoft R&D, India.**

Mohit Sewak is an Artificial Intelligence and Cybersecurity researcher with over 15 years of experience in designing innovative AI software and solutions. Mohit holds more than a dozen patents across the US, India, and worldwide for innovative AI solutions that empower many international products. Mohit is the author of multiple AI book titles on topics including technologies like Deep Reinforcement Learning and Convolutional Neural Networks. Mohit's research is focused on designing AI-based malware and other advanced threat detection and protection systems. Currently, Mohit is serving as a Principal Data Scientist for Security & Compliance Research at Microsoft R&D.

- **Hemant Rathore, Assistant Professor, Department of CS & IS, BITS Pilani, India.**



Hemant Rathore is currently working as Assistant Professor at the Department of CS and IS at BITS Pilani, Goa Campus, India. Before joining academics, he was working in the area of computer security for three years at Symantec, India. His Ph.D. is on the topic of Adversarial Robustness and Explainability in Malware Detection Models. His research interests are in the area of Malware Analysis, Network Security, Machine Learning, and Operating Systems. He has guided several undergraduate and postgraduate students in their independent research projects and published many research papers in reputed journals/conferences.