《Pinpoint: Fast and Precise Sparse Value Flow Analysis for Million Lines of Code》Review

论文信息

Pinpoint: Fast and Precise Sparse Value Flow Analysis for Million Lines of Code 作者：Qingkai Shi, Xiao Xiao, Charles Zhang

解决的问题

Path sensitive analysis vs path non-sensitive: 是否考虑判断语句 interprocedure vs intraprocedure analysis：是否过程间分析 Flow sensitive analysis vs flow non-sensitive analysis：是否考虑语句的先后顺序 Incremental Analysis：增量分析的技术，只分析修改过的部分，减少重复分析 在多数情形下拿不到源代码，只有二进制代码可供分析。

权衡效率（Performance）、召回率（Recall）以及精度（Precision）之间的关系 Pinpoint特点： 中间语言的设计可以进行跨语言分析 支持扩展和自定义检查器

技术概要

传统密集程序分析方法 1、沿着指令执行的顺序访问每条指令 2、在每个指令处储存程序的状态信息

稀疏程序分析 1、每条语句只会更改一小部分量的值 2、变量的值只会被很少的后续语句使用 3、便于进行定理证明，按照实际需求进行路径约束的可满足性求解 稀疏程序分析的关键：获取**数据依赖图**，需要进行指针分析 【指针分析===>建立数据依赖图==>值传递分析缺陷检测】

pinpoint中的技术特点： 稀疏程序分析 独立的分析减少，只预生成程序依赖图 指针分析等专门分析只进行一次，且按照实际需求进行 自下而上：先分析被调用函数 主动求解：使用特殊的线性求解器过滤判定条件（降低分析基数），寻找自相矛盾子句 纯化函数，对参数和返回值进行符号化

学术版获取

发邮件 qshiaa@cse.ust.hk