Chunlin XIONG, Zhenyuan LI, Yan CHEN, Tiantian ZHU, Jian WANG, Hai YANG, Wei RUAN, 2022. Generic, efficient, and effective deobfuscation and semantic-aware attack detection for PowerShell scripts. *Frontiers of Information Technology & Electronic Engineering*, 23(3):361-381. https://doi.org/10.1631/FITEE.2000436

Generic, efficient, and effective deobfuscation and semantic-aware attack detection for PowerShell scripts

Key words: PowerShell; Abstract syntax tree; Obfuscation and deobfuscation; Malicious script detection

Corresponding author: Ruan WEI
E-mail: ruanwei@zju.edu.cn
ORCID: https://orcid.org/0000-0001-8721-4391

Motivation

1. PowerShell is widely used in different tactics in real-world attacks due to its unique features, especially in downloading and executing payloads, establishing reverse shells, and collecting victims' information.

2. However, because the PowerShell language is dynamic by design and can construct script fragments at different levels, state-of-the-art static analysis based PowerShell attack detection approaches are inherently vulnerable to obfuscations.

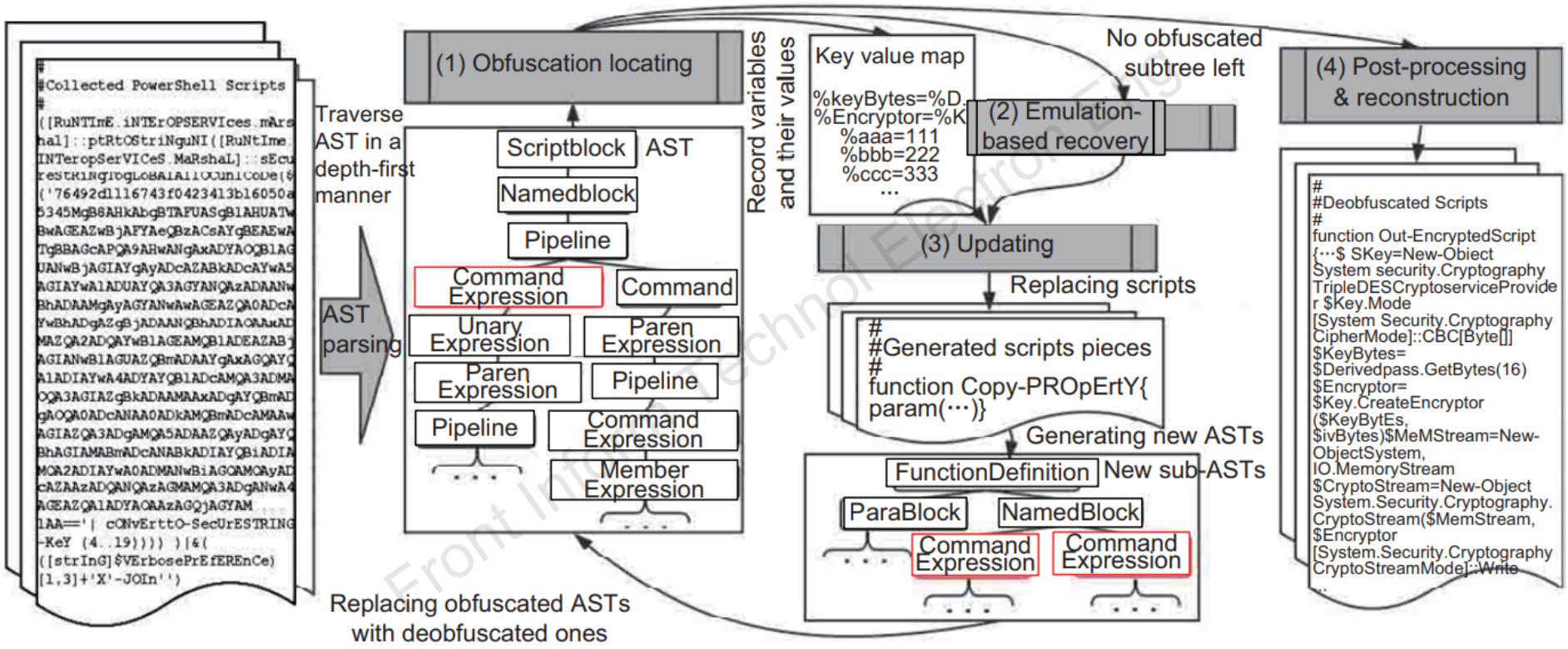
Main idea

1. Analyze the impacts of obfuscation techniques on abstract syntax trees (ASTs). All obfuscation techniques can thus be classified into three categories. Obfuscated PowerShell scripts can be identified based on the features in their AST.

2. Encoding, as one of the obfuscation methods, cannot be used without dynamic execution. To recover the encoded script, the key idea behind our approach is that the obfuscated script fragments must be recovered to the original, unobfuscated scripts before being executed by the Powershell interpreter during runtime.

3. The semantics of the PowerShell scripts can be easily identified in their original form.

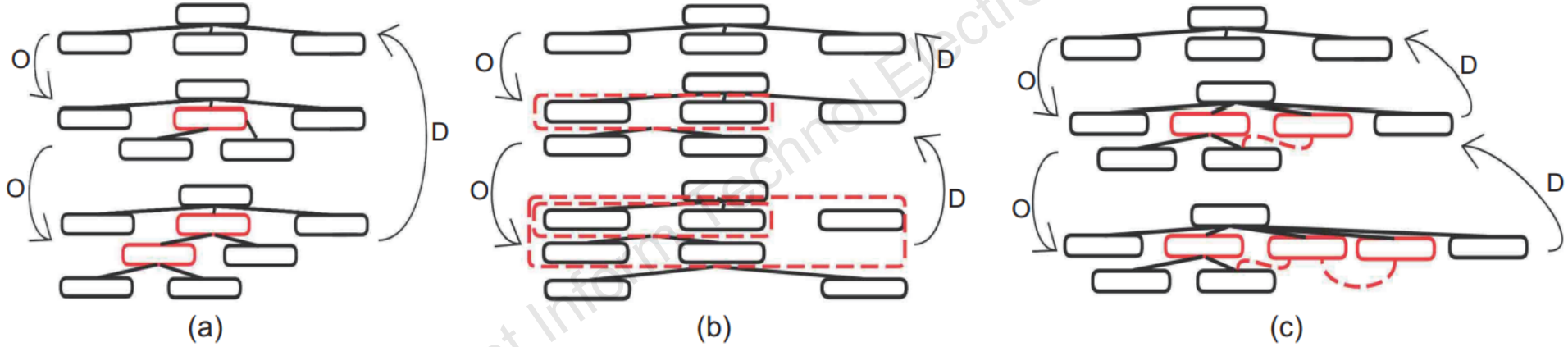
Framework



An overview of the proposed subtree-based deobfuscation for PowerShell scripts

Method

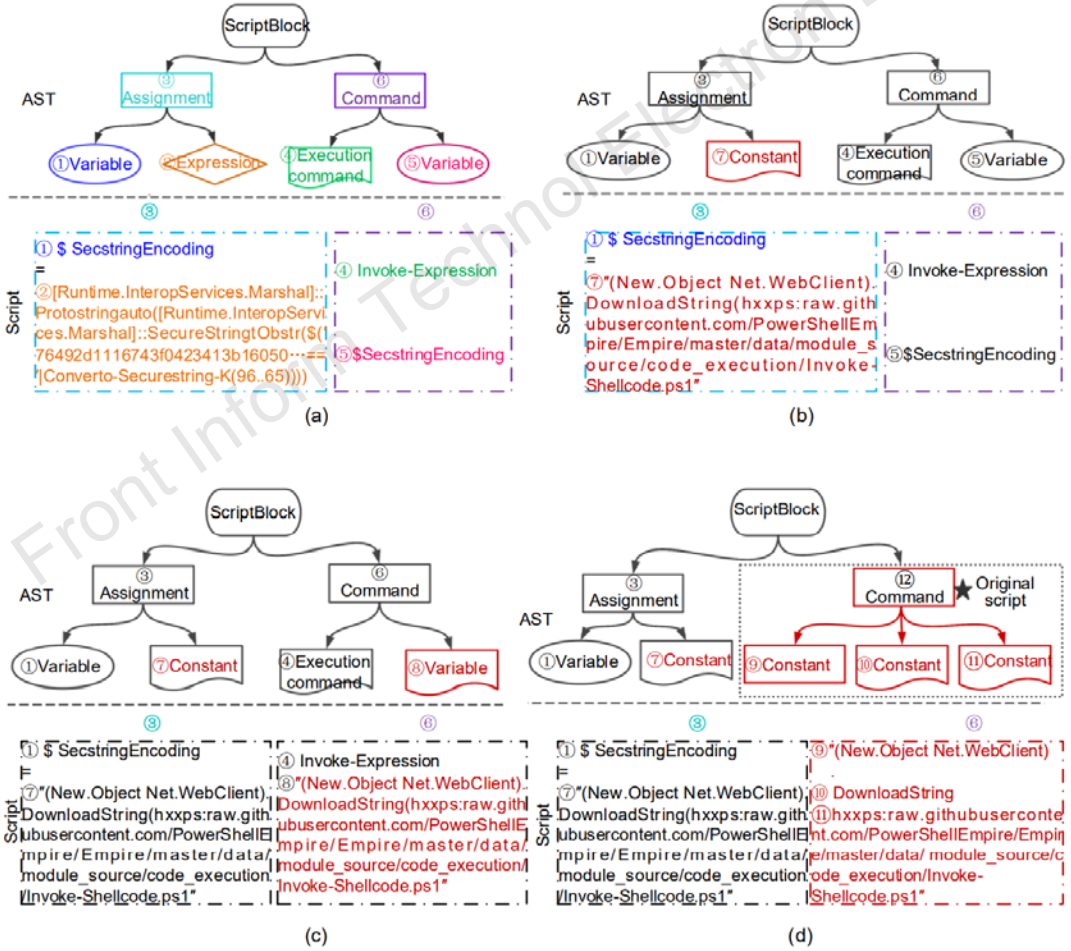
1. Find obfuscated scripts based on the ASTs.



Three kinds of obfuscations based on the impacts on ASTs (O represents obfuscation and D represents deobfuscation)

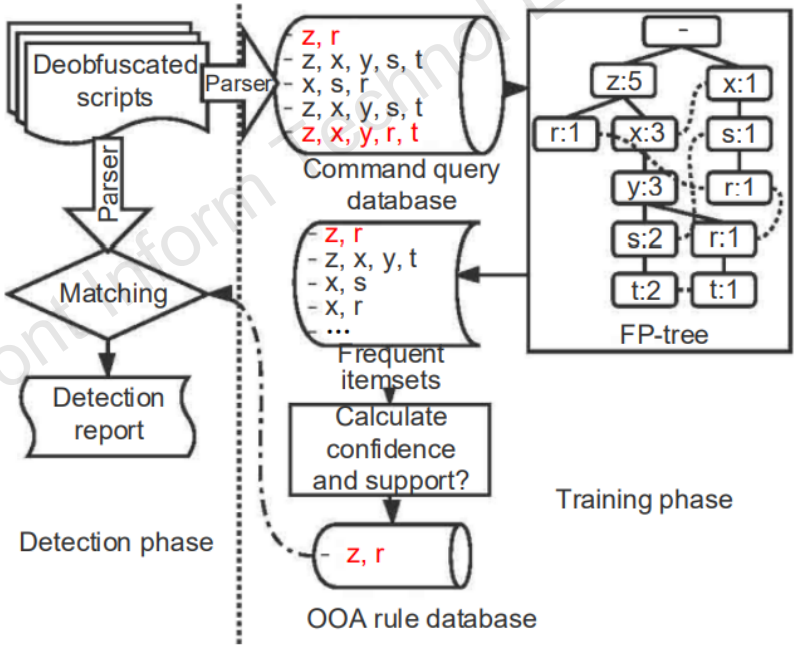
Method

2. An example of the process to deobfuscate a script with emulation-based execution



Method

3. We employ a frequent pattern (FP) growth algorithm based on a frequent pattern tree to generate frequent patterns and a classic classification based on objective-oriented association (OOA) on item sets of commands for detection.



Semantic-aware detection workflow

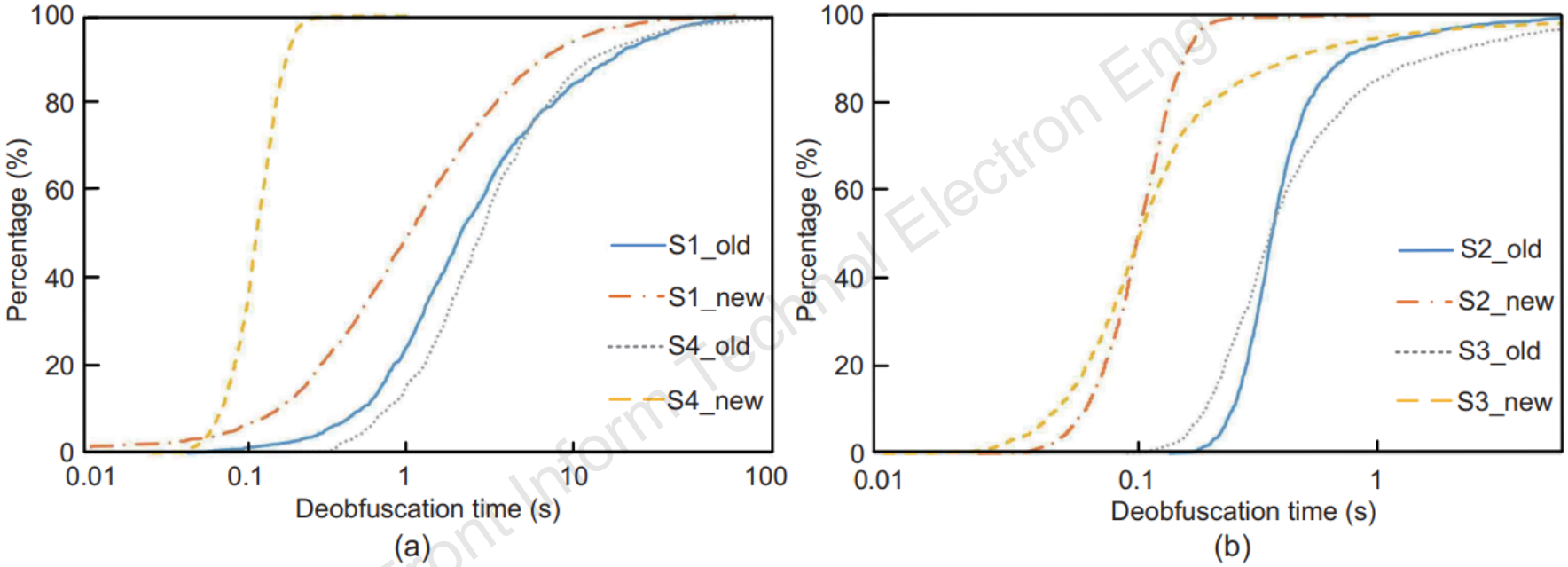
Major results

The average similarities of obfuscated, deobfuscated, and original ASTs

Obf. scheme	Similiarity				
	Obf.	Deobf. (new)*	Deobf. (old)**	Deobf. (PSDEM)	Deobf. (PowerDrive)
S1	1.8%	72.8%	71.5%	70.6%	1.32%
S2	0.1%	100%	79.0%	79.5%	54.2%
S3	0.01%	100%	82.9%	0.01%	84.1%
S4	0.004%	100%	85.2%	0.004%	80.7%
Overall	0.5%	93.2%	79.7%	37.5%	55.1%

Obf.: Obfuscation; Deobf.: Deobfuscation. *This paper; **Li et al. (2019)

Major results



Comparison of deobfuscation efficiency in different schemes: (a) S1 and S4; (b) S2 and S3

The data is represented by the cumulative distribution function (CDF). We limit the maximum time because in some complex cases, the process takes minutes. Results show that the new method has better performance than the old one in all schemes

Conclusions

1. In this paper, we design and implement the first generic, effective, and lightweight deobfuscation approach for PowerShell scripts. To find the obfuscated fragments of scripts accurately, we propose a novel method based on the impact on ASTs. The results show that the new method performs better than the existing methods with almost no false positives, and that it can cover complicated, multilayer, and even unknown obfuscations.

2. To recover the obfuscated scripts, we propose emulation-based recovery at the AST level. This method can recover the scripts perfectly, especially at the script-block level, for which it achieves 100% similarity.

3. We mine 31 newly identified OOA rules automatically with our detection system.

Chunlin XIONG received his BE degree in computer science from Xi’an Jiaotong University, Xi’an, China, in 2015. He is currently a PhD candidate with the College of Cyber Security, Zhejiang University, Hangzhou, China. His research interests include system security, software security, and forensic analysis.

Zhenyuan LI received his BE degree from Xidian University, Xi’an, China, in 2017. He is currently a PhD candidate at Zhejiang University. His research interests include system security, threat detection, and forensic analysis.

Yan CHEN received his PhD degree in computer science from the University of California, Berkeley, CA, USA, in 2003. He is a Professor with the Department of Electrical Engineering and Computer Science, Northwestern University, Evanston, IL, USA. Based on Google Scholar, his papers have been cited over 10000 times and his *h*-index is 49. He won the Department of Energy (DoE) Early CAREER Award in 2005, the Department of Defense (DoD) Young Investigator Award in 2007, and the Microsoft Trustworthy Computing Awards in 2004 and 2005 with his colleagues. His research interests include network security, measurement, and diagnosis for large-scale networks and distributed systems.

Tiantian ZHU received his PhD degree in computer science from Zhejiang University, Hangzhou, China, in 2019. He is currently a lecturer with the College of Computer Science and Technology, Zhejiang University of Technology, China. His research interests include mobile security, system security, and artificial intelligence.

Jian WANG is a senior student at Harbin Engineering University, majoring in information security.

Hai YANG received his ME degree in computer science from Zhejiang University in 1998. Before joining MagicShield, he worked on wireless communications and networking infrastructure for Intel, Broadcom, Skyworks, etc.

Wei RUAN is a professorate senior engineer. After graduating from Shanghai Jiao Tong University in 1991, he received MS and PhD degrees from the Department of Energy, Zhejiang University in 1997 and 2000, respectively. He is currently a teacher with the College of Control Science and Engineering, Zhejiang University, serving as director of the Equipment Automation Center of the Advanced Technology Research Institute of Zhejiang University at the same time.