Softwareprojekt: Iridium

ein Analyseframework für kompilierte Programme

Inokentiy Babushkin

Abteigymnasium Brauweiler

06.05.2015 / Colloquium zur Besonderen Lernleistung

Inhalt

Outline

Erweiterbarkeit

- Modularer Aufbau
- Sinnvolles API-Design
- Multiple Eingabeformate
- Lesbarkeitsorientierter Code

Freie Software

- Verwendung der GPLv3
- Codequalität und Funktionsumfang durch Community unterstützt

Outline

Definition hier: Wiederherstellung von Hochsprachencode aus kompilierten Programmen

- Komplexer und zeitaufändiger Prozess, da
  - Software zunehmend komplex
  - Compiler den Code optimisieren
- sehr nützlich
  - Malware-Analyse
  - Wiederherstellung von altem Programmcode

(Halb-)Automatisierte Ansätze

- Decompiler
- Analyseframeworks

Outline

Analyse anhand von Kontrollflussgraphen

- Schrittbasierter Reduktionsalgorithmus für einzelne Funktionen
- Unterscheidung aller wichtigen Hochsprachen-Programmelemente
- Strukturierte Darstellung

Outline

Heuristiken zur Untersuchung lokaler Variablen

- Untersuchung des Stackframes der jeweiligen Funktion
- (Teilweise) Erkennung von Arrays, Datentypen und Pointern

Outline

Optimisierte Integerdivision

- Von modernen Compiler bei konstantem Divisor angewendet
- Teure Divisions-Instruktion wird durch eine Reihe von Additionen, Shifts und Multiplikationen ersetzt
- Für Menschen keine Rückschlüsse auf den Divisor möglich, deswegen Implementation eines entsprechenden Algorithmus

Zusammenfassung

- Reverse-Engineering und Analyse kompilierter Programme sind häufig notwendig, allerdings meist auch schwieig und mit enormem Aufwand verbunden.
- Gleichzeitig sind die automatisierten Lösungsansätze für diese Aufgaben nicht oder nur kaum für interaktive Arbeit geeignet, weswegen mit dem vorliegenden Projekt ein entsprechender Prototyp vorgelegt wird.
- Noch mehr?

- Ausblick
  - Stabilisierung des Codes
  - Erweiterung um Eingabemodule
  - Weitere Features