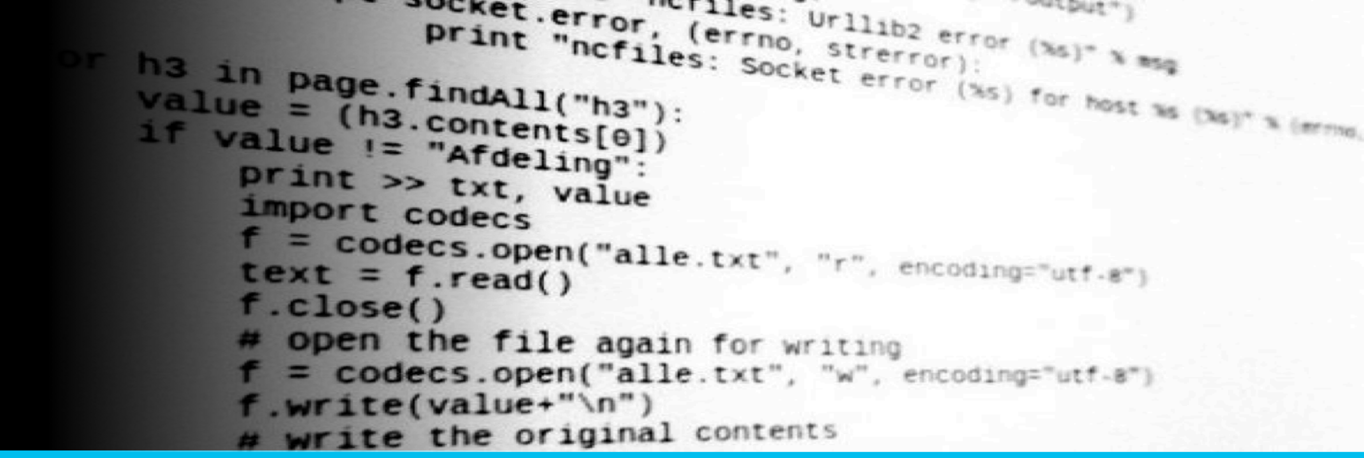
Translingual Obfuscation

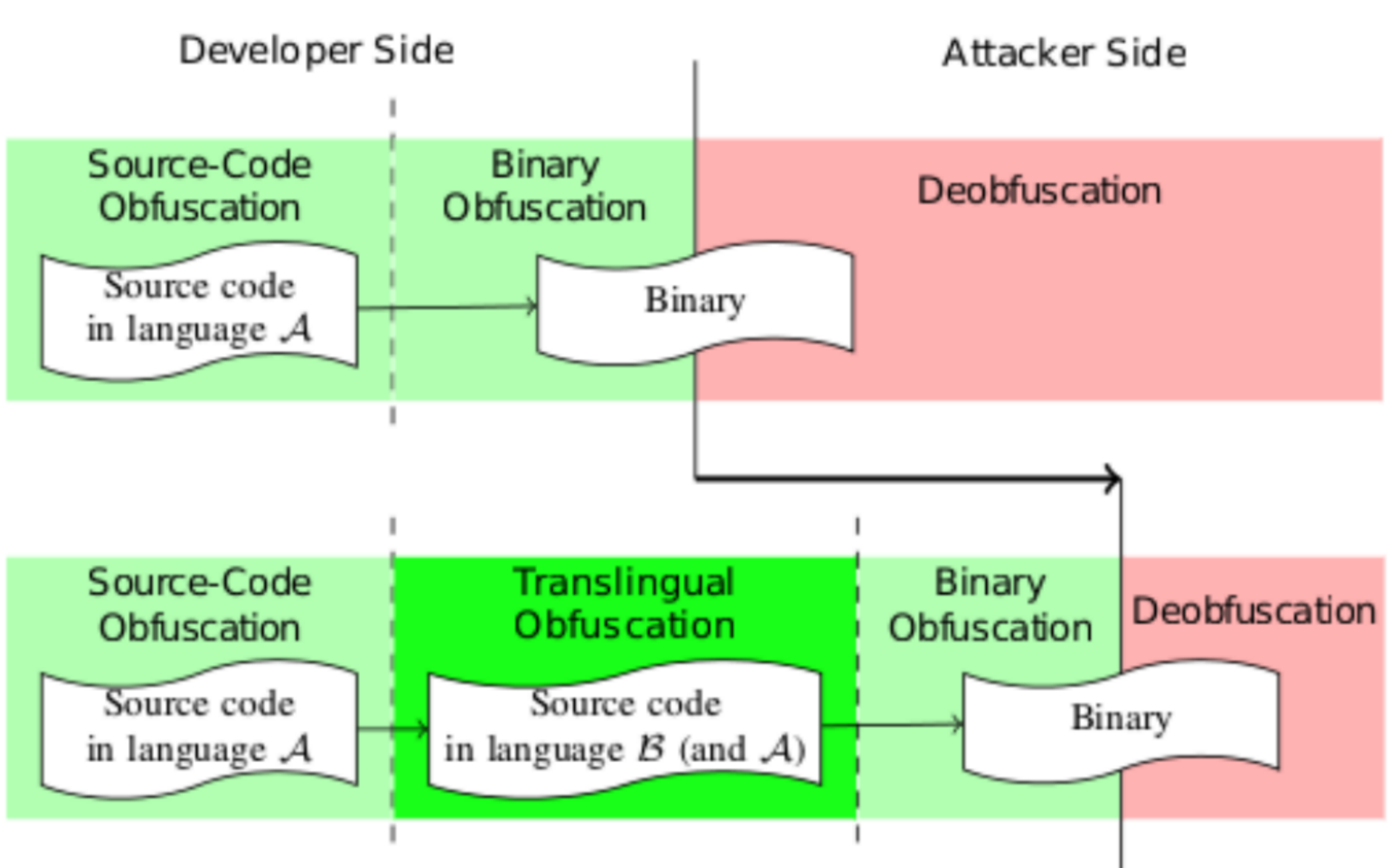


Pei Wang, Shuai Wang, Jiang Ming, Yufei Jiang, and Dinghao Wu, College of Information Sciences and Technology, Pennsylvania State University

Introduction

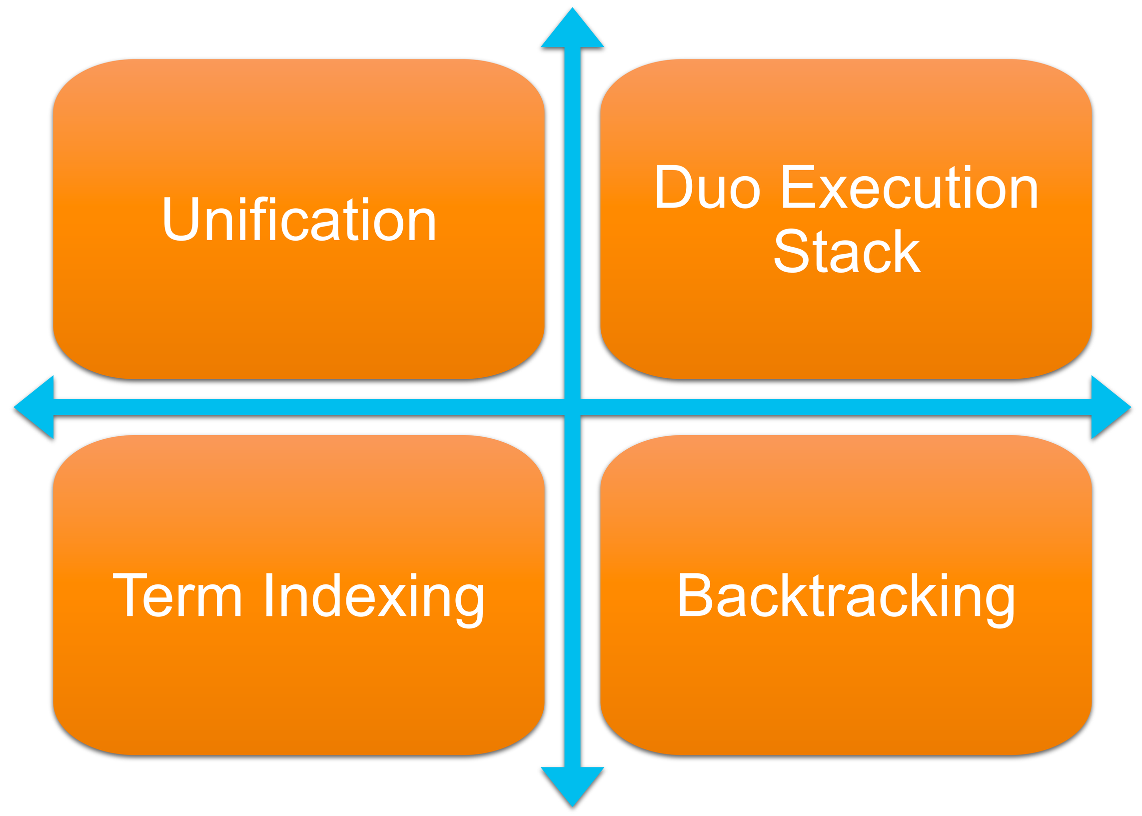
Program obfuscation is an important software protection technique that prevents attackers from revealing the programming logic and design of the software. We introduce *translingual obfuscation*, a new software obfuscation scheme which makes programs obscure by "misusing" the unique features of certain programming languages. Translingual obfuscation translates part of a program from its original language to another language that has a different programming paradigm and execution model, thus increasing program complexity and impeding reverse engineering. In this research, we present the feasibility and effectiveness of translingual obfuscation with Prolog, a logic programming language.

Method Overview

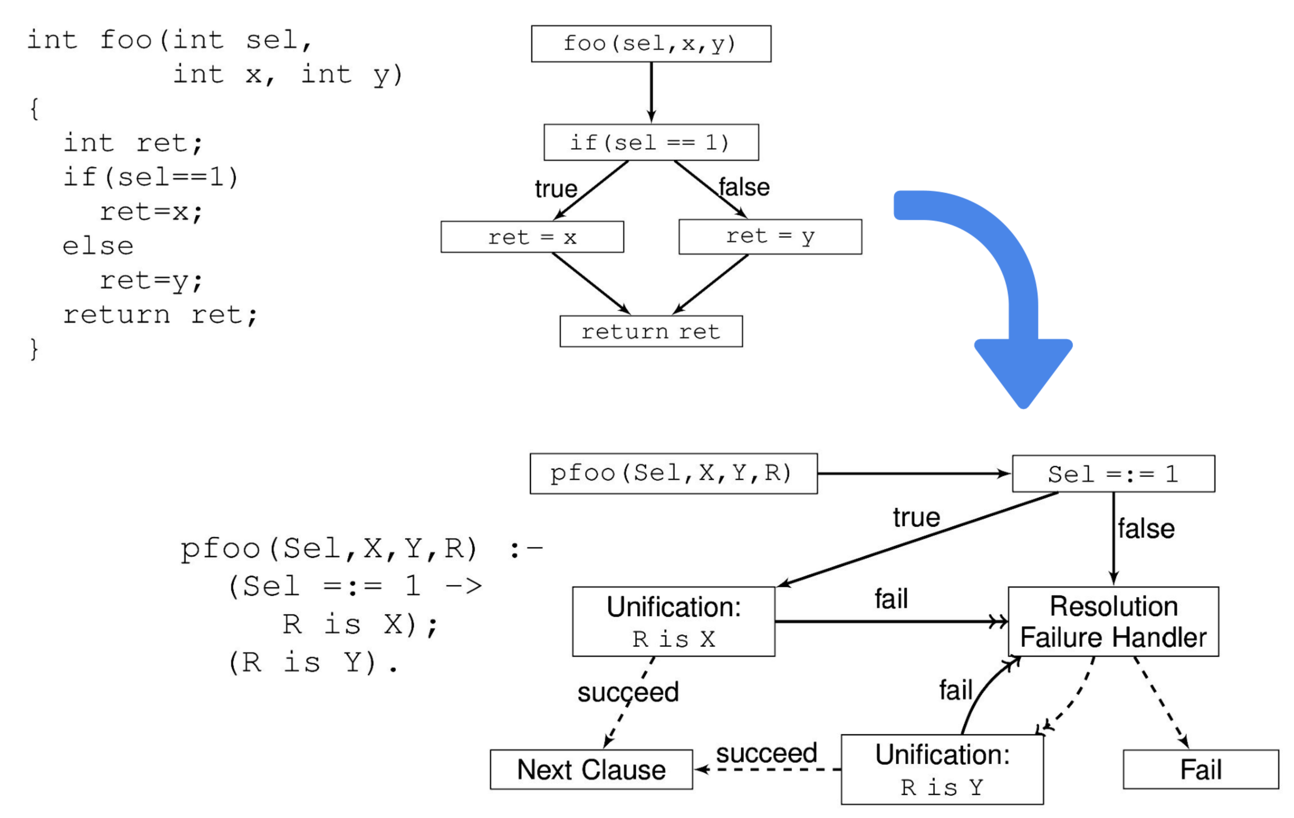


* Supported by NSF Grant No. CNS-1223710 and CCF-1320605.

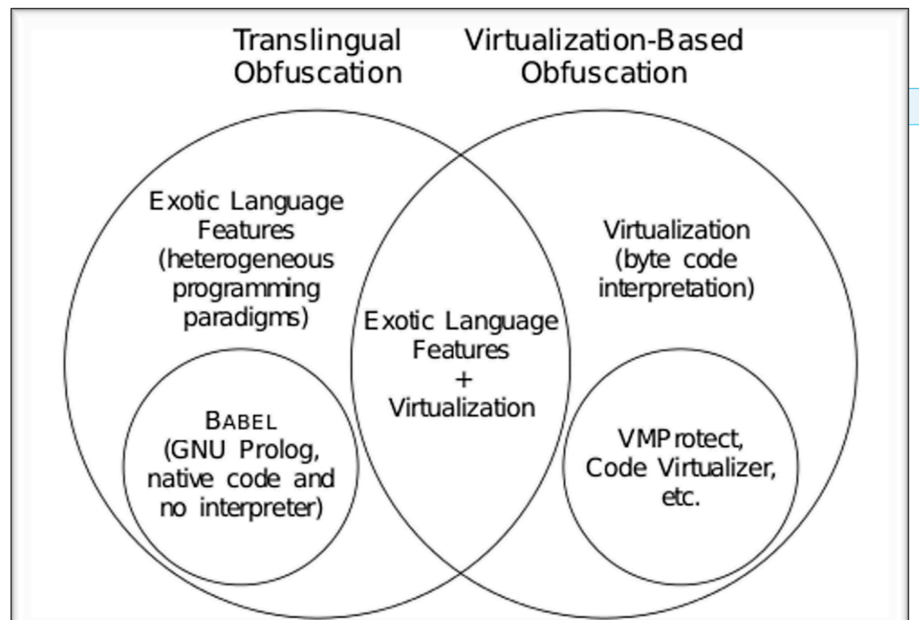
Obfuscating Factors of Prolog



Showcase



V.S. Virtualization-Based Obfuscation



A new direction of software obfuscation

Low Cost

Optional Virtualization

Complementary to other obfuscation techniques

Evaluation Methodology

Obfuscate 6 real-world C programs

- bzip2, mcf, regexp, svm-light, oftpd, monggose

Evaluate in 4 aspects (Collberg et al., 1998)

- Potency: Amount of additional complexity added by obfuscation; measured by software engineering metrics
- Resilience: Security against automated reverse engineering; measured by competing the state-of-the-art binary differs *CoP (Luo et al., 2014)* and *BinDiff*
- Cost: Performance penalty caused by obfuscation; competing with a popular commercial obfuscator (*Code Virtualizer*)
- Stealthy: Difficulty of detecting the presence of obfuscation; measured by instruction distribution

Experiment Results

Potency	• Up to 70x increase on software complexity
Resilience	• *CoP* reported similarity median below 20%; *BinDiff* failed to match over half of the protected functions with their pre-obfuscated versions
Cost	• Overwhelm *Code Virtualizer* in all tests; produce correct binaries in cases where *Code Virtualizer* fails
Stealth	• Instruction distributions fall within normal ranges sampled from SPEC2006

Conclusion

Translingual obfuscation is a new software obfuscation scheme based on translations from one programming language to another. By utilizing certain design and implementation features of the target language, we are able to protect the original program against reverse engineering. The experimental results show that translingual obfuscation is an adequate and practical software protection technique.