Leveraging Program Analysis to Reduce User-Perceived Latency in Mobile Applications

PALOMA: Program Analysis for Latency Optimization of Mobile Apps
Yixue Zhao (yixue.zhao@usc.edu), Marcelo Schmitt Laser, Yingjun Lyu, Nenad Medvidovic

Motivation

Wireless Network Characteristics:
- Low bandwidth
- Unstable connectivity
- Hard to deploy

Key Insights

Solution: Prefetch can achieve "zero" latency!
Wait… But "how" to Prefetch?
- App code has useful information on what & when to prefetch
- "User think time" provides prefetch opportunities

Current State of the Art

Existing Limitations:
- Server-based
- Human-based
- History-based
- Domain-based

PAOMA:
- Client-only
- Automatic
- No historical information
- Domain-independent

PALOMA Overview

Microbenchmark Evaluation

Table Legend:
SD, TP, FFP denote the runtimes of the three PALOMA instrumentation methods.
Orig is the time required to run the original app.
Red/OH represents the reduction/overhead in execution time when applying PALOMA.

Real App Evaluation

Road Ahead

References