Abstract:

Analyzing a binary program is generally viewed as a difficult task. This is because much information has been lost during compilation, especially when optimization techniques are applied. The bar for analyzing binaries is so high that many people believe it is safe to put secrets in binary programs, as few people can analyze the binaries and obtain the secrets. Driven by this misconception, programmers either put secret data in source code, believing no one will be able to read them, or invest little in the security of their programs, believing no one can analyze the binary and discover vulnerabilities. Nonetheless, in the past decade, the technical progress in the field of binary analysis has continuously lowered the bar. Nowadays, attackers can understand binary programs, recover secrets, and discover vulnerabilities faster than ever. The obscurity of binary programs, if it ever existed, is fading away, and thus yields a great impact in software security of both newly developed and legacy binary programs.

In this talk, I will present some interesting research in the past few years that facilitates the reverse engineering of binary programs. First, dynamic analysis, especially dynamic binary instrumentation (DBI), is a powerful technique for monitoring behaviors of a binary program while treating it as a gray box or a black box. As an application, I will show that many software-based DRM solutions can be attacked at a low cost. Second, although static analysis techniques are usually seen as inaccurate, combining them with carefully selected domain-specific heuristics and more expensive techniques, like symbolic execution, yields much more accurate output. On this aspect, I will present our work of binary reassembling using static analysis, which is a key improvement for binary patching, hardening, and binary code reuse. Finally, I will present angr, a popular, flexible, and user-friendly binary analysis platform, and how it can foster doing good research. With angr, our research has advanced the state of the art of binary analysis, showing that it is feasible to solve many tasks that were once believed to be extremely difficult.