ISAMon: A Tool for Cross-Platform Performance Estimation and Clone Detection Using a Novel FFT-based Code Signature

Shubhankar S. Singh and Dr. Smruti R. Sarangi, Indian Institute of Technology Delhi
Biography

• Shubhankar Suman Singh
  • Ph.D. in CSE (2016 – present)
  • IIT Delhi

• Dr. Smruti Ranjan Sarangi
  • Associate Professor in CSE & EE
  • IIT Delhi
Overview

Many ISAs

• x86 (CISC)
• RISC-V, ARM (RISC)

Why compare ISAs?

• Performance
• Energy
• Create a multi-ISA architecture

Performance difference:
0.6 to 1.5
ISAMon Tool

- Application
Code Signature

- Convert sequences of instructions into a raster diagram.
- Apply the Fourier transform on the raster image to obtain a horizontal and a vertical stride.
Code signature is a 5-tuple (H, V, C, P, I)

- H: Horizontal stride of the raster image (loop size)
- V: Vertical stride of the raster image (loop count)
- C: Color composition (histogram of number of instructions for each color)
- P: Phase type (most frequent instruction type)
- I: Instruction set architecture (ARM, RISC-V, x86)
ISA Compare: *deepsjeng* (SPEC 17)

- **ARM**: 1
- **RISC-V**: 0.89
- **x86**: 0.78

### Code signature

<table>
<thead>
<tr>
<th></th>
<th>ARM</th>
<th>RISC-V</th>
<th>x86</th>
</tr>
</thead>
<tbody>
<tr>
<td>Horizontal stride</td>
<td>5</td>
<td>6</td>
<td>7</td>
</tr>
</tbody>
</table>

Smaller loop size leads to a better performance.
Clone Detection

- Cross architecture clone detection for x86, ARM and RISC-V

<table>
<thead>
<tr>
<th>Tools</th>
<th>Accuracy</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAMon</td>
<td>84 %</td>
</tr>
<tr>
<td>kam1no [1]</td>
<td>46 %</td>
</tr>
<tr>
<td>BinDiff [2]</td>
<td>51 %</td>
</tr>
</tbody>
</table>

Confusion matrix
Performance Estimation

- Cross architecture performance estimation for x86, ARM and RISC-V

**Algorithm**: piece-wise linear regression for estimating performance using code signature

- Compared against LACross and SDC

<table>
<thead>
<tr>
<th>Tools</th>
<th>Error</th>
</tr>
</thead>
<tbody>
<tr>
<td>ISAMon</td>
<td>2.8 %</td>
</tr>
<tr>
<td>LACross [3]</td>
<td>3.5 %</td>
</tr>
<tr>
<td>SDC [4]</td>
<td>3 %</td>
</tr>
</tbody>
</table>
Conclusion

• We proposed a novel FFT based code signature
• We proposed a visualization technique for a piece of code based on observed dynamic instructions.
• We used the code signature for comparing ISAs, detecting clones and estimating performance.
• Our algorithms are based on simulation techniques and hence we do not require the real hardware.
• The prediction phase of our algorithm is 5 times faster as compared to previous work.
References


Questions