

1. Combining static analysis with probabilistic models to enable market-scale android inter-component analysis (Open Access)

Octeau, Damien (1, 2); Jha, Somesh (1, 3); Dering, Matthew (2); McDaniel, Patrick (2); Bartel, Alexandre (4); Li, Li (5); Klein, Jacques (5); Le Traon, Yves (5)

Source: *ACM SIGPLAN Notices*, v 51, n 1, p 469-484, April 8, 2016; **ISSN:** 15232867; **DOI:**

10.1145/2837614.2837661; **Publisher:** Association for Computing Machinery

Author affiliation: (1) Department of Computer Sciences, University of Wisconsin, United States (2) Department of Computer Science and Engineering, Pennsylvania State University, United States (3) IMDEA Software Institute, Spain (4) EC SPRIDE, Technische Universität Darmstadt, Germany (5) Interdisciplinary Centre for Security, Reliability and Trust, University of Luxembourg, Luxembourg

Abstract: Static analysis has been successfully used in many areas, from verifying mission-critical software to malware detection. Unfortunately, static analysis often produces false positives, which require significant manual effort to resolve. In this paper, we show how to overlay a probabilistic model, trained using domain knowledge, on top of static analysis results, in order to triage static analysis results. We apply this idea to analyzing mobile applications. Android application components can communicate with each other, both within single applications and between different applications. Unfortunately, techniques to statically infer Inter-Component Communication (ICC) yield many potential inter-component and inter-Application links, most of which are false positives. At large scales, scrutinizing all potential links is simply not feasible. We therefore overlay a probabilistic model of ICC on top of static analysis results. Since computing the inter-component links is a prerequisite to inter-component analysis, we introduce a formalism for inferring ICC links based on set constraints. We design an efficient algorithm for performing link resolution. We compute all potential links in a corpus of 11,267 applications in 30 minutes and triage them using our probabilistic approach. We find that over 95.1% of all 636 million potential links are associated with probability values below 0.01 and are thus likely unfeasible links. Thus, it is possible to consider only a small subset of all links without significant loss of information. This work is the first significant step in making static inter-Application analysis more tractable, even at large scales. © 2016 ACM. (51 refs)

Main heading: Static analysis

Controlled terms: Android (operating system)

Uncontrolled terms: Android applications - Android ICC - Application analysis - Mission critical softwares - Probabilistic approaches - Probabilistic modeling - Probabilistic models - Probabilistic programs

Classification Code: 723 Computer Software, Data Handling and Applications - 723.5 Computer Applications

Open Access type(s): All Open Access, Bronze, Green

Database: Compendex

Data Provider: Engineering Village

Compilation and indexing terms, Copyright 2022 Elsevier Inc.