LLVM Miner: Text Analytics based Static Knowledge Extractor

Hameeza Ahmed, Dr. Muhammad Ali Ismail
NED University of Engineering & Technology, Pakistan

Introduction

• LLVM miner simplifies static IR level analysis in LLVM compiler tool
• Performs text analytics to extract related information from IR code
• IR is passed through the miner for extracting hidden features

Testing Environment

• Bfs, Connected Components, Grep, & Histogram applications
• RStudio tool (R Scripts)
• Instruction frequency & application trend metrics

Results

• Instruction frequency shows count of each instruction by bar graph & word cloud
• Application trend shows proportion of classes of operations using bar graphs
• Determines if application is compute bound, or memory bound or I/O bound
• Clusters instructions in categories like branch, compute, function calls, IO read write, memory consumption, & memory read write operations

Benefits

• Promising direction for studying code hidden features
• Easier & less costly solution in terms of time & efforts as compared to analysis passes

Contact: hameeza@neduet.edu.pk, maismail@neduet.edu.pk